Finite Relational Semantics for Language Kleene Algebra with Complement

Yoshiki Nakamura
Institute of Science Tokyo, Japan

Abstract

We study the equational theory of Kleene algebra (KA) w.r.t. languages (here, meaning the equational theory of regular expressions where each letter maps to any language) by extending the algebraic signature with the language complement. This extension significantly enhances the expressive power of KA. In this paper, we present a *finite relational semantics* completely characterizing the equational theory w.r.t. languages, which extends the relational characterizations known for KA and for KA with top. Based on this relational semantics, we show that the equational theory w.r.t. languages is Π_1^0-complete for KA with complement (with or without Kleene-star) and is PSPACE-complete if the complement only applies to variables or constants.

2012 ACM Subject Classification Theory of computation → Equational logic and rewriting; Theory of computation → Logic and verification; Theory of computation → Formal languages and automata theory

Keywords and phrases Kleene algebra, Language model, Relational model, Complexity

Digital Object Identifier 10.4230/LIPIcs.CSL.2025.37

Related Version *Full Version*: `https://hal.science/hal-04455882` [35]

Funding This work was supported by JSPS KAKENHI Grant Number JP21K13828.

Acknowledgements We thank anonymous reviewers for useful comments.

1 Introduction

Kleene algebra (KA) [24, 11, 25] is an algebraic system for regular expressions consisting of identity (1), empty (0), composition ($;$), union ($+$), and iteration ($\_^*$). As iteration frequently appears in computer science, KA has many applications, e.g., the semantics of programs [46], relation algebra [40], graph query language [12, 21], program verification [29, 23, 48], and program logics [26, 41, 53]. In practice, we often consider extensions of KA. One direction of extensions is to extend equations to formulas, e.g., Horn formulas ($t_1 = s_1 \to \cdots \to t_n = s_n \to t = s$) for considering hypotheses [9, 28, 14, 44]. Another direction is to extend terms by adding some operators. For example, Kleene algebra with tests (KAT) applies to model Hoare logic [26] and KAT with top ($\top$) applies to model incorrectness logic [41, 53, 45]. It is also natural to extend KA with language operators, e.g., reverse [3], residual [8], intersection ($\cap$) [2], top (universality) [53, 45], variable complements ($\overline{x}$) [38, 39], and combinations of some of them [4, 5]. Note that, whereas the class of regular languages is closed under these operators, such extensions strictly enhance the expressive power of KA w.r.t. languages (here, meaning regular expressions where each letter maps to any language); see [38, 39] and Section 2.2 for complement.

In this paper, we study KA w.r.t. languages by extending the algebraic signature with the *language complement* ($\_^-$). Extending with complement and considering its fragments is a natural, comprehensive approach, e.g., in logic, formal language [10, 42], and relation

33rd EACSL Annual Conference on Computer Science Logic (CSL 2025).
Editors: Jörg Endrullis and Sylvain Schmitz; Article No. 37; pp. 37:1–37:23
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

algebra [50, 40] (see also [1, 6, 32, 43, 34]). The language complement[1] in KA w.r.t. languages significantly enhances the expressive power. For instance, we can define $\top$ and $\cap$ using complement: $\top = 0^-$ and $t \cap s = (t^- + s^-)^-$. Additionally, we can encode positive quantifier-free formulas by equations of KA terms with complement (Remark 3.3 and Section A).

Our main contribution is to present *a finite relational semantics* for KA with complement w.r.t. languages: relational subword models RSUB (Section 3). As KA with complement has a high expressive power, our relational semantics can apply to a more broad class of extensions of KA (including KA with $\top$ and $\cap$) than known relational semantics, e.g., REL (for KA) [46, third page] and GREL (for KA with $\top$) [53, 45] (see Remark 3.5). A good point of RSUB is its form; each model is finite and totally ordered (with minimal and maximal vertices). For instance, the Π^0_1 upper bound result of the equational theory of KA with complement w.r.t. languages is immediate from the finiteness of RSUB. Another good point is that we can naturally consider lifting techniques known in REL to LANG. For instance, by using the techniques in [34] w.r.t. REL, we can show the following complexity results: the equational theory w.r.t. languages is Π^0_1-complete for KA with intersection and variable complements (Theorem 4.12) and for KA with complement and without Kleene-star (i.e., star-free regular expressions w.r.t. LANG) (Theorem 4.15); and PSPACE-complete for KA with variable and constant complements (Theorem 6.10). The PSPACE decidability result above positively settles the open problem posed in [38, Sect. 7].

This paper is structured as follows. In Section 2, we give basic definitions, including language models (LANG) and generalized relational models (GREL). In Section 3, we introduce RSUB (a subclass of GREL) and show that the equational theory w.r.t. LANG coincides with that w.r.t. RSUB. In Section 4, by using RSUB, we give a reduction from the quantifier-free theory w.r.t. LANG into the equational theory w.r.t. LANG. Using this reduction, we show that the equational theory w.r.t. LANG is Π^0_1-complete for KA with complement (moreover, for KA with intersection and variable complements and for KA with complement and without Kleene-star). In Section 5, by using RSUB, we give a graph characterization for KA terms with variable and constant complements. In Section 6, by using this characterization, we show that the equational theory for KA terms with variable and constant complements is PSPACE-complete. In Section 7, we conclude this paper.

2 Preliminaries

We write $\mathbb{N}$ for the set of non-negative integers. For $l, r \in \mathbb{N}$, we write $[l, r]$ for the set $\{i \in \mathbb{N} \mid l \leq i \leq r\}$. For a set X, we write $\wp(X)$ for the power set of X.

For a set X (of letters), we write X^* for the set of words over X. A *language* over X is a subset of X^*. We use w, v to denote words and use L, K to denote languages, respectively. We write $\|w\|$ for the *length* of a word w. We write ε for the empty word. We write wv for the concatenation of words w and v. For languages $L, K \subseteq X^*$, the concatenation $L \mathbin{;} K$ and the Kleene-star L^* is defined by:

$$L \mathbin{;} K \triangleq \{wv \mid w \in L \ \wedge \ v \in K\}, \qquad L^* \triangleq \bigcup_{n \geq 0} \{\varepsilon\} \mathbin{;} \underbrace{L \mathbin{;} \cdots \mathbin{;} L}_{n \text{ times}}.$$

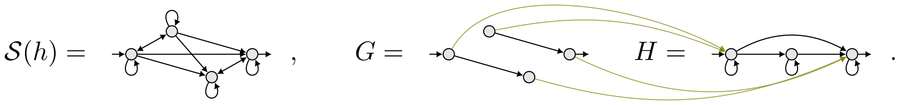
A *(2-pointed) graph* G over a set A is a tuple $\langle |G|, \{a^G\}_{a \in A}, 1^G, 2^G \rangle$, where $|G|$ is a non-empty set (of vertices), each $a^G \subseteq |G|^2$ is a binary relation, and $1^G, 2^G \in |G|$ are vertices. Let G, H be graphs over a set A. For a map $f \colon |G| \to |H|$, we say that f is a *graph*

[1] KAT [29] is also an extension of KA with complement, but this complement is not the language complement.

homomorphism from G to H, written $f\colon G \longrightarrow H$, if for all x, y, and a, $\langle x, y\rangle \in a^G$ implies $\langle f(x), f(y)\rangle \in a^H$, $f(1^G) = 1^H$, and $f(2^G) = 2^H$. We say that f is a *graph isomorphism* from G to H if f is a bijective graph homomorphism and for all x, y, and a, $\langle x, y\rangle \in a^G$ iff $\langle f(x), f(y)\rangle \in a^H$. We say that H is a (canonical) *edge-extension* of G if $|H| = |G|$ and the identity map is a graph homomorphism from G to H. For a set $\{1^G, 2^G\} \subseteq X \subseteq |G|$, the *induced subgraph* of G on X is the graph $\langle X, \{a^G \cap X^2\}_{a \in A}, 1^G, 2^G\rangle$. For an equivalence relation E on $|G|$, the *quotient graph* of G w.r.t. E is the graph $G/E \triangleq \langle |G|/E, \{\{\langle X, Y\rangle \mid \exists x \in X, y \in Y, \langle x, y\rangle \in a^G\}_{a \in A}, [1^G]_E, [2^G]_E\rangle$ where X/E denotes the set of equivalence classes of X by E and $[x]_E$ denotes the equivalence class of x. Additionally, we use the following operation:

▶ **Definition 2.1.** *For a graph homomorphism $h\colon G \longrightarrow H$ where G, H are graphs over a set A, the* edge-saturation *of G w.r.t. h is the graph $\mathcal{S}(h) \triangleq \langle |G|, \{\{\langle x, y\rangle \in |G|^2 \mid \langle h(x), h(y)\rangle \in a^H\}\}_{a \in A}, 1^G, 2^G\rangle$.*

▶ **Example 2.2.** Let $h\colon G \longrightarrow H$ be the graph homomorphism indicated by green colored arrows (graphs are depicted as unlabeled graphs for simplicity). Then $\mathcal{S}(h)$ is the following graph in the left-hand side, which is an edge-extension of G where the extended edges are derived from edges of H:



2.1 Syntax: terms of KA with complement

We consider *terms* over the signature $\mathrm{S} \triangleq \{1_{(0)}, 0_{(0)}, ;_{(2)}, +_{(2)}, \_^{*}_{(1)}, \_^{-}_{(1)}\}$. Let $\mathbf{V}$ be a countably infinite set of variables. For a term t over S, let $\overline{t}$ be s if $t = s^-$ for some s and be t^- otherwise. We use the abbreviations: $\top \triangleq 0^-$ and $t \cap s \triangleq (t^- + s^-)^-$.

For $X \subseteq \{\overline{x}, \overline{1}, \top, \cap, -\}$, let KA_X be the minimal set A of terms over S satisfying:

$$\frac{y \in \mathbf{V}}{y \in A} \qquad \frac{}{1 \in A} \qquad \frac{}{0 \in A} \qquad \frac{t \in A \quad s \in A}{t\,;s \in A} \qquad \frac{t \in A \quad s \in A}{t + s \in A} \qquad \frac{t \in A}{t^* \in A}$$

$$\frac{\overline{x} \in X \quad y \in \mathbf{V}}{\overline{y} \in A} \qquad \frac{\overline{1} \in X}{\overline{1} \in A} \qquad \frac{\top \in X}{\top \in A} \qquad \frac{\cap \in X \quad t \in A \quad s \in A}{t \cap s \in A} \qquad \frac{- \in X \quad t \in A}{t^- \in A}.$$

We often abbreviate $t\,;s$ to ts. We use parentheses in ambiguous situations (where $+$ and $;$ are left-associative). We write $\sum_{i=1}^{n} t_i$ for the term $0 + t_1 + \dots + t_n$.

An *equation* $t = s$ is a pair of terms. An *inequation* $t \le s$ abbreviates the equation $t + s = s$. The set of *quantifier-free formulas* of KA_X is defined by the following grammar:

$$\varphi, \psi \quad ::= \quad t = s \mid \varphi \land \psi \mid \neg\varphi. \qquad (t, s \in \mathrm{KA}_X)$$

We use the following abbreviations, as usual: $\varphi \lor \psi \triangleq \neg(\neg\varphi \land \neg\psi)$, $\varphi \to \psi \triangleq \neg\varphi \lor \psi$, $\varphi \leftrightarrow \psi \triangleq (\varphi \to \psi) \land (\psi \to \varphi)$, $\mathsf{f} \triangleq \neg\varphi \land \varphi$, and $\mathsf{t} \triangleq \neg\mathsf{f}$. We use parentheses in ambiguous situations (where $\lor$ and $\land$ are left-associative). We write $\bigwedge_{i=1}^{n} \varphi_i$ for $\mathsf{t} \land \varphi_1 \land \dots \land \varphi_n$ and $\bigvee_{i=1}^{n} \varphi_i$ for $\mathsf{f} \lor \varphi_1 \lor \dots \lor \varphi_n$.

We say that a quantifier-free formula is *positive* if the formula in the following set A:

$$\varphi, \psi \in A \quad ::= \quad t = s \mid \varphi \land \psi \mid \varphi \lor \psi \qquad (t, s \in \mathrm{KA}_X)$$

where $\varphi \lor \psi$ expresses $\neg(\neg\varphi \land \neg\psi)$ in the above. We say that a quantifier-free formula is a *Horn formula* if the formula is of the form $(\bigwedge_{i=1}^{n} \varphi_i) \to \psi$ where $n \ge 0$.

2.2 Semantics: language models

An *S-algebra* $\mathcal{A}$ is a tuple $\langle |\mathcal{A}|, \{f^{\mathcal{A}}\}_{f_{(k)} \in \mathrm{S}} \rangle$, where $|\mathcal{A}|$ is a non-empty set and $f^{\mathcal{A}} \colon |\mathcal{A}|^k \to |\mathcal{A}|$ is a k-ary map for each $f_{(k)} \in \mathrm{S}$. A *valuation* $\mathfrak{v}$ of an S-algebra $\mathcal{A}$ is a map $\mathfrak{v} \colon \mathbf{V} \to |\mathcal{A}|$. For a valuation $\mathfrak{v}$, we write $\hat{\mathfrak{v}} \colon \mathrm{KA}_{\{-\}} \to |\mathcal{A}|$ for the unique homomorphism extending $\mathfrak{v}$. Moreover, for a quantifier-free formula φ, we define $\hat{\mathfrak{v}}(\varphi) \in \{\mathsf{true}, \mathsf{false}\}$ by:

$$\hat{\mathfrak{v}}(t = s) \Leftrightarrow (\hat{\mathfrak{v}}(t) = \hat{\mathfrak{v}}(s)), \qquad \hat{\mathfrak{v}}(\varphi \wedge \psi) \Leftrightarrow (\hat{\mathfrak{v}}(\varphi) \text{ and } \hat{\mathfrak{v}}(\psi)), \qquad \hat{\mathfrak{v}}(\neg\varphi) \Leftrightarrow (\text{not } \hat{\mathfrak{v}}(\varphi)).$$

For a quantifier-free formula φ and a class of valuations (of S-algebra) $\mathcal{C}$,[2] we write

$$\mathcal{C} \models \varphi \quad \Leftrightarrow \quad \hat{\mathfrak{v}}(\varphi) \text{ holds for all valuations } \mathfrak{v} \in \mathcal{C}.$$

We abbreviate $\{\mathfrak{v}\} \models \varphi$ to $\mathfrak{v} \models \varphi$. The *equational theory w.r.t. $\mathcal{C}$* is the set of all equations $t = s$ such that $\mathcal{C} \models t = s$. The *quantifier-free theory w.r.t. $\mathcal{C}$* is the set of all quantifier-free formulas φ such that $\mathcal{C} \models \varphi$.

The *language model* $\mathcal{A}$ over a set X, written lang_X, is the S-algebra defined by $|\mathcal{A}| = \wp(X^*)$, $\mathsf{1}^{\mathcal{A}} = \{\varepsilon\}$, $\mathsf{0}^{\mathcal{A}} = \emptyset$, and for all $L, K \subseteq X^*$,

$$L \,;^{\mathcal{A}} K = L \,; K, \qquad L +^{\mathcal{A}} K = L \cup K, \qquad L^{*^{\mathcal{A}}} = L^*, \qquad L^{-^{\mathcal{A}}} = X^* \setminus L.$$

We write LANG_X for the class of all valuations of lang_X and write LANG for $\bigcup_X \mathsf{LANG}_X$. The *equational theory (resp. quantifier-free theory) w.r.t. languages* expresses that w.r.t. LANG.

The language $[t] \subseteq \mathbf{V}^*$ of a term t is $\hat{\mathfrak{v}}_{\mathrm{st}}(t)$ where $\mathfrak{v}_{\mathrm{st}}$ is the standard language valuation on the language model over the set $\mathbf{V}$, which is defined by $\mathfrak{v}_{\mathrm{st}}(x) = \{x\}$ for $x \in \mathbf{V}$. Since $\mathfrak{v}_{\mathrm{st}} \in \mathsf{LANG}$, we have

$$\mathsf{LANG} \models t = s \quad \Rightarrow \quad [t] = [s] \tag{†}$$

The converse direction fails; e.g., when $x \neq y$, we have $[y] \subseteq [\overline{x}]$ and $\mathsf{LANG} \not\models y \leq \overline{x}$, because $[y] = \{y\} \subseteq \mathbf{V}^* \setminus \{x\} = [\overline{x}]$ and $\hat{\mathfrak{v}}(y) = \{\varepsilon\} \not\subseteq \mathbf{V}^* \setminus \{\varepsilon\} = \hat{\mathfrak{v}}(\overline{x})$ where $\mathfrak{v}$ is a valuation of lang_X s.t. $\mathfrak{v}(x) = \mathfrak{v}(y) = \{\varepsilon\}$. See [38] for more counter-examples.

▶ Remark 2.3. For (non-extended) KA, the equational theory w.r.t. languages coincides with the language equivalence [25, 2] (i.e., the converse direction of Equation (†) also holds). This is an easy consequence of the completeness theorem of KA [25] (see also [38, Appendix A] for a direct proof). From this, KA with complement (even with variable complements) has a strictly more expressive power than KA.

In the sequel, we consider the equational theory w.r.t. languages.

2.3 (Generalized) relational models

We write $\triangle_A$ for the identity relation on a set A: $\triangle_A \triangleq \{\langle x, x\rangle \mid x \in A\}$. For binary relations R, S on a set B, the composition $R \,; S$, and the reflexive transitive closure R^* are defined by:

$$R \,; S \triangleq \{\langle x, z\rangle \mid \exists y, \langle x, y\rangle \in R \wedge \langle y, z\rangle \in S\}, \qquad R^* \triangleq \bigcup_{n \geq 0} \triangle_B \,; \underbrace{R \,; \cdots ; R}_{n \text{ times}}.$$

[2] This paper considers classes of valuations rather than classes of S-algebras (cf. Theorem 3.6).

Let U be a binary relation on a non-empty set B. A *generalized relational model*[3] $\mathcal{A}$ on U is an S-algebra such that $|\mathcal{A}| \subseteq \wp(U)$, $1^{\mathcal{A}} = \triangle_B$, $0^{\mathcal{A}} = \emptyset$, and for all $R, S \subseteq U$,

$$R \mathbin{;^{\mathcal{A}}} S = R \mathbin{;} S, \qquad R +^{\mathcal{A}} S = R \cup S, \qquad R^{*^{\mathcal{A}}} = R^{*}, \qquad R^{-^{\mathcal{A}}} = U \setminus R.$$

We say that $\mathcal{A}$ is a *relational model* if $U = B^2$ and $|\mathcal{A}| = \wp(B^2)$. We write GREL (resp. REL) for the class of all valuations of generalized relational models (resp. relational models).[4]

Let $\mathcal{A}$ be a generalized relational model on a binary relation U on a set A. For a non-empty subset $B \subseteq A$, the (induced) *submodel* $\mathcal{A} \restriction B$ of $\mathcal{A}$ w.r.t. B is the generalized relational model on the binary relation $U \cap B^2$ on the set B with the universe $\{R \cap B^2 \mid R \in |\mathcal{A}|\}$. We say that a non-empty subset $B \subseteq A$ is *$\top$-closed* if for all $x, y, z \in A$, if $\langle x, z\rangle, \langle z, y\rangle \in U$ and $x, y \in B$, then $z \in B$. When B is $\top$-closed, it is easy to see that the map

$$\kappa_B \colon R \mapsto R \cap B^2$$

forms an S-homomorphism from $\mathcal{A}$ to $\mathcal{A} \restriction B$ (the condition is used for preserving ; and $*$). Similarly, for a valuation $\mathfrak{v}$ of $\mathcal{A}$, let $\mathfrak{v} \restriction B$ be the valuation of $\mathcal{A} \restriction B$ given by the map κ_B.

3 RSUB: finite relational models for language models

In this section, we define the class RSUB of *relational subword models*, for the equational theory w.r.t. languages of $\mathsf{KA}_{\{-\}}$. RSUB is a subclass of finite generalized relational models where the universe relation U is a total order.

▶ **Definition 3.1.** *Let $n \in \mathbb{N}$. The relational subword language model $\mathcal{A}$ of length n, written rsub_n, is the generalized relational model on the set $U = \{\langle i, j\rangle \in [0, n]^2 \mid i \le j\}$ s.t.*

$$|\mathcal{A}| \quad = \quad \{R \in \wp(U) \mid R \supseteq \triangle_{[0,n]} \ \vee\ U \setminus R \supseteq \triangle_{[0,n]}\}.$$

We write RSUB_n for the class of all valuations of rsub_n and write RSUB for $\bigcup_{n \ge 0} \mathsf{RSUB}_n$. ⌟

Each rsub_n is based on the image of Pratt's embedding (or called Cayley map) [46][5]:

$$\iota_X \colon L \mapsto \{\langle w, wv\rangle \mid w \in X^* \wedge v \in L\}$$

where we restrict the universe X^* of words into the subwords of a word of length n with *pairwise distinct* letters (i.e., a subword of length i corresponds to the vertex i in rsub_n).

Let rlang_X be the generalized relational model on $\iota_X(X^*)$ with the universe $\{\iota_X(L) \mid L \subseteq X^*\}$. It is easy to see that the map ι_X forms an S-isomorphism from lang_X to rlang_X. For a word w, let $\mathrm{Subw}(w)$ be the set of subwords of w. By Definition 3.1, it is easily shown that

- for a word $w \in X^*$ of length n, the generalized relational model $\mathsf{rlang}_X \restriction \mathrm{Subw}(w)$ is isomorphic to a subalgebra of rsub_n,
- for a word $w = a_1 \dots a_n \in X$ where $a_1, \dots, a_n$ are pairwise distinct letters, the generalized relational model $\mathsf{rlang}_X \restriction \mathrm{Subw}(w)$ is isomorphic to rsub_n,

[3] By definition, for each generalized relational model, U is a preorder: (Reflexivity): By $\triangle_B = 1^{\mathcal{A}} \in |\mathcal{A}| \subseteq \wp(U)$, we have $\triangle_B \subseteq U$; (Transitivity): By $\emptyset = 0^{\mathcal{A}} \in |\mathcal{A}|$, $U = \emptyset^{-^{\mathcal{A}}} \in |\mathcal{A}|$, and $U \mathbin{;} U = U \mathbin{;^{\mathcal{A}}} U \in |\mathcal{A}| \subseteq \wp(U)$, we have $U \mathbin{;} U \subseteq U$.

[4] Generalized relational models and relational models are variants of proper relation algebras and full proper relation algebras (see, e.g., [22]), respectively, where B is non-empty set and the converse operator is not introduced (due to this, U is possibly not symmetric, cf. [22, Lem. 3.4]) here.

[5] This trick itself is already used to prove equivalences between relational and language models, e.g., for KAT [29] and for $\mathsf{KA}_{\{\top\}}$ [53, 45].

by the map

$$\theta \colon R \mapsto \{\langle \|w\|, \|v\| \rangle \mid \langle w, v \rangle \in R\}.$$

We then have that the equational theory w.r.t. languages coincides with that w.r.t. RSUB.

▶ **Theorem 3.2.** *For all* $\mathrm{KA}_{\{-\}}$ *terms* t *and* s*, we have:* $\mathsf{LANG} \models t \leq s \Leftrightarrow \mathsf{RSUB} \models t \leq s$.

Proof. ($\Rightarrow$): For each $n \in \mathbb{N}$, by the *surjective* S-homomorphism given by:

$$\mathsf{lang}_X \xrightarrow{\iota_X} \mathsf{rlang}_X \xrightarrow{\kappa_{\mathrm{Subw}(a_1 \dots a_n)}} \mathsf{rlang}_X \restriction \mathrm{Subw}(a_1 \dots a_n) \xrightarrow{\theta} \mathsf{rsub}_n$$

where $a_1, \dots, a_n$ are any pairwise distinct letters and $X = \{a_1, \dots, a_n\}$. (As $\mathrm{Subw}(a_1 \dots a_n)$ is $\top$-closed, $\kappa_{\mathrm{Subw}(a_1 \dots a_n)}$ is indeed an S-homomorphism.) ($\Leftarrow$): We prove the contraposition. By $\mathsf{LANG} \not\models t \leq s$, there are X, $\mathfrak{v} \in \mathsf{LANG}_X$, and $w_0 \in X^*$ such that $w_0 \in \hat{\mathfrak{v}}(t) \setminus \hat{\mathfrak{v}}(s)$. We then consider the S-homomorphism given by:

$$\mathsf{lang}_X \xrightarrow{\iota_X} \mathsf{rlang}_X \xrightarrow{\kappa_{\mathrm{Subw}(w_0)}} \mathsf{rlang}_X \restriction \mathrm{Subw}(w_0) \xrightarrow{\theta} \mathsf{rsub}_{\|w_0\|}$$

Let $\mathfrak{v}'$, $\mathfrak{v}''$, and $\mathfrak{v}'''$ be the valuations of rlang_X, $\mathsf{rlang}_X \restriction \mathrm{Subw}(w_0)$, and rsub_n, given by $\iota_X \circ \mathfrak{v}$, $\kappa_{\mathrm{Subw}(w_0)} \circ \mathfrak{v}'$, and $\theta \circ \mathfrak{v}''$, respectively. We then have:

$$\begin{aligned}
w_0 \in \hat{\mathfrak{v}}(t) \setminus \hat{\mathfrak{v}}(s) \quad &\Rightarrow \quad \langle \varepsilon, w_0 \rangle \in \hat{\mathfrak{v}}'(t) \setminus \hat{\mathfrak{v}}'(s) && (\text{By } w_0 \in L \text{ iff } \langle \varepsilon, w_0 \rangle \in \iota_X(L)) \\
\Rightarrow \quad & \langle \varepsilon, w_0 \rangle \in \hat{\mathfrak{v}}''(t) \setminus \hat{\mathfrak{v}}''(s) && (\text{By } \varepsilon, w_0 \in \mathrm{Subw}(w_0)) \\
\Rightarrow \quad & \langle 0, \|w_0\| \rangle \in \hat{\mathfrak{v}}'''(t) \setminus \hat{\mathfrak{v}}'''(s). && (\text{By } \langle \varepsilon, w_0 \rangle \in R \text{ iff } \langle 0, \|w_0\| \rangle \in \theta(R))
\end{aligned}$$

Hence, $\mathsf{RSUB} \not\models t \leq s$. ◀

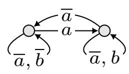
▶ Remark 3.3. By almost the same argument as Theorem 3.2, we can extend the coincidence between LANG and RSUB from the equational theory to the *positive* quantifier-free theory (see Section A for more details). However, this coincidence is broken (only $\mathsf{LANG} \models \varphi \Leftarrow \mathsf{RSUB} \models \varphi$ holds) for the quantifier-free theory and even for Horn theory. For instance, $\varphi \triangleq xx \leq 0 \to x \leq 0$ is a counter-example ($\mathsf{LANG} \models \varphi$ holds because, if $w \in \hat{\mathfrak{v}}(x)$, then $ww \in \hat{\mathfrak{v}}(xx)$; however, $\mathsf{RSUB}_1 \not\models \varphi$ under the valuation $x \mapsto \{\langle 0, 1 \rangle\}$). ⌟

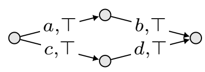
▶ **Corollary 3.4.** *The equational theory w.r.t. languages is in* Π^0_1 *for* $\mathrm{KA}_{\{-\}}$ *terms.*

Proof. By the finite model property of RSUB (the universe $|\mathsf{rsub}_n|$ is finite for each n). ◀

Comparison to other semantics

▶ Remark 3.5 (RSUB and GREL). For $\mathrm{KA}_{\{\top\}}$, the equational theory of LANG coincides with that of GREL [45, REL′ in Sect. 5][53]. However for $\mathrm{KA}_{\{-\}}$, this coincidence is broken. For instance, the following equations are valid w.r.t. LANG but not valid w.r.t. GREL (the first equation is not valid also w.r.t. REL):

$a \leq \bar{b}a\bar{b} + bab$ 

| $ab \cap cd \leq a\top d + c\top b$ 

(Each figure expresses a valuation for (G)REL $\not\models$ _ where some edges are omitted.) Here, $\mathsf{LANG} \models a \leq \bar{b}a\bar{b} + bab$ is shown by distinguishing the cases based on $\mathsf{LANG} \models 1 \leq b \vee 1 \leq \bar{b}$. The inequation $ab \cap cd \leq a\top d + c\top b$ is Levi's inequation [30][5, Example 26]. ⌟

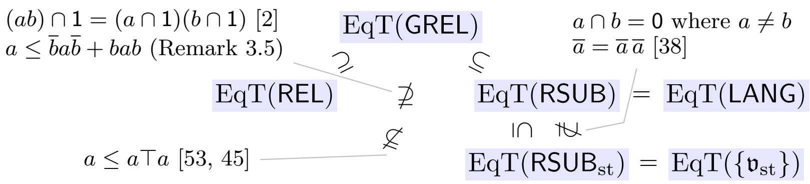


Figure 1 Equational theories for $\mathrm{KA}_{\{-\}}$ under GREL.

Additionally, the standard language valuation can also be given as a subclass of RSUB (cf. Theorem 3.2), based on the following correspondence between words and relations:

$$a_1 a_2 \dots a_n \quad \Big| \quad \to\circ \xrightarrow{a_1} \circ \xrightarrow{a_2} \circ \cdots \xrightarrow{a_n} \circ \to \quad .$$

▶ **Theorem 3.6.** *For all terms t and s, $[t] = [s]$ iff $\mathsf{RSUB}_{\mathrm{st}} \models t = s$ where*

$$\mathsf{RSUB}_{\mathrm{st}} \triangleq \bigcup_{n \ge 0} \left\{ \mathfrak{v} \in \mathsf{RSUB}_n \;\middle|\; \begin{array}{r} \bigcup_{a \in \mathbf{V}} \mathfrak{v}(a) = \{\langle i-1, i \rangle \mid i \in [1, n]\} \\ \mathfrak{v}(a) \text{ (where } a \text{ ranges over } \mathbf{V}\text{) are disjoint sets} \end{array} \right\}.$$

Proof. By the same construction in the proof of Theorem 3.2, as $\mathsf{RSUB}_{\mathrm{st}}$ is the subclass of RSUB obtained by restricting valuations to the standard language valuation $\{\mathfrak{v}_{\mathrm{st}}\}$. ◀

Figure 1 summarizes the equational theories above where the inclusions are shown by $\mathsf{REL} \subseteq \mathsf{GREL} \supseteq \mathsf{RSUB} \supseteq \mathsf{RSUB}_{\mathrm{st}}$ (and Theorem 3.2) and the non-inclusions are shown by counter-examples. Additionally, note that $\mathrm{EqT}(\{\mathfrak{v}_{\mathrm{st}}\}) = \mathrm{EqT}(\mathsf{GREL})$ for KA [25] and $\mathrm{EqT}(\mathsf{LANG}) = \mathrm{EqT}(\mathsf{GREL})$ for $\mathrm{KA}_{\{\top\}}$ [53, 45].

4 From quantifier-free formulas to equations

In this section, we show that there is a (polynomial-time) reduction from the quantifier-free theory into the equational theory, w.r.t. RSUB. Slightly more generally, we show this characterization for *submodel-closed* classes. We say that a class $\mathcal{C} \subseteq \mathsf{GREL}$ is *submodel-closed* if for all $\mathfrak{v} \in \mathcal{C}$ (on a binary relation U on a set A) and all non-empty subsets $B \subseteq A$, we have $(\mathfrak{v} \upharpoonright B) \in \mathcal{C}$. By definition, RSUB is a submodel-closed class up to isomorphism. Also, REL and GREL are submodel-closed. Additionally, for $\mathfrak{v} \in \mathsf{GREL}$ (on a binary relation U on a set A), we say that a vertex $x \in A$ is *minimal* on $\mathfrak{v}$ if $\langle x, y \rangle \in \hat{\mathfrak{v}}(\top)$ for all $y \in A$ and that a vertex $x \in A$ is *maximal* on $\mathfrak{v}$ if $\langle y, x \rangle \in \hat{\mathfrak{v}}(\top)$ for all $y \in A$. In the following lemma, we have that, to check whether a given equation is valid, it suffices to check for minimal and maximal pairs of vertices.

▶ **Lemma 4.1.** *Let $\mathcal{C} \subseteq \mathsf{GREL}$ be submodel-closed. For all terms t, s, we have: $\mathcal{C} \models t \le s \Leftrightarrow \forall \mathfrak{v} \in \mathcal{C}, \forall l, r$ s.t. l is minimal and r is maximal on $\mathfrak{v}$, $\langle l, r \rangle \notin \hat{\mathfrak{v}}(t) \setminus \hat{\mathfrak{v}}(s)$.*

Proof. ($\Rightarrow$): Trivial. ($\Leftarrow$): We prove the contraposition. Let $\mathfrak{v} \in \mathcal{C}$ (on a binary relation U on a set A), l, and r be s.t. $\langle l, r \rangle \in \hat{\mathfrak{v}}(t) \setminus \hat{\mathfrak{v}}(s)$. Let $B \triangleq \{z \in A \mid \langle l, z \rangle, \langle z, r \rangle \in U\}$. By letting $\mathfrak{v}' \triangleq \mathfrak{v} \upharpoonright B$, we have $\langle l, r \rangle \in \hat{\mathfrak{v}}'(t) \setminus \hat{\mathfrak{v}}'(s)$ $(= (\hat{\mathfrak{v}}(t) \cap B^2) \setminus (\hat{\mathfrak{v}}(s) \cap B^2))$. Hence, this completes the proof. ◀

Next, using minimal vertex l and maximal vertex r, we consider replacing each inequation $u \le 0$ with $\top u \top \le 0$, based on that $\mathfrak{v} \models u \le 0$ iff $\langle l, r \rangle \notin \hat{\mathfrak{v}}(\top u \top)$. More generally, for a quantifier-free formula φ, let $\mathrm{Tr}(\varphi)$ be the $\mathsf{KA}_{\{-\}}$ term defined by:[6]

[6] $\mathrm{Tr}(t = s)$ can be simplified for specific cases, e.g., $\mathrm{Tr}(t \le s) = \top(t \cap s^{-})\top$ and $\mathrm{Tr}(t \le 0) = \top t \top$.

$$\mathrm{Tr}(t = s) \triangleq \top((t \cap s^-) + (t^- \cap s))\top, \quad \mathrm{Tr}(\varphi \wedge \psi) \triangleq \mathrm{Tr}(\varphi) + \mathrm{Tr}(\psi), \quad \mathrm{Tr}(\neg\varphi) \triangleq \mathrm{Tr}(\varphi)^-.$$

(For the case of $t = s$, we use the fact $\mathsf{GREL} \models t = s \leftrightarrow (t \cap s^-) + (t^- \cap s) \leq \mathsf{0}$.) We then have the following.

▶ **Lemma 4.2.** *Let $\mathfrak{v} \in \mathsf{GREL}$, l be a minimal vertex on $\mathfrak{v}$, and r be a maximal vertex on $\mathfrak{v}$. For all quantifier-free formulas φ (of $\mathsf{KA}_{\{-\}}$ terms), we have:*

$$\mathfrak{v} \models \varphi \quad \Leftrightarrow \quad \langle l, r\rangle \notin \hat{\mathfrak{v}}(\mathrm{Tr}(\varphi)).$$

Proof. By easy induction on φ. Case ($t = s$): Let $u = (t \cap s^-) + (t^- \cap s)$. Then $\mathfrak{v} \models t = s$ iff $\hat{\mathfrak{v}}(u) = \emptyset$ iff $\langle l, r\rangle \notin \hat{\mathfrak{v}}(\top u \top)$ iff $\langle l, r\rangle \notin \hat{\mathfrak{v}}(\mathrm{Tr}(t = s))$. Case $\psi \wedge \rho$: By ($\langle l, r\rangle \notin \hat{\mathfrak{v}}(\mathrm{Tr}(\psi))$ and $\langle l, r\rangle \notin \hat{\mathfrak{v}}(\mathrm{Tr}(\rho))$) iff $\langle l, r\rangle \notin \hat{\mathfrak{v}}(\mathrm{Tr}(\psi) + \mathrm{Tr}(\rho))$. Case $\neg\psi$: By (not $\langle l, r\rangle \notin \hat{\mathfrak{v}}(\mathrm{Tr}(\psi))$) iff $\langle l, r\rangle \notin \hat{\mathfrak{v}}(\mathrm{Tr}(\psi)^-)$. ◀

▶ **Theorem 4.3.** *Let $\mathcal{C} \subseteq \mathsf{GREL}$ be submodel-closed. For all quantifier-free formulas φ,*

$$\mathcal{C} \models \varphi \quad \Leftrightarrow \quad \mathcal{C} \models \mathrm{Tr}(\varphi) \leq \mathsf{0}.$$

Proof. By Lemmas 4.1 and 4.2. ◀

By the reduction of Theorem 4.3, we have the following complexity results.

▶ **Corollary 4.4.** *The quantifier-free theory w.r.t. RSUB for $\mathsf{KA}_{\{-\}}$ terms is in Π^0_1.*

Proof. By Theorem 4.3 with Corollary 3.4. (The Π^0_1-hardness will be derived from Theorem 4.12.) ◀

▶ **Corollary 4.5.** *The equational theory w.r.t. $\mathsf{REL}/\mathsf{GREL}$ for $\mathsf{KA}_{\{-\}}$ terms is Π^1_1-complete.*

Proof. (Π^1_1-hard): By Theorem 4.3 with that the Horn theory of KA w.r.t. $\mathsf{REL}/\mathsf{GREL}$ is Π^1_1-complete [20]. (In Π^1_1): By the same argument as [20]. ◀

▶ Remark 4.6. In cotrast to Corollary 4.4, the authors do not know the complexity of the quantifier-free theory (resp. Horn theory) w.r.t. LANG for $\mathsf{KA}/\mathsf{KA}_{\{-\}}$ terms, cf. the Horn theory is Π^1_1-complete for $*$-*continuous* KA [27] and for KA w.r.t. $\mathsf{REL}/\mathsf{GREL}$ [20]. (E.g., in the proof of [27], *quotient models* of the standard language valuation are used, but they are not in LANG in general.) ⌟

Also, as a special case of Theorem 4.3, we have the following Hoare hypothesis elimination.

▶ **Corollary 4.7** (Hoare hypothesis elimination)**.** *Let $\mathcal{C} \subseteq \mathsf{GREL}$ be submodel-closed. For all terms t, s, u, we have:*

$$\mathcal{C} \models u \leq \mathsf{0} \to t \leq s \quad \Leftrightarrow \quad \mathcal{C} \models t \leq s + \top u \top.$$

Proof. By Theorem 4.3 with easy inequations, we have:

$$\begin{aligned} \mathcal{C} \models u \leq \mathsf{0} \to t \leq s \quad &\Leftrightarrow \quad \mathcal{C} \models \top(t \cap s^-)\top \leq \top u \top && \text{(By Theorem 4.3)} \\ &\Leftrightarrow \quad \mathcal{C} \models t \cap s^- \leq \top u \top && (\Rightarrow\text{: By } \mathsf{1} \leq \top \ \Leftarrow\text{: By } \top\top \leq \top) \\ &\Leftrightarrow \quad \mathcal{C} \models t \leq s + \top u \top. \end{aligned}$$

◀

▶ Remark 4.8. Theorem 4.3 and Corollary 4.7 fail w.r.t. LANG; for Corollary 4.7, for instance, we have:

$$\mathsf{LANG} \models xx \leq 0 \to x \leq 0, \qquad\qquad \mathsf{LANG} \not\models x \leq \top xx\top.$$

Hence, to use Hoare hypothesis elimination, it is essential to use RSUB instead of LANG. ⌟

▶ Remark 4.9. When $\mathcal{C} = \mathsf{REL}$, we have $\mathcal{C} \models \varphi \leftrightarrow \mathrm{Tr}(\varphi) \leq 0$ (cf. Theorem 4.3) and $\mathcal{C} \models (u \leq 0 \to t \leq s) \leftrightarrow (t \leq s + \top u \top)$ (cf. Corollary 4.7) by the Schröder-Tarski translation [50, XXXII.][19, p. 390, 391]. However, they fail in general when $\mathcal{C}$ is RSUB or GREL. For instance, when $\mathcal{C} = \mathsf{RSUB}$, $t = \top$, $s = 0$, and $u = x$, the second above is equivalent to "$\mathsf{RSUB} \not\models (\neg x \leq 0) \leftrightarrow \top \leq \top x \top$", but this fails; when $\mathfrak{v} \in \mathsf{RSUB}_1$ satisfies $\mathfrak{v}(x) = \{\langle 0, 1\rangle\}$, we have $\hat{\mathfrak{v}}(\top) = \{\langle 0,0\rangle, \langle 0,1\rangle, \langle 1,1\rangle\}$ but $\hat{\mathfrak{v}}(\top x \top) = \{\langle 0,1\rangle\}$. This is why we go via "$\langle l, r\rangle \notin \hat{\mathfrak{v}}(\_)$". ⌟

▶ Remark 4.10. We say that a class $\mathcal{C} \subseteq \mathsf{GREL}$ is *$\top$-submodel-closed* if, for all $\mathfrak{v} \in \mathcal{C}$ (on a binary relation U on a set A) and all $\top$-closed non-empty subsets $B \subseteq A$, we have $(\mathfrak{v} \upharpoonright B) \in \mathcal{C}$. By definition, if $\mathcal{C}$ is submodel-closed, thne $\mathcal{C}$ is $\top$-submodel-closed. Lemma 4.1, Theorem 4.3, and Corollary 4.7 can be straight-forwardly generalized for $\top$-submodel-closed classes. ⌟

4.1 Undecidability via Hoare hypothesis elimination

Using Hoare hypothesis elimination w.r.t. RSUB (Corollary 4.7) (see also Remark 4.8), we show the undecidability of the equational theory w.r.t. LANG. The proof can be obtained by the same argument as [34, Lem. 47] by replacing REL with RSUB.

A *context-free grammar* (CFG) $\mathfrak{C}$ over a finite set A is a tuple $\langle X, \mathcal{R}, s\rangle$, where

- X is a finite set of non-terminal labels s.t. $A \cap X = \emptyset$,
- $\mathcal{R}$ is a finite set of rewriting rules $x \leftarrow w$ of $x \in X$ and $w \in (A \cup X)^*$,
- $s \in X$ is the start label.

The relation $x \vdash_{\mathfrak{C}} w$, where $x \in X$ and $w \in A^*$, is defined as the minimal relation closed under the following rule: for all $n \in \mathbb{N}$, $x, x_1, \dots, x_n \in X$ and $w_0, \dots, w_n, v_1, \dots, v_n \in A^*$, if $x \leftarrow w_0 x_1 w_1 \dots x_n w_n \in \mathcal{R}$, then $\dfrac{x_1 \vdash_{\mathfrak{C}} v_1 \quad \dots \quad x_n \vdash_{\mathfrak{C}} v_n}{x \vdash_{\mathfrak{C}} w_0 v_1 w_1 \dots v_n w_n}$. The *language* $[\mathfrak{C}]$ is defined by $[\mathfrak{C}] \triangleq \{w \in A^* \mid s \vdash_{\mathfrak{C}} w\}$. It is well-known that the *universality problem* for CFGs – given a CFG $\mathfrak{C}$, does $[\mathfrak{C}] = A^*$ hold? – is Π^0_1-complete. We can naturally encode this problem by the quantifier-free theory w.r.t. RSUB as follows.

▶ **Lemma 4.11.** *Let $\mathfrak{C} = \langle X, \mathcal{R}, s\rangle$ be a CFG over a finite set $A = \{a_1, \dots, a_n\}$. Then,*

$$[\mathfrak{C}] = A^* \quad \Leftrightarrow \quad \mathsf{RSUB} \models (\bigwedge_{(x \leftarrow w) \in \mathcal{R}} w \leq x) \to ((\sum_{i=1}^{n} a_i)^* \leq s).$$

Proof. By [34, Lem. 47] with replacing REL with RSUB, because the valuations used in the proof are of the form of RSUB and the operators $\top$ and $\_^-$ do not occur in the formula. (See a full version [35] for an explicit proof.) ◀

▶ **Theorem 4.12.** *The equational theory w.r.t. languages is Π^0_1-complete for $\mathrm{KA}_{\{\overline{x}, \cap\}}$.*

Proof. (in Π^0_1): By Corollary 3.4. (Π^0_1-hard): Let $\mathfrak{C} = \langle X, \{x_i \leftarrow w_i \mid i \in [1, m]\}, s\rangle$ be a CFG over a finite set $A = \{a_1, \dots, a_n\}$. Based on $(\bigwedge_{i=1}^m w_i \leq x_i) \leftrightarrow (\sum_{i=1}^m w_i \cap \overline{x}_i \leq 0)$, by applying the Hoare hypothesis elimination (Corollary 4.7) to Lemma 4.11, we have: $[\mathfrak{C}] = A^*$ iff $\mathsf{RSUB} \models (\sum_{i=1}^n a_i)^* \leq s + \top(\sum_{i=1}^m w_i \cap \overline{x}_i)\top$. Thus, we can give a reduction from the universality problem of CFGs. ◀

Moreover, by the following fact, we can eliminate Kleene-star from Lemma 4.11.

▶ **Proposition 4.13.** $\mathsf{RSUB} \models \overline{1} = x\top \to x^* = \top$.

Proof. Let $n \in \mathbb{N}$ and $\mathfrak{v} \in \mathsf{RSUB}_n$. Let $i \in [1, n]$ be arbitrary. By $\langle i-1, i-1\rangle \not\in \hat{\mathfrak{v}}(\overline{1}) = \hat{\mathfrak{v}}(x\top)$, we have $\langle i-1, i-1\rangle \not\in \hat{\mathfrak{v}}(x)$. By $\langle i-1, i\rangle \in \hat{\mathfrak{v}}(\overline{1}) = \hat{\mathfrak{v}}(x\top)$, we have $\langle i-1, i\rangle \in \hat{\mathfrak{v}}(x)$ (by $\langle i-1, i-1\rangle \not\in \hat{\mathfrak{v}}(x)$). Thus, we have $\hat{\mathfrak{v}}(x^*) = \{\langle i, j\rangle \mid 0 \le i \le j \le n\} = \hat{\mathfrak{v}}(\top)$. ◀

▶ **Lemma 4.14.** *Let $\mathfrak{C} = \langle X, \mathcal{R}, s\rangle$ be a CFG over a finite set $A = \{a_1, \ldots, a_n\}$. Then,*

$$[\mathfrak{C}] = A^* \quad\Leftrightarrow\quad \mathsf{RSUB} \models (\overline{1} = (\sum_{i=1}^{n} a_i)\top \wedge \bigwedge_{(x \leftarrow w)\in\mathcal{R}} w \le x) \to (\top \le s).$$

Proof Sketch. By the same argument as Lemma 4.11 with replacing $(\sum_{i=1}^{n} a_i)^*$ with $\top$ using Proposition 4.13 (see [35], for a detail). ◀

Hence, the undecidability above still holds even without Kleene-star.

▶ **Theorem 4.15.** *The equational theory w.r.t. languages is Π^0_1-complete for $\mathrm{KA}_{\{-\}}$ without Kleene-star.*

Proof. By the same way as Theorem 4.12 using Lemma 4.14 instead of with Lemma 4.11. ◀

▶ Remark 4.16. Theorem 4.15 is close to Trakhtenbrot's theorem [52] in first-order logic. By a similar Kleene-star elimination via an encoding of connectivity in finite models [15, p. 30], we can also give a reduction from the universality problem of CFGs into the theory of the finite validity problem of first-order logic (resp. the calculus of relations). (See [35], for a detail.) ⌟

5 Graph characterization for $\mathrm{KA}_{\{\overline{x}, \overline{1}, \top, \cap\}}$ terms

In Sections 5 and 6, we show that the equational theory w.r.t. languages for $\mathrm{KA}_{\{\overline{x}, \overline{1}, \top\}}$ is decidable and PSPACE-complete. We recall Section 2 for graphs. In this section, we give a graph characterization of the equational theory of RSUB for $\mathrm{KA}_{\{\overline{x}, \overline{1}, \top, \cap\}}$, by generalizing the graph characterization of REL [34, Thm. 18] (and also [1, 6, 7]). Slightly more generally, we show this characterization for *submodel-closed* classes (Section 4).

5.1 Graph languages for $\mathrm{KA}_{\{\overline{x}, \overline{1}, \top, \cap\}}$

Let $\tilde{\mathbf{V}} \triangleq \{x, \overline{x} \mid x \in \mathbf{V}\} \cup \{\overline{1}, \top\}$ and $\tilde{\mathbf{V}}_1 \triangleq \tilde{\mathbf{V}} \cup \{1\}$. For a $\mathrm{KA}_{\{\overline{x}, \overline{1}, \top, \cap\}}$ term t, the graph language $\mathcal{G}(t)$ [1, 7, 34] is a set of graphs over $\tilde{\mathbf{V}}_1$ defined by:[7]

$$\mathcal{G}(x) \triangleq \{\ \to\circ\xrightarrow{x}\circ\to\ \} \text{ where } x \in \tilde{\mathbf{V}}, \qquad \mathcal{G}(0) \triangleq \emptyset, \quad \mathcal{G}(1) \triangleq \{\ \to\circ\to\ \},$$

$$\mathcal{G}(t \cap s) \triangleq \{\ \to\circ\begin{smallmatrix}G\\H\end{smallmatrix}\circ\to \mid G \in \mathcal{G}(t) \wedge H \in \mathcal{G}(s)\}, \qquad \mathcal{G}(t + s) \triangleq \mathcal{G}(t) \cup \mathcal{G}(s),$$

$$\mathcal{G}(t\,;s) \triangleq \{\ \to\circ\xrightarrow{G}\circ\xrightarrow{H}\circ\to \mid G \in \mathcal{G}(t) \wedge H \in \mathcal{G}(s)\}, \qquad \mathcal{G}(t^*) \triangleq \bigcup_{n\ge 0} \mathcal{G}(t^n).$$

For a valuation $\mathfrak{v} \in \mathsf{GREL}$ on a binary relation on a set B and $\langle x, y\rangle \in \hat{\mathfrak{v}}(\top)$, let $\mathsf{G}(\mathfrak{v}, x, y)$ be the graph defined by: $\mathsf{G}(\mathfrak{v}, x, y) \triangleq \langle B, \{\hat{\mathfrak{v}}(a)\}_{a\in\tilde{\mathbf{V}}_1}, x, y\rangle$. For a class $\mathcal{C} \subseteq \mathsf{GREL}$, let $\mathrm{GR}_{\mathcal{C}}$

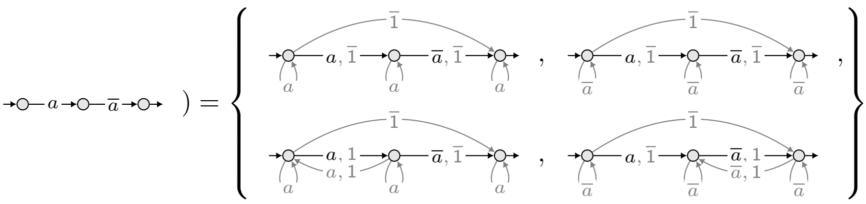
[7] We introduce $\top$-labeled edges, cf. [34, Def. 6], because $\top$ is not fixed to the full relation.

be the graph language $\{\mathsf{G}(\mathfrak{v}, x, y) \mid \mathfrak{v} \in \mathcal{C} \text{ and } \langle x, y\rangle \in \hat{\mathfrak{v}}(\top)\}$. Note that if $\mathcal{C} \subseteq \mathsf{GREL}$ is submodel-closed, then $\mathrm{GR}_{\mathcal{C}}$ is induced subgraph-closed (i.e., every induced subgraph of every $G \in \mathcal{G}$ is isomorphic to a member of $\mathcal{G}$).

We recall edge-saturations $\mathcal{S}(h)$ of Definition 2.1. For a graph G and graph language $\mathcal{G}$, let

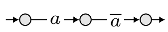
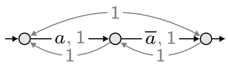
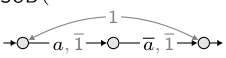
$$\mathcal{S}_{\mathcal{C}}(G) \triangleq \{\mathcal{S}(h) \mid \exists H \in \mathrm{GR}_{\mathcal{C}}, h\colon G \longrightarrow H\}, \qquad \mathcal{S}_{\mathcal{C}}(\mathcal{G}) \triangleq \bigcup_{H \in \mathcal{G}} \mathcal{S}_{\mathcal{C}}(H).$$

▶ **Example 5.1.** The following is an instance of $\mathcal{S}_{\mathsf{RSUB}}(G)$ where $\mathbf{V} = \{a\}$:



(Here, gray-colored edges are the edges extended by edge-saturations $\mathcal{S}_{\mathsf{RSUB}}$. We omit unimportant edges.)

For instance, the below right graph above can be obtained from the following map:

Note that, for instance, $\mathcal{S}_{\mathsf{RSUB}}($  $)$ does not contain graphs of the forms  nor  because we cannot make a graph homomorphism into graphs of RSUB. ⌟

By the form of $\mathrm{GR}_{\mathsf{RSUB}}$, each graph $H \in \mathcal{S}_{\mathsf{RSUB}}(\_)$ satisfies the following:

- $\top^H$ is a total preorder (possibly not a total order);
- $a^H \supseteq 1^H$ or $\overline{a}^H \supseteq 1^H$ holds for each $a \in \mathbf{V}$.

Let $H^{\mathcal{Q}} \triangleq H/(1^H)^{=}$ and $\mathcal{G}^{\mathcal{Q}} \triangleq \{H^{\mathcal{Q}} \mid H \in \mathcal{G}\}$ where $R^{=}$ denotes the equivalence closure of R. We then have the following graph language characterization, which is an analog of [34, Thm. 18], but is slightly generalized for including RSUB (see [35], for an explicit proof).

▶ **Theorem 5.2.** *Let $\mathcal{C} \subseteq \mathsf{GREL}$ be submodel-closed. For all $\mathrm{KA}_{\{\overline{x}, \overline{1}, \top, \cap\}}$ terms t, s,*

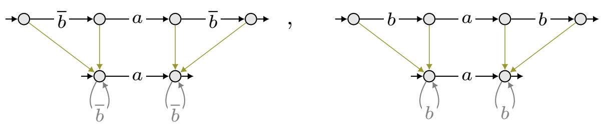
$$\mathcal{C} \models t \leq s \quad \Leftrightarrow \quad \forall H \in \mathcal{S}_{\mathcal{C}}(\mathcal{G}(t))^{\mathcal{Q}}, \exists G \in \mathcal{G}(s), G \longrightarrow H.$$

▶ **Example 5.3.** (We recall the inequations in Remark 3.5.) Here are examples to show $\mathrm{KA}_{\{\overline{x}, \overline{1}, \top, \cap\}}$ equations on RSUB using Theorem 5.2. (Gray-colored edges are the edges extended by edge-saturations $\mathcal{S}_{\mathsf{RSUB}}$. We omit unimportant edges.)

- $\mathsf{LANG} \models a \leq \overline{b}a\overline{b} + bab$: This equation is shown by the following graph homomorphisms:

$\mathcal{G}(\overline{b}a\overline{b} + bab) \quad = \quad \{ \quad \}$

$\mathcal{S}_{\mathsf{RSUB}}(\mathcal{G}(a))^{\mathcal{Q}} \ni H \quad :$

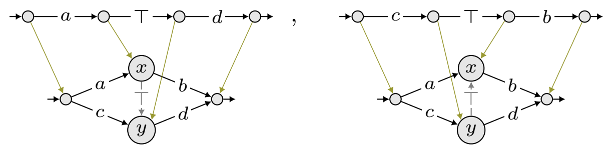


(Case $\overline{b}^H \supseteq 1^H$) (Case $b^H \supseteq 1^H$)

- LANG $\models ab \cap cd \leq a\top d + c\top b$: For each graph $H \in \mathcal{S}_{\mathsf{RSUB}}(\mathcal{G}(ab \cap cd))^{\mathcal{Q}}$, we can give a graph homomorphism from some graph in $\mathcal{G}(a\top d + c\top b)$ as follows:

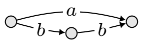
$\mathcal{G}(a\top d + c\top b) \quad = \{ \quad , \quad \}$

$\mathcal{S}_{\mathsf{RSUB}}(\mathcal{G}(ab \cap cd))^{\mathcal{Q}} \ni H \quad :$



(Case $\langle x, y \rangle \in \top^H$) (Case $\langle x, y \rangle \in \top^H$)

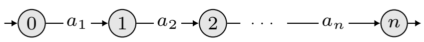
Additionally, note that $\_^{\mathcal{Q}}$ is necessary in general, e.g., for $\top \leq 1 + \bar{1}$ [34, Remark 19]. ⌟

▶ Remark 5.4. Theorem 5.2 fails when $\mathcal{C}$ is not submodel-closed. E.g., if $\mathcal{C}$ consists of the one valuation given by , $t = a$, and $s = bb$, then $\mathcal{C} \models t \leq s$ holds but the right-hand side does not hold. ⌟

5.2 Word languages for $\mathrm{KA}_{\{\overline{x}, \bar{1}, \top\}}$

Particularly for $\mathrm{KA}_{\{\overline{x}, \bar{1}, \top\}}$, Theorem 5.2 can be rephrased by *word* languages.

For a word $w = a_1 \dots a_n$ over $\tilde{\mathbf{V}}$, let $\mathsf{G}(w)$ be the following graph where $|\mathsf{G}(w)| = [0, n]$:



$\mathsf{G}(w)$ is the unique graph in $\mathcal{G}(w)$ up to graph isomorphisms.

For a $\mathrm{KA}_{\{\overline{x}, \bar{1}, \top\}}$ term t, we write $[t]_{\tilde{\mathbf{V}}}$ for the word language $[t]$ over $\tilde{\mathbf{V}}$ (namely, $\overline{x}$, $\bar{1}$, and $\top$ are also viewed as letters); e.g., $[\overline{a}]_{\tilde{\mathbf{V}}} = \{\overline{a}\}$ and $[\overline{a}] = \mathbf{V}^* \setminus \{a\}$ for $a \in \mathbf{V}$. Note that $\mathcal{G}(t) = \{\mathsf{G}(w) \mid w \in [t]_{\tilde{\mathbf{V}}}\}$; thus, for $\mathrm{KA}_{\{\overline{x}, \bar{1}, \top\}}$ terms, graph languages are expressible by using word languages.

Additionally, we introduce nondeterministic finite word automata with epsilon transitions (*NFAs*). NFAs are (2-pointed) graphs over $\tilde{\mathbf{V}}_1$ where the source and target vertices denote the initial and (single) accepting states, respectively, and 1-labeled edges denote epsilon transitions. For a graph H and a word $w = a_1 \dots a_n$, we write δ_w^H for the binary relation $(1^H)^* ; a_1^H ; (1^H)^* ; \dots ; a_n^H ; (1^H)^*$. For $q \in |H|$, we let $\delta_w^H(q) \triangleq \{q' \mid \langle q, q' \rangle \in \delta_w^H\}$. For $Q \subseteq |H|$, we let $\delta_w^H(Q) \triangleq \bigcup_{q \in Q} \delta_w^H(q)$. The word language $[H]_{\tilde{\mathbf{V}}}$ is defined as $\{w \in \tilde{\mathbf{V}}^* \mid \langle 1^H, 2^H \rangle \in \delta_w^H\}$. Note that $[H]_{\tilde{\mathbf{V}}} = \{w \in \tilde{\mathbf{V}}^* \mid \mathsf{G}(w) \longrightarrow H^{\mathcal{Q}}\}$ if 1^H is an equivalence relation. We then have the following, which a rephrasing of Theorem 5.2 (see Section B for an explicit proof). This shows that RSUB $\models t \leq s$ is equivalent to that every NFA obtained from a word w of t by an edge-saturation w.r.t. RSUB has an intersection with $[s]_{\tilde{\mathbf{V}}}$.

▶ **Corollary 5.5.** *Let $\mathcal{C} \subseteq$ GREL be submodel-closed. For all $\mathrm{KA}_{\{\overline{x}, \bar{1}, \top\}}$ terms t and s,*

$$\mathcal{C} \models t \leq s \quad \Leftrightarrow \quad [t]_{\tilde{\mathbf{V}}} \subseteq \{w \in \tilde{\mathbf{V}}^* \mid \forall H \in \mathcal{S}_{\mathcal{C}}(\mathsf{G}(w)), [s]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} \neq \emptyset\}.$$

▶ **Example 5.6.** Here are examples to show $\mathrm{KA}_{\{\overline{x}, \bar{1}, \top\}}$ equations on RSUB using Corollary 5.5.

- LANG $\models a \leq \bar{b}a\bar{b} + bab$ (the first example in Example 5.3): For all NFAs $H \in \mathcal{S}_{\mathsf{RSUB}}(\mathsf{G}(a))$, we have $[\bar{b}a\bar{b} + bab]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} \ni \begin{cases} \bar{b}a\bar{b} & (\bar{b}^H \supseteq 1^H) \\ bab & (b^H \supseteq 1^H) \end{cases}$ by the following paths:

(Case $\bar{b}^H \supseteq 1^H$) (Case $b^H \supseteq 1^H$)

- LANG $\models a \leq \bar{1} + aa$ [34, (3)]: For all NFAs $H \in \mathcal{S}_{\mathsf{RSUB}}(\mathsf{G}(a))$, we have $[\bar{1} + aa]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} \ni \begin{cases} \bar{1} & (1^H = \Delta_{|H|}) \\ aa & (1^H = \top^H) \end{cases}$ by the following paths:

(Case $1^H = \Delta_{|H|}$) (Case $1^H = \top^H$)

- LANG $\models \bar{1}a\bar{a}\bar{1} \leq \bar{1}\bar{a}a\bar{1}$ [39]: For all NFAs $H \in \mathcal{S}_{\mathsf{RSUB}}(\mathsf{G}(\bar{1}a\bar{a}\bar{1}))$, we have $[\bar{1}\bar{a}a\bar{1}]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} \ni \bar{1}\bar{a}a\bar{1}$ in either $\bar{a}^H \supseteq 1^H$ or $a^H \supseteq 1^H$ by the following paths:

(Case $\bar{a}^H \supseteq 1^H$) (Case $a^H \supseteq 1^H$)

Next, we use the NFA characterization of Corollary 5.5 for an automata construction.

6 PSPACE decidability for $\mathrm{KA}_{\{\overline{x},\bar{1},\top\}}$ terms

In this section, based on the graph characterization (Section 5), we present an NFA construction for deciding the equational theory for $\mathrm{KA}_{\{\overline{x},\bar{1},\top\}}$ terms. Here, we will use NFAs (graphs over $\tilde{\mathbf{V}}_1$) instead of $\mathrm{KA}_{\{\overline{x},\bar{1},\top\}}$ terms (regular expressions over the alphabet $\tilde{\mathbf{V}}$). To be more precise, relying on the graph characterization (Corollary 5.5), we consider the following: given an NFA J (having the same language as the term s in Corollary 5.5), we construct an NFA recognizing the following word language:

$$\mathsf{L}_J \triangleq \{w \in \tilde{\mathbf{V}}^* \mid \exists H \in \mathcal{S}_{\mathsf{RSUB}}(\mathsf{G}(w)), [J]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} = \emptyset\}.$$

Note that RSUB $\models t \leq s \Leftrightarrow [t]_{\tilde{\mathbf{V}}} \cap \mathsf{L}_J = \emptyset$ when $[s]_{\tilde{\mathbf{V}}} = [J]_{\tilde{\mathbf{V}}}$. We first present an equivalent notion of "$w \in \mathsf{L}_J$" in Section 6.1, and then we give an NFA construction in Section 6.2. Our approach in this section is based on [34] where we consider RSUB instead of REL.

6.1 Saturable paths for RSUB

We first give an equivalent notion of $[J]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} = \emptyset$ in the definition of L_J.

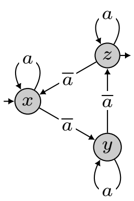
▶ **Definition 6.1.** *Let J and H be NFAs. A map $U\colon |H| \to \wp(|J|)$ is an emptiness-witness for $[J]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} = \emptyset$ if the following hold where $U_x \triangleq U(x)$:*
- $1^J \in U_{1^H}$ *and* $\forall a \in \tilde{\mathbf{V}}_1, \forall \langle x, y\rangle \in a^H$, $\delta_a^J(U_x) \subseteq U_y$,
- $2^J \notin U_{2^H}$. ⌟

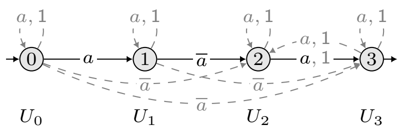
Intuitively, the first condition denotes that U is a cover of the reachable states from the pair "$1^J \in U_{1^H}$". If the second condition holds, we can see that the pair "$2^J \in U_{2^H}$" is unreachable. As expected, we have the following (see Section C, for a proof).

▶ **Proposition 6.2.** *Let J and H be NFAs where 1^H is reflexive. Then, we have:*

$$[J]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} = \emptyset \quad \Leftrightarrow \quad \exists U\colon |H| \to \wp(|J|),\ U \text{ is an emptiness-witness for } [J]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} = \emptyset.$$

▶ **Example 6.3.** We consider the following NFAs J and H. The NFA J satisfies $[J]_{\tilde{\mathbf{V}}} = \{w \in \{a, \bar{a}\}^* \mid \exists n \in \mathbb{N}, \bar{a} \text{ occurs } 3n+2 \text{ times in } w\}$ and the NFA H is a graph in $\mathcal{S}_{\mathsf{RSUB}}(\mathsf{G}(a\bar{a}a))$, where $\top$- or $\bar{1}$-labeled edges are omitted, and gray-colored edges are the edges edge-saturated from the graph $\mathsf{G}(a\bar{a}a)$. From the form of H, one can see that $[J]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} = \emptyset$.

$J =$ 

$H =$ 

.

If $U_0 = U_1 = \{\textcircled{x}\}$ and $U_2 = U_3 = \{\textcircled{y}\}$, then this U is an emptiness-witness; e.g., for $\langle 1, 2\rangle \in \overline{a}^H$, $\delta^J_{\overline{a}}(U_1) = \{\textcircled{y}\} \subseteq U_2$. By the witnesses, we have $[J]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} = \emptyset$. Besides this, if $U_0 = U_1 = \{\textcircled{x}\}$ and $U_2 = U_3 = \{\textcircled{x}, \textcircled{y}\}$, then this U is also an emptiness-witness; so, U may not coincide with the reachable states from the pair "$1^J \in U_{1^H}$". ⌟

Next, we give an equivalent notion of "$w \in \mathsf{L}_J$", by forgetting saturated edges (gray-colored edges in Example 6.3) using "U" of Proposition 6.2.

▶ **Definition 6.4.** *Let J be a NFA and w be a word. A pair $P = \langle H, U\rangle$ is a saturable path for $w \in \mathsf{L}_J$ if the following hold:*

(P-Ext) *H is an edge-extension*[8] *of $\mathsf{G}(w)$ such that*

- *$\top^H$ is a total preorder and $\top^H \supseteq \{\langle i-1, i\rangle \mid i \in [1, n]\}$ where $w = a_1 \dots a_n$,*
- *$1^H = \top^H \cap \{\langle j, i\rangle \mid \langle i, j\rangle \in \top^H\}$ and $\overline{1}^H = \top^H \setminus 1^H$,*
- *$\forall a \in \mathbf{V}$, $\langle a^H, \overline{a}^H\rangle$ is either $\langle a^{\mathsf{G}(w)} \cup 1^H, \overline{a}^{\mathsf{G}(w)}\rangle$ or $\langle a^{\mathsf{G}(w)}, \overline{a}^{\mathsf{G}(w)} \cup 1^H\rangle$.*

(P-Con) *H is consistent: $\forall a \in \mathbf{V}$, $a^{H^{\mathsf{Q}}} \cap \overline{a}^{H^{\mathsf{Q}}} = \emptyset$.*

(P-Wit) *$U \colon |H| \to \wp(|J|)$ is an emptiness-witness for $[J]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} = \emptyset$.*

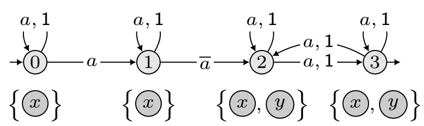
(P-Sat) *H is saturable: $\forall a \in \mathbf{V}, \forall \langle i, j\rangle \in \overline{1}^H$, $\delta^J_a(U_i) \subseteq U_j$ or $\delta^J_{\overline{a}}(U_i) \subseteq U_j$.* ⌟

Then, as expected, the existence of saturable path can characterize "$w \in \mathsf{L}_J$".

▶ **Lemma 6.5** (Section D). *Let J be a NFA and w be a word. Then,*

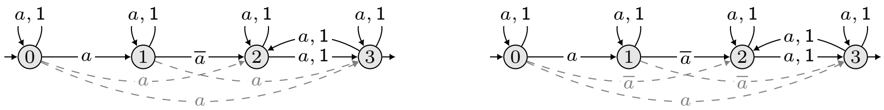
$$w \in \mathsf{L}_J \quad \Leftrightarrow \quad \text{there is a saturable path for } w \in \mathsf{L}_J.$$

▶ **Example 6.6.** We recall the NFAs J and $H \in \mathcal{S}_{\mathsf{RSUB}}(\mathsf{G}(a\overline{a}a))$ in Example 6.3. The following P is a saturable path for $a\overline{a}a \in \mathsf{L}_J$ where $\top$- or $\overline{1}$-labeled edges are omitted:

$P = \Bigg($ 

$\Bigg).$

(P is of the form of a *path graph* by taking the quotient graph w.r.t. 1-labeled edges.) P is an abstraction of edge-saturated graphs. From P, we can construct a graph $H \in \mathcal{S}_{\mathsf{RSUB}}(\mathsf{G}(a\overline{a}a))$ s.t. $[J]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} = \emptyset$. Because both $\delta^J_{\overline{a}}(\{\textcircled{x}\}) \subseteq \{\textcircled{x}, \textcircled{y}\}$ and $\delta^J_a(\{\textcircled{x}\}) \subseteq \{\textcircled{x}, \textcircled{y}\}$ hold, in addition to the graph H in Example 6.3, for instance, the following are also possible edge-saturated graphs:



. ⌟

By using saturable paths, we can replace the existence of such gray-colored edges connecting distant vertices with a "locally" defined witness U. This rephrasing will be useful for our automata construction.

To give an NFA construction, let

$$\varphi^J(\mathcal{U}, U) \;\triangleq\; \forall a \in \mathbf{V}, \forall \langle u, u'\rangle \in \mathcal{U}, \delta^J_a(u) \subseteq U \vee \delta^J_{\overline{a}}(u') \subseteq U$$

and we also replace (P-Sat) with a "local" condition.

[8] In this definition, $\top^H$-, 1^H-, and $\overline{1}^H$-edges are edge-saturated and a- and $\overline{a}$-edges in 1^H (for $a \in \mathbf{V}$) are also edge-saturated. This is for preserving (P-Con) easily.

▶ **Proposition 6.7.** *Let J and H be graphs. Let $i \in |H|$. Then we have:*

$$(\forall a \in \mathbf{V}, \forall j \text{ s.t. } \langle j, i\rangle \in \overline{1}^H, \delta^J_a(U_j) \subseteq U_i \vee \delta^J_{\overline{a}}(U_j) \subseteq U_i) \quad \Leftrightarrow \quad \varphi^J(\bigcup_{j;\langle j,i\rangle \in \overline{1}^H} U_j^2, U_i).$$

Proof. For each i, j, we have: $(\forall a \in \mathbf{V}, \delta^J_a(U_j) \subseteq U_i \vee \delta^J_{\overline{a}}(U_j) \subseteq U_i)$ iff $(\forall a \in \mathbf{V}, (\forall u \in U_j, \delta^J_a(u) \subseteq U_i) \vee (\forall u' \in U_j, \delta^J_{\overline{a}}(u') \subseteq U_i))$ iff $\varphi^J(U_j^2, U_i)$ (by taking the prenex normal form). By $(\forall j \text{ s.t. } \langle j, i\rangle \in \overline{1}^H, \varphi^J(U_j^2, U_i))$ iff $\varphi^J(\bigcup_{j;\langle j,i\rangle \in \overline{1}^H} U_j^2, U_i)$, this completes the proof. ◀

6.2 Automata from saturable paths

Let $\mathcal{X} \triangleq \{X \in \wp(\tilde{\mathbf{V}}_1) \mid 1, \top \in X, \overline{1} \notin X, \text{ and } \forall x \in \mathbf{V}, x \in X \leftrightarrow \overline{x} \notin X\}$. (This set is equivalent to the set $\{\{x \in \tilde{\mathbf{V}}_1 \mid 1^H \subseteq x^H\} \mid H \in \mathrm{GR}_{\mathsf{RSUB}}\}$.)

▶ **Definition 6.8** (NFA construction). *Let ▶ and ◀ be two fresh symbols. For a graph J and a set $X \in \mathcal{X}$, let $J^{\mathcal{S}_X}$ be the graph G defined as follows:*

- $|G| \triangleq \{▶, ◀\} \cup Q$ *where* $Q \triangleq \{\langle \mathcal{U}, U\rangle \in \wp(|J|^2) \times \wp(|J|) \mid \varphi^J(\mathcal{U}, U) \wedge \forall x \in X, \delta^J_x(U) \subseteq U\}$,
- $1^G \triangleq (\{▶\} \times \{\langle \mathcal{U}, U\rangle \in Q \mid 1^J \in U \wedge \mathcal{U} = \emptyset\}) \cup (\{\langle \mathcal{U}, U\rangle \in Q \mid 2^J \notin U\} \times \{◀\})$,
- $x^G \triangleq \{\langle\langle \mathcal{U}, U\rangle, \langle \mathcal{U}', U'\rangle\rangle \in Q^2 \mid \psi^X_{x,\overline{1}}(\mathcal{U}, U, \mathcal{U}', U') \vee \psi^X_{x,1}(\mathcal{U}, U, \mathcal{U}', U')\}$ *for* $x \in \tilde{\mathbf{V}}$,
- $1^G \triangleq ▶$,
- $2^G \triangleq ◀$.

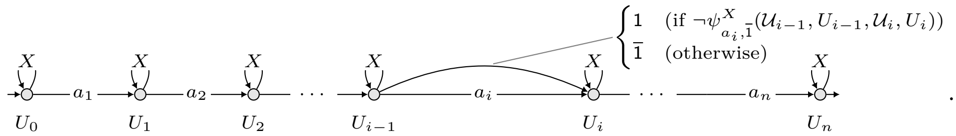
Here, $\psi^X_{x,\overline{1}}(\mathcal{U}, U, \mathcal{U}', U')$ *and* $\psi^X_{x,1}(\mathcal{U}, U, \mathcal{U}', U')$ *are defined as follows:*

- $\psi^X_{x,\overline{1}}(\mathcal{U}, U, \mathcal{U}', U') \quad \Leftrightarrow \quad \left(\mathcal{U}' = \mathcal{U} \cup U^2 \wedge \bigwedge \left\{\begin{array}{r} \delta^J_x(U) \subseteq U', \\ \delta^J_{\top}(\{u \mid \langle u, u\rangle \in \mathcal{U}'\}) \subseteq U', \\ \delta^J_{\overline{1}}(\{u \mid \langle u, u\rangle \in \mathcal{U}'\}) \subseteq U' \end{array}\right\}\right)$,
- $\psi^X_{x,1}(\mathcal{U}, U, \mathcal{U}', U') \quad \Leftrightarrow \quad (\mathcal{U}' = \mathcal{U} \wedge U' = U \wedge x \in X)$. ⌟

By the form of $J^{\mathcal{S}_X}$, if $a_1 \dots a_n \in [J^{\mathcal{S}_X}]_{\tilde{\mathbf{V}}}$, then its run is of the following form:

$$\to ▶ \xrightarrow{1} \langle \mathcal{U}_0, U_0\rangle \xrightarrow{a_1} \langle \mathcal{U}_1, U_1\rangle \xrightarrow{a_2} \langle \mathcal{U}_2, U_2\rangle \longrightarrow \cdots \xrightarrow{a_n} \langle \mathcal{U}_n, U_n\rangle \xrightarrow{1} ◀ \to .$$

Intuitively, this run corresponds to the following saturable path where some $\top$-, $\overline{1}$-, or 1-labeled edges are omitted and $\overset{X}{\circlearrowleft}$ denotes x-labeled edges for $x \in X$:



Here, $\mathcal{U}_i$ is used to denote the set $\bigcup_{j;\langle j,i\rangle \in \overline{1}^H} U_j^2$ (cf. Proposition 6.7) where H is the graph of the saturable path above. Additionally, we have $\psi^X_{a_i,\overline{1}}(\mathcal{U}_{i-1}, U_{i-1}, \mathcal{U}_i, U_i)$ if $\langle i-1, i\rangle \in \overline{1}^H$ and we have $\psi^X_{a_i,1}(\mathcal{U}_{i-1}, U_{i-1}, \mathcal{U}_i, U_i)$ if $\langle i-1, i\rangle \in 1^H$ by construction. Based on this correspondence, from a word $w \in \bigcup_{X \in \mathcal{X}} [J^{\mathcal{S}_X}]_{\tilde{\mathbf{V}}}$, we can construct a saturable path for $w \in \mathsf{L}_J$, and conversely, from a saturable path for $w \in \mathsf{L}_J$, we can show $w \in \bigcup_{X \in \mathcal{X}} [J^{\mathcal{S}_X}]_{\tilde{\mathbf{V}}}$ (see Section E, for details). Thus we have the following.

▶ **Lemma 6.9** (Section E). *Let J be a graph. Then we have* $\mathsf{L}_J = \bigcup_{X \in \mathcal{X}} [J^{\mathcal{S}_X}]_{\tilde{\mathbf{V}}}$.

▶ **Theorem 6.10.** *The equational theory w.r.t. languages for $\mathrm{KA}_{\{\overline{x}, \overline{1}, \top\}}$ is PSPACE-complete.*

Table 1 Summary of our complexity results for equational theories w.r.t. languages, with comparison to other semantics.

	KA	$\mathsf{KA}_{\{\overline{x}\}}$	$\mathsf{KA}_{\{\overline{1}\}}$	$\mathsf{KA}_{\{\overline{x},\overline{1}\}}$	$\mathsf{KA}_{\{\cap\}}$
LANG RSUB	PSPACE-c [25]	PSPACE-c (Theorem 6.10)			EXPSPACE-c [4]
$\{\mathfrak{v}_{\mathrm{st}}\}$	PSPACE-c [31]				EXPSPACE-c [18]
REL	PSPACE-c [25]	PSPACE-c [34]		in coNEXP [34]	EXPSPACE-c [6, 7, 32, 37]

	$\mathsf{KA}_{\{\overline{1},\cap\}}$	$\mathsf{KA}_{\{\overline{x},\cap\}}$	$\mathsf{KA}_{\{\overline{x},\overline{1},\cap\}}$	$\mathsf{KA}_{\{-\}}$
LANG RSUB	(open)	Π^0_1-c (Theorem 4.12 and Corollary 3.4)		
$\{\mathfrak{v}_{\mathrm{st}}\}$	EXPSPACE-c [18]			TOWER-c [49, 47]
REL	Π^0_1-c [36]	Π^0_1-c [34]		Π^1_1-c [20] (Corollary 4.5)

Proof. (in PSPACE): Let t and s be $\mathrm{KA}_{\{\overline{x},\overline{1},\top\}}$ terms. Let G and J be NFAs s.t. $[G]_{\tilde{\mathbf{V}}} = [t]_{\tilde{\mathbf{V}}}$ and $[J]_{\tilde{\mathbf{V}}} = [s]_{\tilde{\mathbf{V}}}$. By Corollary 5.5 and Lemma 6.9, we have: $\mathsf{RSUB} \models t \leq s \Leftrightarrow [G]_{\tilde{\mathbf{V}}} \cap \mathsf{L}_J = \emptyset \Leftrightarrow [G]_{\tilde{\mathbf{V}}} \cap (\bigcup_{X \in \mathcal{X}} [J^{\mathcal{S}_X}]_{\tilde{\mathbf{V}}}) = \emptyset$. Thus we can reduce the equational theory into the emptiness problem of NFAs of size exponential to the size of the input inequation, where we use the union construction for $\cup$ and the product construction for $\cap$ in NFAs. In this reduction, using a standard on-the-fly algorithm for the non-emptiness problem of NFAs (essentially the graph reachability problem), we can give a non-deterministic polynomial space algorithm. (Note that the membership of "$a \in |J^{\mathcal{S}_X}|$" and "$\langle a, b\rangle \in x^{J^{\mathcal{S}_X}}$" for each $x \in \tilde{\mathbf{V}}_1$ can be easily determined in polynomial space; so, we can construct such an on-the-fly algorithm indeed.) (Hardness): The equational theory of KA w.r.t. languages coincides with the language equivalence problem of regular expressions (Remark 2.3), which is PSPACE-complete [31]. Hence, the equational theory of $\mathrm{KA}_{\{\overline{x},\overline{1},\top\}}$ is PSPACE-hard. ◀

▶ Remark 6.11. W.r.t. REL, it is open the complexity of the equational theory for $\mathrm{KA}_{\{\overline{x},\overline{1},\top\}}$ [34, Remark 45]. W.r.t. RSUB, each equivalence class induced from 1-labeled edges is always an interval; so, the problematic case of [34, Remark 45] (w.r.t. REL) does not appear in Theorem 6.10 (w.r.t. RSUB). ⌟

7 Conclusion and Future directions

We have introduced RSUB for the equational theory w.r.t. languages for $\mathsf{KA}_{\{-\}}$ terms. Using RSUB, we have shown some complexity results for the equational theory w.r.t. languages for fragments of $\mathsf{KA}_{\{-\}}$ terms (Table 1). We leave open the decidability and complexity of the equational theory w.r.t. languages for $\mathsf{KA}_{\{\overline{1},\cap\}}$ (cf. Remark 6.11). A natural interest is to consider variants or fragments of $\mathrm{KA}_{\{-\}}$, e.g., with reverse [3], with tests [29] (by considering guarded strings) or with (anti-)domain [13]. It would also be interesting to consider the combination of variables and letters (cf. Theorems 3.2 and 3.6) in the context of language/string constraints.

Additionally, to separate the expressive power w.r.t. languages, it would also be interesting to consider games like Ehrenfeucht-Fraïssé games [16, 17] on RSUB, cf., e.g., on REL for the calculus of relations [33] and on languages for star-free expressions [51].

References

1 Hajnal Andréka and D. A. Bredikhin. The equational theory of union-free algebras of relations. *Algebra Universalis*, 33(4):516–532, 1995. doi:10.1007/BF01225472.

2 Hajnal Andréka, Szabolcs Mikulás, and István Németi. The equational theory of Kleene lattices. *Theoretical Computer Science*, 412(52):7099–7108, 2011. doi:10.1016/J.TCS.2011.09.024.

3 S. L. Bloom, Z. Ésik, and Gh. Stefanescu. Notes on equational theories of relations. *algebra universalis*, 33(1):98–126, 1995. doi:10.1007/BF01190768.

4 Paul Brunet. Reversible Kleene lattices. In *MFCS*, volume 83 of *LIPIcs*, pages 66:1–66:14. Schloss Dagstuhl, 2017. doi:10.4230/LIPICS.MFCS.2017.66.

5 Paul Brunet. A complete axiomatisation of a fragment of language algebra. In *CSL*, volume 152 of *LIPIcs*, pages 11:1–11:15. Schloss Dagstuhl, 2020. doi:10.4230/LIPIcs.CSL.2020.11.

6 Paul Brunet and Damien Pous. Petri automata for Kleene allegories. In *LICS*, pages 68–79. IEEE, 2015. doi:10.1109/LICS.2015.17.

7 Paul Brunet and Damien Pous. Petri automata. *Logical Methods in Computer Science*, 13(3), 2017. doi:10.23638/LMCS-13(3:33)2017.

8 Wojciech Buszkowski. On the complexity of the equational theory of relational action algebras. In *RAMICS*, volume 4136 of *LNTCS*, pages 106–119. Springer, 2006. doi:10.1007/11828563_7.

9 Ernie Cohen. Hypotheses in Kleene algebra. *Unpublished manuscript*, 1994.

10 Rina S. Cohen and J. A. Brzozowski. Dot-depth of star-free events. *Journal of Computer and System Sciences*, 5(1):1–16, 1971. doi:10.1016/S0022-0000(71)80003-X.

11 John H. Conway. *Regular Algebra and Finite Machines*. Chapman and Hall, 1971.

12 Isabel F. Cruz, Alberto O. Mendelzon, and Peter T. Wood. A graphical query language supporting recursion. *ACM SIGMOD Record*, 16(3):323–330, 1987. doi:10.1145/38714.38749.

13 Jules Desharnais, Bernhard Möller, and Georg Struth. Kleene algebra with domain. *ACM Transactions on Computational Logic*, 7(4):798–833, 2006. doi:10.1145/1183278.1183285.

14 Amina Doumane, Denis Kuperberg, Damien Pous, and Pierre Pradic. Kleene algebra with hypotheses. In *FoSSaCS*, volume 11425 of *LNTCS*, pages 207–223. Springer, 2019. doi:10.1007/978-3-030-17127-8_12.

15 Heinz-Dieter Ebbinghaus and Jörg Flum. *Finite Model Theory*. Springer, 1995. doi:10.1007/3-540-28788-4.

16 Andrzej Ehrenfeucht. An application of games to the completeness problem for formalized theories. *Fundamenta Mathematicae*, 49(2):129–141, 1961. doi:10.4064/fm-49-2-129-141.

17 Roland Fraïssé. Sur les classifications des systems de relations. *Publ. Sci. Univ. Alger I.*, 1954.

18 Martin Fürer. The complexity of the inequivalence problem for regular expressions with intersection. In *ICALP*, volume 85 of *LNCS*, pages 234–245. Springer, 1980. doi:10.1007/3-540-10003-2_74.

19 Steven Givant. The calculus of relations. In *Introduction to Relation Algebras*, volume 1, pages 1–34. Springer International Publishing, 2017. doi:10.1007/978-3-319-65235-1_1.

20 Chris Hardin and Dexter Kozen. On the complexity of the Horn theory of REL. Technical report, Cornell University, 2003. URL: https://hdl.handle.net/1813/5612.

21 Jelle Hellings, Catherine L. Pilachowski, Dirk Van Gucht, Marc Gyssens, and Yuqing Wu. From relation algebra to semi-join algebra: An approach to graph query optimization. *The Computer Journal*, 64(5):789–811, 2021. doi:10.1093/comjnl/bxaa031.

22 Robin Hirsch and Ian Hodkinson. *Relation Algebras by Games*, volume 147 of *Studies in logic and the foundations of mathematics*. Elsevier, 1 edition, 2002.

23 Tony Hoare, Bernhard Möller, Georg Struth, and Ian Wehrman. Concurrent Kleene algebra and its foundations. *The Journal of Logic and Algebraic Programming*, 80(6):266–296, 2011. doi:10.1016/j.jlap.2011.04.005.

24 Stephen C. Kleene. Representation of events in nerve nets and finite automata. In *Automata Studies. (AM-34)*, pages 3–42. Princeton University Press, 1956. doi:10.1515/9781400882618-002.

25 Dexter Kozen. A completeness theorem for Kleene algebras and the algebra of regular events. In *LICS*, pages 214–225. IEEE, 1991. doi:10.1109/LICS.1991.151646.

26 Dexter Kozen. On Hoare logic and Kleene algebra with tests. *ACM Transactions on Computational Logic*, 1(1):60–76, 2000. doi:10.1145/343369.343378.

27 Dexter Kozen. On the complexity of reasoning in Kleene algebra. *Information and Computation*, 179(2):152–162, 2002. doi:10.1006/INCO.2001.2960.

28 Dexter Kozen and Konstantinos Mamouras. Kleene algebra with equations. In *ICALP*, volume 8573 of *LNCS*, pages 280–292. Springer, 2014. doi:10.1007/978-3-662-43951-7_24.

29 Dexter Kozen and Frederick Smith. Kleene algebra with tests: Completeness and decidability. In *CSL*, volume 1258 of *LNCS*, pages 244–259. Springer, 1996. doi:10.1007/3-540-63172-0_43.

30 F. W. Levi. On semigroups. *Bulletin of the Calcutta Mathematical Society*, 36(36):141–146, 1944.

31 A. R. Meyer and L. J. Stockmeyer. The equivalence problem for regular expressions with squaring requires exponential space. In *SWAT*, pages 125–129. IEEE, 1972. doi:10.1109/SWAT.1972.29.

32 Yoshiki Nakamura. Partial derivatives on graphs for Kleene allegories. In *LICS*, pages 1–12. IEEE, 2017. doi:10.1109/LICS.2017.8005132.

33 Yoshiki Nakamura. Expressive power and succinctness of the positive calculus of binary relations. *Journal of Logical and Algebraic Methods in Programming*, 127:100760, 2022. doi:10.1016/j.jlamp.2022.100760.

34 Yoshiki Nakamura. Existential calculi of relations with transitive closure: Complexity and edge saturations. In *LICS*, pages 1–13. IEEE, 2023. doi:10.1109/LICS56636.2023.10175811.

35 Yoshiki Nakamura. Finite relational semantics for language Kleene algebra with complement, 2024. URL: https://hal.science/hal-04455882.

36 Yoshiki Nakamura. Undecidability of the positive calculus of relations with transitive closure and difference: Hypothesis elimination using graph loops. In *RAMICS*, volume 14787 of *LNTCS*, pages 207–224. Springer, 2024. doi:10.1007/978-3-031-68279-7_13.

37 Yoshiki Nakamura. Derivatives on graphs for the positive calculus of relations with transitive closure, 2024 (submitted, journal version of [32]). doi:10.48550/arXiv.2408.08236.

38 Yoshiki Nakamura and Ryoma Sin'ya. Words-to-letters valuations for language Kleene algebras with variable complements. In *AFL*, volume 386 of *EPTCS*, pages 185–199. EPTCS, 2023. doi:10.4204/EPTCS.386.15.

39 Yoshiki Nakamura and Ryoma Sin'ya. Words-to-letters valuations for language Kleene algebras with variable and constant complements, 2024 (accepted, journal version of [38]). URL: https://arxiv.org/abs/2309.02760.

40 Kan Ching Ng. *Relation algebras with transitive closure*. PhD thesis, University of California, 1984.

41 Peter W. O'Hearn. Incorrectness logic. *Proceedings of the ACM on Programming Languages*, 4(POPL):10:1–10:32, 2019. doi:10.1145/3371078.

42 Jean-Éric Pin. The dot-depth hierarchy, 45 years later. In *The Role of Theory in Computer Science*, pages 177–201. WORLD SCIENTIFIC, 2016. doi:10.1142/9789813148208_0008.

43 Damien Pous. On the positive calculus of relations with transitive closure. In *STACS*, volume 96 of *LIPIcs*, pages 3:1–3:16. Schloss Dagstuhl, 2018. doi:10.4230/LIPICS.STACS.2018.3.

44 Damien Pous, Jurriaan Rot, and Jana Wagemaker. On tools for completeness of kleene algebra with hypotheses. *Logical Methods in Computer Science*, Volume 20, Issue 2, 2024. doi:10.46298/lmcs-20(2:8)2024.

45 Damien Pous and Jana Wagemaker. Completeness theorems for Kleene algebra with top. In *CONCUR*, volume 243 of *LIPIcs*, pages 26:1–26:18. Schloss Dagstuhl, 2022. doi:10.4230/LIPICS.CONCUR.2022.26.

46 V. R. Pratt. Dynamic algebras and the nature of induction. In *STOC*, pages 22–28. ACM, 1980. doi:10.1145/800141.804649.

47 Sylvain Schmitz. Complexity hierarchies beyond elementary. *ACM Transactions on Computation Theory*, 8(1):1–36, 2016. doi:10.1145/2858784.

48 Steffen Smolka, Nate Foster, Justin Hsu, Tobias Kappé, Dexter Kozen, and Alexandra Silva. Guarded Kleene algebra with tests: verification of uninterpreted programs in nearly linear time. *Proceedings of the ACM on Programming Languages*, 4(POPL):61:1–61:28, 2019. doi:10.1145/3371129.

49 Larry J. Stockmeyer. *The Complexity of Decision Problems in Automata Theory and Logic*. PhD thesis, Massachusetts Institute of Technology, 1974. doi:1721.1/15540.

50 Alfred Tarski. On the calculus of relations. *The Journal of Symbolic Logic*, 6(3):73–89, 1941. doi:10.2307/2268577.

51 Wolfgang Thomas. A concatenation game and the dot-depth hierarchy. In *Computation Theory and Logic*, number 270 in LNCS, pages 415–426. Springer, 1987. doi:10.1007/3-540-18170-9_183.

52 B. A. Trakhtenbrot. The impossibility of an algorithm for the decision problem in finite classes. *Doklady Akademii Nauk SSSR*, 70(4):569–572, 1950.

53 Cheng Zhang, Arthur Azevedo de Amorim, and Marco Gaboardi. On incorrectness logic and Kleene algebra with top and tests. *Proceedings of the ACM on Programming Languages*, 6(POPL):29:1–29:30, 2022. doi:10.1145/3498690.

A Slight Extensions of Theorem 3.2

In this section, we note that we can extend Theorem 3.2 in the following two:

▶ **Theorem A.1.** *For all positive quantifier-free formulas φ of $\mathsf{KA}_{\{-\}}$ terms, we have:* $\mathsf{LANG} \models \varphi \Rightarrow \mathsf{RSUB} \models \varphi$.

Proof Sketch. By the same surjective S-homomorphism in the proof of Theorem 3.2(⇒). ◀

▶ **Theorem A.2.** *For all quantifier-free formulas φ of $\mathsf{KA}_{\{-\}}$ terms, we have:* $\mathsf{RSUB} \models \varphi \Rightarrow \mathsf{LANG} \models \varphi$.

Proof. Because the formulas $t = s \leftrightarrow (t \leq s \wedge s \leq t)$ and $t \leq s \leftrightarrow t \cap s^{-} \leq \mathsf{0}$ are valid on $\mathsf{LANG} \cup \mathsf{SUB}$, without loss of generality, we can assume that each equation in φ is of the form $u \leq \mathsf{0}$. By taking the conjunctive normal form, it suffices to prove when φ is of the form $(\bigvee_{i=1}^{n} t_i \leq \mathsf{0}) \vee (\bigvee_{j=1}^{m} \neg s_j \leq \mathsf{0})$. We prove the contraposition. By $\mathsf{LANG} \not\models \varphi$, there are $X, \mathfrak{v} \in \mathsf{LANG}_X, w_1, \dots, w_n \in X^*$ such that $w_i \in \hat{\mathfrak{v}}(t_i)$ for $i \in [1, n]$ and $\hat{\mathfrak{v}}(s_j) = \emptyset$ for $j \in [1, m]$. By letting $w_0 \triangleq w_1 \dots w_n$ and considering the same S-homomorphism as Theorem 3.2(⇐), we have $\mathsf{RSUB} \not\models \varphi$. ◀

For Theorem A.2, note that the converse direction fails (Remark 3.3), cf. Theorem A.1.

B Proof of Corollary 5.5

Proof. We have:

$$\mathcal{C} \models t \leq s \;\Leftrightarrow\; \forall w \in [t]_{\tilde{\mathbf{V}}}, \forall H \in \mathcal{S}_{\mathcal{C}}(\mathsf{G}(w)), \exists v \in [s]_{\tilde{\mathbf{V}}}, \mathsf{G}(v) \longrightarrow H^{\mathcal{Q}}$$

(Theorem 5.2 and $\mathcal{G}(s) = \{\mathsf{G}(v) \mid v \in [s]_{\tilde{\mathbf{V}}}\}$)

$$\Leftrightarrow \quad \forall w \in [t]_{\tilde{\mathbf{V}}}, \forall H \in \mathcal{S}_{\mathcal{C}}(\mathsf{G}(w)), [s]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} \neq \emptyset \qquad ([H]_{\tilde{\mathbf{V}}} = \{v \in \tilde{\mathbf{V}}^* \mid \mathsf{G}(v) \longrightarrow H^{\mathcal{Q}}\})$$

$$\Leftrightarrow \quad [t]_{\tilde{\mathbf{V}}} \subseteq \{w \in \tilde{\mathbf{V}}^* \mid \forall H \in \mathcal{S}_{\mathcal{C}}(\mathsf{G}(w)), [s]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} \neq \emptyset\}.$$

◀

C Proof of Proposition 6.2

Let $R' \subseteq |H| \times |J|$ be the minimal set such that

- $\langle 1^H, 1^J \rangle \in R'$,
- $\forall a \in \tilde{\mathbf{V}}_1, \forall x, x' \in |H|, \forall y, y' \in |J|, (\langle x, y \rangle \in R' \land \langle x, x' \rangle \in \delta_a^H \land \langle y, y' \rangle \in \delta_a^J) \Rightarrow \langle x', y' \rangle \in R'$.

▷ Claim C.1. $[J]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} \neq \emptyset \quad \Leftrightarrow \quad \langle 2^H, 2^J \rangle \in R'$.

Proof. By definition, R' coincides with the set of all reachable states of the product NFA of H and J. ◁

Let $R \subseteq |H| \times |J|$ be the minimal set such that

- $\langle 1^H, 1^J \rangle \in R$,
- $\forall a \in \tilde{\mathbf{V}}_1, \forall x, x' \in |H|, \forall y, y' \in |J|, (\langle x, y \rangle \in R \land \langle x, x' \rangle \in a^H \land \langle y, y' \rangle \in \delta_a^J) \Rightarrow \langle x', y' \rangle \in R$.

▷ Claim C.2. $R = R'$.

Proof. ($\subseteq$): Clear, by $a^H \subseteq \delta_a^H$. ($\supseteq$): By induction on derivations of R'.

- Case $\langle 1^H, 1^J \rangle \in R'$: Trivial, by $\langle 1^H, 1^J \rangle \in R$.
- Case $(\langle x, y \rangle \in R' \land \langle x, x' \rangle \in \delta_a^H \land \langle y, y' \rangle \in \delta_a^J) \Rightarrow \langle x', y' \rangle \in R'$: By IH, $\langle x, y \rangle \in R$.
  - Sub-Case $a \neq 1$: Let $x_0, \dots, x_{n-1}, x_n, \dots, x_m$ be s.t. $\langle x, x' \rangle = \langle x_0, x_m \rangle$ and
    * for all $i \in [1, n-1]$, $\langle x_{i-1}, x_i \rangle \in 1^H$,
    * $\langle x_{n-1}, x_n \rangle \in a^H$,
    * for all $i \in [n+1, m]$, $\langle x_{i-1}, x_i \rangle \in 1^H$.

    Let $y_0 = \dots = y_{n-1} = y$ and $y_n = \dots = y_m = y'$. Then by applying the second rule multiply, we have $\langle x', y' \rangle \in R$.
  - Sub-Case $a = 1$: By reflexivity of 1^H, $\langle x, x' \rangle \in (1^H)^+$. Let $x_0, \dots, x_m$ ($m > 0$) be s.t. $\langle x, x' \rangle = \langle x_0, x_m \rangle$ and
    * for all $i \in [1, m]$, $\langle x_{i-1}, x_i \rangle \in 1^H$.

    Let $y_0 = y$ and $y_1 = \dots = y_m = y'$. Then by applying the second rule multiply, we have $\langle x', y' \rangle \in R$. ◁

Proof of Proposition 6.2. ($\Rightarrow$): By letting U as the map defined by $U(x) \triangleq \{y \mid \langle x, y \rangle \in R\}$. Here, $2^J \notin U_{2^H}$ is shown by $[J]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} = \emptyset$ with Claim C.1 and C.2. ($\Leftarrow$): Let $R'' \triangleq \{\langle x, y \rangle \mid y \in U(x)\}$. By the minimality of R, we have $R \subseteq R''$. By $\langle 2^H, 2^J \rangle \notin R''$, we have $\langle 2^H, 2^J \rangle \notin R$. Hence by Claim C.1 and Claim C.2, we have $[J]_{\tilde{\mathbf{V}}} \cap [H]_{\tilde{\mathbf{V}}} = \emptyset$. ◀

D Proof of Lemma 6.5

Proof. ($\Rightarrow$): By Proposition 6.2, let $H' \in \mathcal{S}_{\mathsf{RSUB}}(\mathsf{G}(w))$ and let U be an emptiness-witness for $[J]_{\tilde{\mathbf{V}}} \cap [H']_{\tilde{\mathbf{V}}} = \emptyset$. We define the graph H as follows:

- $|H| = |H'|$,
- $a^H = a^{H'}$ for $a \in \{\top, 1, \bar{1}\}$,
- $a^H = a^{\mathsf{G}(w)} \cup (a^{H'} \cap 1^{H'})$ for $a \in \tilde{\mathbf{V}}_1 \setminus \{\top, 1, \bar{1}\}$.

We then have that the pair $P \triangleq \langle H, U \rangle$ is a saturable path for $w \in \mathsf{L}_J$, as follows:

- (P-Ext): By that H' is an edge-saturation w.r.t. RSUB.
- (P-Con): Because H' is consistent by $H' \in \mathcal{S}_{\mathsf{RSUB}}(\mathsf{G}(w))$.
- (P-Wit): Because U is an emptiness-witness for $[J]_{\tilde{\mathbf{V}}} \cap [H']_{\tilde{\mathbf{V}}} = \emptyset$.
- (P-Sat): Because $a^{H'} \cup \bar{a}^{H'} = \top^{H'}$ and U is an emptiness-witness for $[J]_{\tilde{\mathbf{V}}} \cap [H']_{\tilde{\mathbf{V}}} = \emptyset$.

($\Leftarrow$): Let $P = \langle H, U \rangle$ be a saturable path for $w \in \mathsf{L}_J$. By (P-Ext), 1^H is an equivalence relation. We define the graph H' as follows:

- $|H'| = |H|$,
- $a^{H'} = a^H$ for $a \in \{\top, 1, \overline{1}\}$,
- for $a \in \mathbf{V}$ and $\langle x, y\rangle \in \top^H$,
  - if $\langle [x]_{1^H}, [y]_{1^H}\rangle \in a^{H^{\mathcal{Q}}}$, then $\langle x, y\rangle \in a^{H'} \setminus \overline{a}^{H'}$,
  - else if $\langle [x]_{1^H}, [y]_{1^H}\rangle \in \overline{a}^{H^{\mathcal{Q}}}$, then $\langle x, y\rangle \in \overline{a}^{H'} \setminus a^{H'}$,
  - else if $U_y \subseteq \delta^J_a(U_x)$, then $\langle x, y\rangle \in a^{H'} \setminus \overline{a}^{H'}$,
  - else $\langle x, y\rangle \in \overline{a}^{H'} \setminus a^{H'}$.

By the construction of H', we have the following:

- H' is an edge-extension of H: By (P-Con), if $\langle [x]_{1^H}, [y]_{1^H}\rangle \in \overline{a}^{H^{\mathcal{Q}}}$, then $\langle [x]_{1^H}, [y]_{1^H}\rangle \notin a^{H^{\mathcal{Q}}}$.
- H' is consistent: If $[x]_{1^H} = [y]_{1^H}$ then $U_x = U_y$, because $U_x \subseteq \delta^J_1(U_x) \subseteq U_y \subseteq \delta^J_1(U_y) \subseteq U_x$ by (P-Wit); thus, if $[x]_{1^H} = [x']_{1^H}$ and $[y]_{1^H} = [y']_{1^H}$, then $\langle x, y\rangle \in a^{H'}$ iff $\langle x', y'\rangle \in a^{H'}$.
- for $a \in \mathbf{V}$, $\overline{a}^{H'} = \top^{H'} \setminus a^{H'}$: Because $a^{H'} \cup \overline{a}^{H'} = \top^{H'}$ and H' is consistent.

From them and (P-Ext), we have $H' \in \mathcal{S}_{\mathsf{RSUB}}(\mathsf{G}(w))$. Also, U is an emptiness-witness for $[J]_{\tilde{\mathbf{V}}} \cap [H']_{\tilde{\mathbf{V}}} = \emptyset$ as follows. For edges already in H, it is shown by (P-Wit). For extended edges from H, it is shown by the construction of H' (for the last case of the four cases above, by $U_y \not\subseteq \delta^J_a(U_x)$ and (P-Sat), we have $U_y \subseteq \delta^J_{\overline{a}}(U_x)$). Hence, this completes the proof. ◀

E Proof of Lemma 6.9

Proof. ($\subseteq$): Let $w = a_1 \dots a_n \in \mathsf{L}_J$. Let $P = \langle H, U\rangle$ be a saturable path for $w \in \mathsf{L}_J$. Let $X \triangleq \{a \in \tilde{\mathbf{V}}_1 \mid a^H \supseteq 1^H\}$ (note that $X \in \mathcal{X}$). For each i, let $\mathcal{U}_i \triangleq \bigcup_{j;\langle j,i\rangle \in \overline{1}^H} U^2_j$. Then we have:

- $\varphi^J(\mathcal{U}_i, U_i)$: By (P-Sat) and Proposition 6.7.
- $\forall a \in X, \delta^J_a(U_i) \subseteq U_i$: By $a^H \supseteq 1^H \supseteq \Delta_{|H|}$ and (P-Wit).

Thus $\langle \mathcal{U}_i, U_i\rangle \in |J^{\mathcal{S}_X}|$. We consider the following run of the NFA $J^{\mathcal{S}_X}$ on w:

$$\rightarrow \blacktriangleright \xrightarrow{1} \langle \mathcal{U}_0, U_0\rangle \xrightarrow{a_1} \langle \mathcal{U}_1, U_1\rangle \xrightarrow{a_2} \langle \mathcal{U}_2, U_2\rangle \longrightarrow \cdots \xrightarrow{a_n} \langle \mathcal{U}_n, U_n\rangle \xrightarrow{1} \blacktriangleleft \rightarrow \quad .$$

This is indeed a run of the NFA $J^{\mathcal{S}_X}$ as follows:

- $\langle \blacktriangleright, \langle \mathcal{U}_0, U_0\rangle\rangle \in 1^{J^{\mathcal{S}_X}}$: By $1^J \in U_0$ (P-Wit) and $\mathcal{U}_0 = \emptyset$.
- $\langle \langle \mathcal{U}_n, U_n\rangle, \blacktriangleleft\rangle \in 1^{J^{\mathcal{S}_X}}$: By $2^J \notin U_n$ (P-Wit).
- $\forall i \in [1, n], \langle \langle \mathcal{U}_{i-1}, U_{i-1}\rangle, \langle \mathcal{U}_i, U_i\rangle\rangle \in a_i^{J^{\mathcal{S}_X}}$: We distinguish the following cases:
  - Case $\langle i-1, i\rangle \in 1^H$:
    - $\mathcal{U}_i = \mathcal{U}_{i-1}$: By $\langle j, i\rangle \in \overline{1}^H$ iff $\langle j, i-1\rangle \in \overline{1}^H$, for all j.
    - $U_i = U_{i-1}$: By (P-Wit), we have $U_{i-1} \subseteq \delta^J_1(U_{i-1}) \subseteq U_i \subseteq \delta^J_1(U_i) \subseteq U_{i-1}$.
    - $a_i \in X$ ($a_i^H \supseteq 1^H$): By $a_i^H \cap 1^H \neq \emptyset$ and (P-Ext), we have $a_i^H = a_i^{\mathsf{G}(w)} \cup 1^H$ (if not, this contradicts to (P-Con)).

    Thus by $\psi^X_{a_i,1}(\mathcal{U}_{i-1}, U_{i-1}, \mathcal{U}_i, U_i)$, we have $\langle \langle \mathcal{U}_{i-1}, U_{i-1}\rangle, \langle \mathcal{U}_i, U_i\rangle\rangle \in a_i^{J^{\mathcal{S}_X}}$.
  - Case $\langle i-1, i\rangle \in \overline{1}^H$:
    - $\mathcal{U}_i = \mathcal{U}_{i-1} \cup U^2_{i-1}$: By $\langle j, i\rangle \in \overline{1}^H$ iff $j < i$ iff $\langle j, i-1\rangle \in \overline{1}^H \vee \langle j, i-1\rangle \in 1^H$, for all j. (Intuitively, $\mathcal{U}_{i-1}$ corresponds to the case $\langle j, i-1\rangle \in \overline{1}^H$ and U^2_{i-1} corresponds to the case $\langle j, i-1\rangle \in 1^H$.)
    - $\delta^J_{a_i}(U_{i-1}) \subseteq U_i$: By (P-Wit).
    - $\delta^J_\top(\{u \mid \langle u, u\rangle \in \mathcal{U}_i\}) \subseteq U_i$: We have $\delta^J_\top(\{u \mid \langle u, u\rangle \in \mathcal{U}_i\}) = \delta^J_\top(\bigcup_{j;\langle j,i\rangle \in \overline{1}^H} U_j) = \bigcup_{j<i} \delta^J_\top(U_j) \subseteq U_i$ by (P-Wit).

  * $\delta^J_{\overline{1}}(\{u \mid \langle u, u\rangle \in \mathcal{U}_i\}) \subseteq U_i$: We have $\delta^J_{\overline{1}}(\{u \mid \langle u, u\rangle \in \mathcal{U}_i\}) = \delta^J_{\overline{1}}(\bigcup_{j;\langle j,i\rangle \in \overline{1}^H} U_j) = \bigcup_{j<i} \delta^J_{\overline{1}}(U_j) \subseteq U_i$ by (P-Wit).

  Thus by $\psi^X_{a_i,\overline{1}}(\mathcal{U}_{i-1}, U_{i-1}, \mathcal{U}_i, U_i)$, we have $\langle\langle \mathcal{U}_{i-1}, U_{i-1}\rangle, \langle \mathcal{U}_i, U_i\rangle\rangle \in a_i^{J^{\mathcal{S}_X}}$.

Hence, $w \in [J^{\mathcal{S}_X}]$.

($\supseteq$): Let $X \subseteq \mathcal{X}$ and $w = a_1 \dots a_n \in [J^{\mathcal{S}_X}]_{\tilde{\mathbf{V}}}$. Let the run of $J^{\mathcal{S}_X}$ on w be as follows:

$$\to \blacktriangleright \xrightarrow{1} \langle \mathcal{U}_0, U_0\rangle \xrightarrow{a_1} \langle \mathcal{U}_1, U_1\rangle \xrightarrow{a_2} \langle \mathcal{U}_2, U_2\rangle \longrightarrow \cdots \xrightarrow{a_n} \langle \mathcal{U}_n, U_n\rangle \xrightarrow{1} \blacktriangleleft \to .$$

Let H be the edge-extension of $\mathsf{G}(w)$ defined as follows:

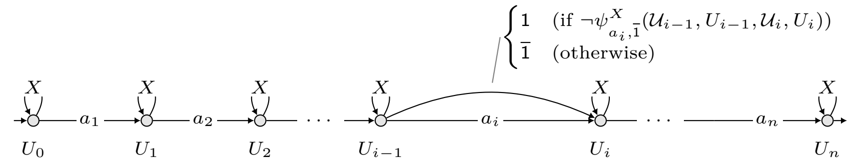
- $\top^H = \{\langle x, y\rangle \in [0, n]^2 \mid \forall i \in [y+1, x],\ \neg\psi^X_{a_i,\overline{1}}(\mathcal{U}_{i-1}, U_{i-1}, \mathcal{U}_i, U_i)\}$,
- $1^H = \top^H \cap \{\langle x, y\rangle \mid \langle y, x\rangle \in \top^H\}$ and $\overline{1}^H = \top^H \setminus 1^H$,
- $\forall a \in \mathbf{V} \cap X$, $\langle a^H, \overline{a}^H\rangle = \langle a^{\mathsf{G}(w)} \cup 1^H, \overline{a}^{\mathsf{G}(w)}\rangle$.

Note that by definition of $\top^H$, we have

- $\top^H \supseteq \{\langle x, y\rangle \mid x \le y\}$,
- $\top^H$ is transitive (by case analysis).

Hence, $\top^H$ is a total preorder and each equivalence class w.r.t. 1^H is an interval $[l, r]$.

Let $P \triangleq \langle H, U\rangle$ where U is defined as $i \mapsto U_i$ for $i \in [0, n]$. The following depicts P.



Then P is a saturable path for $w \in \mathsf{L}_J$ as follows:

- (P-Ext): By the definition of H.
- (P-Con): Assume that $a^{H^{\mathcal{Q}}} \cap \overline{a}^{H^{\mathcal{Q}}} \neq \emptyset$. Let x, x', y, y' be s.t. $[x]_{1^H} = [x']_{1^H}$, $[y]_{1^H} = [y']_{1^H}$, $\langle x, y\rangle \in a^H$, and $\langle x', y'\rangle \in \overline{a}^H$. WLOG, we can assume that $a \in X$ and $\overline{a} \notin X$. Then, we have the following:
  - $\langle x', y'\rangle \in \overline{a}^{\mathsf{G}(w)}$ (so, $x' = y' - 1$ and $a_{y'} = \overline{a}$): By $\overline{a}^H = \overline{a}^{\mathsf{G}(w)}$ (since $\overline{a} \notin X$).
  - $\langle x, y\rangle \in a^{\mathsf{G}(w)}$ (so, $x = y - 1$ and $a_y = a$): If not, then by $a^H = a^{\mathsf{G}(w)} \cup 1^H$, we have $[x]_{1^H} = [y]_{1^H}$. Thus, $\langle y', y' - 1\rangle \in 1^H (\subseteq \top^H)$. By the definition of $\top^H$, we have $\neg\psi^X_{a_{y'},\overline{1}}(\mathcal{U}_{y'-1}, U_{y'-1}, \mathcal{U}_{y'}, U_{y'})$. By the definition of $a^{J^{\mathcal{S}_X}}$, we have $\psi^X_{a_{y'},1}(\mathcal{U}_{y'-1}, U_{y'-1}, \mathcal{U}_{y'}, U_{y'})$, so $\overline{a} \in X$. This contradicts $\overline{a} \notin X$.
  - $([x, x'] \cup [x', x]) \cap ([y, y'] \cup [y', y]) = \emptyset$ (so, $x = x'$ and $y = y'$): If not, then because the interval between x and x' and that between y and y' have an intersection, we have $[x]_{1^H} = [y]_{1^H}$. Then, in the same manner as above, we have $\overline{a} \in X$. This contradicts $\overline{a} \notin X$.

  Thus, we reach a contradiction, because $a = a_y = a_{y'} = \overline{a}$ (by $y = y'$). Hence, $a^{H^{\mathcal{Q}}} \cap \overline{a}^{H^{\mathcal{Q}}} = \emptyset$.
- (P-Sat): By the form of $J^{\mathcal{S}_X}$, we have $\mathcal{U}_x = \begin{cases} \mathcal{U}_{x-1} & (\langle x-1, x\rangle \in 1^H) \\ \mathcal{U}_{x-1} \cup U^2_{x-1} & (\langle x-1, x\rangle \in \overline{1}^H) \end{cases}$. Thus, $\mathcal{U}_y = \bigcup_{x;\langle x,y\rangle \in \overline{1}^H} U^2_x$ ($\bigstar$). By Proposition 6.7, this completes the proof.
- (P-Wit): For $1^J \in U_0$ and $2^J \notin U_n$, they are shown by the form of $J^{\mathcal{S}_X}$. For $\forall a \in \tilde{\mathbf{V}}_1, \forall \langle x, y\rangle \in a^H, \delta^J_a(U_x) \subseteq U_y$, we distinguish the following cases:
  - Case $a = 1$: Then we have
    * $U_x = U_y$: By $\langle x, y\rangle \in 1^H$ and the form of $J^{\mathcal{S}_X}$, we have the following: $\forall z \in [y+1, x]$, $\psi^X_{a_z,1}(\mathcal{U}_{z-1}, U_{z-1}, \mathcal{U}_z, U_z)$. Thus, $U_y = U_{y+1} = \dots = U_x$.

  - $\delta^J_{1}(U_x) \subseteq U_x$: By $\langle \mathcal{U}_x, U_x \rangle \in |J^{\mathcal{S}_X}|$.

  Hence, $\delta^J_{1}(U_x) \subseteq U_y$.

- Case $a = \overline{1}$: Let $z \in [x+1, y]$ be such that $\psi^X_{a_z, \overline{1}}(\mathcal{U}_{z-1}, U_{z-1}, \mathcal{U}_z, U_z)$ and $\forall z' \in [z+1, y], \neg\psi^X_{a_{z'}, \overline{1}}(\mathcal{U}_{z'-1}, U_{z'-1}, \mathcal{U}_{z'}, U_{z'})$. Then we have

$$\begin{aligned}\delta^J_{\overline{1}}(U_x) &\subseteq \delta^J_{\overline{1}}(\{u \mid \langle u, u\rangle \in \mathcal{U}_z\}) && \text{(by (}\bigstar\text{) and } \langle x, z\rangle \in \overline{1}^H \text{ (by } \langle z-1, z\rangle \in \overline{1}^H))\\ &\subseteq U_z && \text{(by } \psi^X_{a_z, \overline{1}}(\mathcal{U}_{z-1}, U_{z-1}, \mathcal{U}_z, U_z))\\ &\subseteq U_{z+1} = \cdots = U_y. && \text{(by the form of } J^{\mathcal{S}_X}, \psi^X_{a_{z'}, 1}(\mathcal{U}_{z'-1}, U_{z'-1}, \mathcal{U}_{z'}, U_{z'}))\end{aligned}$$

- Case $a = \top$: We distinguish the following two sub-cases:
  - Case $\langle x, y\rangle \in \overline{1}^H$: By the similar argument as Case $a = \overline{1}$.
  - Case $\langle x, y\rangle \in 1^H$: By the similar argument as Case $a = 1$, we have $U_x = U_y$ and $\delta^J_{\top}(U_x) \subseteq U_x$, and thus $\delta^J_{\top}(U_x) \subseteq U_y$.
- Case $a \in \{a, \overline{a} \mid a \in \mathbf{V}\}$: We distinguish the following sub-cases:
  - Case $\langle x, y\rangle \in \overline{1}^H$: By $\langle x, y\rangle \in a^H \cap \overline{1}^H = a^{\mathsf{G}(w)}$, we have $x = y - 1$ and $a_y = a$. Thus by $\psi^X_{a_y, \overline{1}}(\mathcal{U}_{y-1}, U_{y-1}, \mathcal{U}_y, U_y)$, we have $\delta^J_a(U_x) \subseteq U_y$.
  - Case $a \notin X$: By $a^H = a^{\mathsf{G}(w)}$, we have $x = y - 1$ and $a_y = a$. By the form of $J^{\mathcal{S}_X}$ with $\neg\psi^X_{a_y, 1}(\mathcal{U}_{y-1}, U_{y-1}, \mathcal{U}_y, U_y)$ (since $a_y \notin X$), we have $\psi^X_{a_y, \overline{1}}(\mathcal{U}_{y-1}, U_{y-1}, \mathcal{U}_y, U_y)$. Hence, $\delta^J_a(U_x) \subseteq U_y$.
  - Case $\langle x, y\rangle \in 1^H$ and $a \in X$: By the similar argument as Case $a = 1$, we have $U_x = U_y$ (by $\langle x, y\rangle \in 1^H$) and $\delta^J_a(U_x) \subseteq U_x$ (by $a \in X$). Thus, $\delta^J_a(U_x) \subseteq U_y$. ◀