Continuous Algebras with Hypotheses

Lukas Mulder
Radboud University, Nijmegen, The Netherlands

Damien Pous
CNRS, LIP, ENS de Lyon, France

Jana Wagemaker
Radboud University, Nijmegen, The Netherlands

Abstract

In the literature on Kleene algebra (KA), a number of variants have been proposed such as Kleene algebra with tests, commutative KA, bi-KA, and concurrent KA. The equational theories of some of these structures have then been studied in the presence of additional assumptions, called *hypotheses*. We propose a unifying framework encompassing all the previous structures, as well as regular tree languages. This is done by considering algebras ordered by complete lattices, where least fixpoints can be computed. We provide a canonical model consisting of closed languages, which we prove sound and complete with respect to all *continuous* models. Then we study quasi-equational axiomatisations. It is illusory to hope for a generic axiomatisation which would be sound and complete for all instances. Instead, we provide a generic axiomatisation which we prove sound and we setup tools that make it possible to get complete ones in a modular way, building on previous works from the literature. We showcase these tools by proving new completeness results for commutative KA, bi-KA, and regular tree languages, in each case extended with various hypotheses.

2012 ACM Subject Classification Theory of computation → Algebraic language theory

Keywords and phrases Kleene algebra, complete lattices, languages, completeness

Digital Object Identifier 10.4230/LIPIcs.CONCUR.2026.42

Related Version *Full Version with all proofs in appendices*: `https://hal.science/hal-05620463` [49]

Funding This work was partially supported by the Dutch research council (NWO) under grant no. VI.Veni.242.134 (VerHyp) and by the LABEX MILYON (ANR-10-LABX-0070).

Acknowledgements We thank Paul Brunet and Amina Doumane for early discussions on this work.

1 Introduction

Kleene algebras (KA) [27, 13] are structures with sequential composition, iteration and choice. They were first proposed as the algebra of regular expressions, and their axioms are sound and complete w.r.t. relational models and language models [5, 28, 35]. Their equational theory is decidable via automata algorithms, which makes it possible to design automation tactics in proof assistants [7, 34, 39]. In such applications, we often want to reason under assumptions, or hypotheses. Unfortunately, the Horn theory of Kleene algebras is undecidable [30, 36] and we may only automate reasoning under specific forms of hypotheses [11, 33, 21, 31].

In each case, proving completeness and extending decidability is a delicate issue. This has motivated the development of a general theory of *Kleene algebra with hypotheses* [31, 15, 41]. There, models of *closed languages* are shown to characterise the Horn theory of *star-continuous* Kleene algebras, and a notion of *reduction* makes it possible to approach completeness and decidability proofs in a principled and modular way.

Beside hypotheses, Kleene algebras have also been extended to deal with common programming or logical constructs: conditionals in Kleene algebra with tests (KAT) [29], mutable tests [20], concurrency in bi-KA [37], concurrent KA [22] and synchronous KA

(SKA) [44, 48], intersection [16], or converse [18, 3]. Here again, completeness proofs are notoriously hard, and while some of the above extensions do fit – with some effort – into the framework of KA with hypotheses (e.g., KAT, SKA and KA with converse), some of them do not, typically when operations such as parallel composition or intersection are added, or when the required closures do not preserve regularity. An extended framework was developed [25, 23, 47] to deal with hypotheses in concurrent versions of KA, but a unifying theory is still lacking.

We propose such a theory in the present paper. Before introducing this abstract theory, observe that the operations from KA can be organised into three categories:

1. dot ($\cdot$) and one (1), which form a monoid,
2. plus (+) and zero (0), which form a semilattice,
3. star, which is the least fixpoint of a function expressible from $\cdot, +$ and 1 ($e^* = \mu x.1 + e \cdot x$).

When moving from one variant of KA to another, we keep the semilattice operations, but we change the first layer and the set of allowed fixpoints. For instance, to get bi-KA, we use bimonoids and two fixpoint operators; to get commutative KA [13, 8], we simply require the monoid to be commutative. Moreover, the standard language models from the literature consist of sets of elements of the free algebraic structure from the first layer: words for KA, series-parallel pomsets for bi-KA, finite multisets for commutative KA.

Accordingly, our theory is parameterised by an arbitrary signature Σ and a set of equations E for the first layer. We define our syntax by adding plus and zero to the operations from Σ, as well as arbitrary fixpoints. We interpret the resulting expressions in *continuous models*, which are models of E over complete lattices and where all operations from Σ are continuous. In particular, we get such a model by considering languages of *atoms*: terms built only from Σ and quotiented by E. The main requirement here is that the equations in E should be linear and regular: each variable must appear exactly once on each side of an equation, so that the algebra of languages we obtain remains a model of E [45, Theorem 3]. We show that every interpretation factors through this language model, witnessing its canonicity (Theorem 3.15).

At this point we have captured all the variants discussed above in a uniform way (and more, such as regular tree languages [17] by taking the empty set for E). The next step consists of dealing with hypotheses. Following and generalising [15, 25, 41], we associate to every set H of hypotheses a closure operator H^* on languages. This operator makes it possible to refine the former language model to take hypotheses into account, and we prove the following characterisation (Theorem 4.1):

$$\mathcal{C}E \models H \to e \leq f \iff H^* \llbracket e \rrbracket_E \subseteq H^* \llbracket f \rrbracket_E \quad .$$

In words, an inequation $e \leq f$ between expressions follows from a set H of hypotheses in all continuous models of E if and only if the H-closed language of e is contained in that of f. This characterisation is powerful as it reduces the validity of a Horn sentence in all continuous models to a language-theoretic problem. For instance, the right-hand side of the above equivalence often turns out to be decidable via automata algorithms.

It remains to study potential axiomatisations. It is easy to come up with quasi-equational axiomatisations Q which are *sound* for all continuous models:

$$Q \vdash H \to e \leq f \implies \mathcal{C}E \models H \to e \leq f \quad .$$

In particular, the "naive" axiomatisation consisting of the equations in E, the fact that symbols from Σ distribute over finite joins, and the fact that fixpoint expressions are indeed least fixpoints, is always sound (Theorem 5.4 and Proposition 5.5).

Obtaining the converse implication for all Horn sentences (i.e., *Horn completeness*) is difficult, if not impossible [30, 31]. Thus we focus on specific sets H of hypotheses: we say that an axiomatisation Q is *complete for H (on E)* if for all expressions e, f, we have

$$Q \vdash H \to e \leq f \Longleftarrow \mathcal{C}E \models H \to e \leq f \enspace .$$

Actually, completeness is difficult to achieve even when H is empty. Ésik has proved that the aforementioned naive axiomatisation is complete when E is also empty [17]. Unfortunately, this axiomatisation is incomplete on monoids: there we need the axioms of Conway algebras [13, 35], (left-handed) Kleene algebras [28, 32, 14, 24], or Boffa algebras [6]. Still, we can import all these completeness results for the empty set of hypotheses in our framework, as well as the ones for bi-KA on series-parallel pomsets [37], commutative KA on finite multisets [38, 13], KA with converse on involutive words [18], and identity-free distributive Kleene lattices on series-parallel graphs [16].

There are also completeness results for specific non-empty sets of hypotheses, which fit our framework: on monoids, KA is complete for $\{e \leq 0\}$ [11, 29], for $\{a \leq w\}$ [31, 15] (with a one or a letter, and w a word), for $\{S = 1\}$ [15] (with S a sum of letters), and for $\{e \leq \top \mid \forall e\}$ and $\{e \leq \top,\ e \leq e\top e \mid \forall e\}$ [42, 43]; on involutive monoids, KA with converse is complete for $\{e \leq ee^\circ e \mid \forall e\}$ [18]; on pomsets, bi-KA is complete for $\{a \leq w \mid |w| \geq 1\}$ (with a a letter, w a word) and for the *exchange law* $\{(e \parallel f) \cdot (g \parallel h) \leq (e \cdot g) \parallel (f \cdot h) \mid \forall e, f, g, h\}$ [23], which yields completeness of concurrent KA [26]. To deal with the diversity of these completeness results and be able to combine some of them together, we follow the literature and develop an appropriate notion of *reduction* [25, 40, 41, 9].

Contributions. We propose a framework for studying least fixpoints in algebras on complete lattices. We propose a model of languages which is canonical for continuous models of sets of linear and regular equations (Section 3). We show how add hypotheses in such structures, by defining a notion of closed languages that capture their Horn theory (Section 4). We finally study axiomatisations (Section 5): we provide a naive axiomatisation which is always sound, as well as tools to establish completeness, encompassing all results from the literature (Section 6). We illustrate our framework by proving new completeness results for commutative KA, bi-KA, and regular tree languages in the presence of specific hypotheses (Section 7).

On the distinction between equations and hypotheses. In the proposed framework, equations (E) and hypotheses (H) are quite close from the end-user perspective. The main differences are the following:

- equations are rather constrained (universally quantified, only between terms, linear regular, not inequations), but they can be integrated into the language semantics in a direct manner;
- in contrast, hypotheses are arbitrary (possibly about specific variables, between expressions, not necessarily linear/regular, can be inequations), but they have to be dealt with using the machinery of closed languages.

2 Preliminaries

Given a set X, we write $\mathcal{P}(X)$ for the set of its subsets (its *powerset*), and X^n for its n-fold Cartesian product. Given a partial map f, an element x and a value y, we write $f, x \mapsto y$ for the partial map whose value is y at x, and the same as f on the other elements. We also may write $\vec{x} = x_1, \ldots, x_n$ for a sequence of elements of a set X.

An *algebra* for a signature Σ is a set A together with a function $s_A \colon A^n \to A$ for each symbol of arity n in Σ. Symbols of arity zero are *constants*; they are just elements in the algebras. When it is clear from the context, we omit the subscript in s_A. The *powerset lifting* $\mathcal{P}(A)$ of such an algebra A is the algebra obtained by lifting its operations to sets in a pointwise manner: for a symbol s of arity n and for all $L_i \in \mathcal{P}(A)$ we set

$$s_{\mathcal{P}(A)}(L_1, \dots, L_n) = \{s_A(a_1, \dots, a_n) \mid \forall i, a_i \in L_i\} \ .$$

Given two algebras A and B, a *homomorphism from A to B* is a function $f \colon A \to B$ such that $f(s_A(a_1, \dots, a_n)) = s_B(f(a_1), \dots, f(a_n))$ for all symbols s of arity n and all elements $a_1, \dots, a_n$ from A.

A *partial order* is a set S with a binary relation $\leq$ which is reflexive, transitive, and antisymmetric. The functions from a set into a partial order S, are partially ordered pointwise ($f \leq g \triangleq \forall x,\ f(x) \leq g(x)$). A function $f \colon S \to S'$ between two partial orders is *monotone* if $f(x) \leq f(y)$ in S' whenever $x \leq y$ in S. A *partially ordered algebra* is an algebra whose carrier is a partial order and whose operations are monotone.

A *semilattice* is a partial order S with a least element ($\bot$) and all binary joins ($\vee$). If every subset S' of S has a supremum $\bigvee S'$, then the semilattice is called a *complete lattice*, and we denote it by $(S, \bigvee)$. Every monotone function f on a complete lattice admits a least fixpoint, written μf, which is also its least pre-fixpoint [46]. The powerset $\mathcal{P}(X)$ of a set X, ordered by inclusion, is the free complete lattice on X in the sense that every function $f \colon X \to S$ from X to a complete lattice S extends to a (unique) complete lattice homomorphism $\bigvee f \colon \mathcal{P}(X) \to S$ by setting $\bigvee f(Y) \triangleq \bigvee_{y \in Y} f(y)$. A function f between complete lattices is *continuous* if it preserves all suprema: $f \bigvee Y = \bigvee_{x \in Y} f(x)$ for all Y.

3 Framework

3.1 Terms and expressions

We fix throughout the paper a signature Σ and a set X of *variables*.

▶ **Definition 3.1.** *We write* Term *for the set of* terms *with variables in X, which are defined by the following grammar.*

$$t ::= x \mid s(t_1, \dots, t_n) \qquad (x \in X,\ s^{(n)} \in \Sigma)$$

Note that terms are trees whose leaves are either constants from Σ or variables from X; they form the free algebra over X, in the sense that every *valuation* $\sigma \colon X \to A$ into an algebra A for Σ extends to a unique homomorphism $\widehat{\sigma} \colon \mathrm{Term} \to A$.

We extend terms to *expressions* by adding a constant 0, a binary symbol $+$, and a least fixpoint operator μ. For the latter, we assume a second set R of *recursion variables*, countably infinite and disjoint from X.

▶ **Definition 3.2.** Expressions *are defined by the following grammar*

$$e, f ::= x \mid s(e_1, \dots, e_n) \mid 0 \mid e + f \mid \mu y.e \qquad (x \in X \cup R,\ y \in R,\ s^{(n)} \in \Sigma)$$

The recursion variable y in the expression $\mu y.e$ is bound *in e. An expression is* closed *if all its recursion variables are bound. An x-*expression *is an expression containing at most one unbound recursion variable, x. We denote the set of closed expressions by* Expr.

Every term can be seen as a closed expression: we have a canonical embedding Term $\hookrightarrow$ Expr. Also note that closed expressions may contain variables in X, which are never bound, and that x-expressions over X can also be seen as closed expressions over $X \cup \{x\}$. This change of perspective is necessary for some proofs found in the appendix. However, throughout the main text we keep X fixed and omit it from notation for the sake of clarity.

A *(closed) substitution* is a partial map from variables in $X \cup R$ to closed expressions; given an expression e and a substitution ρ, we write $e\rho$ for the *application* of ρ to e:

$$0\rho \triangleq 0 \qquad x\rho \triangleq \begin{cases} \rho(x) & \text{if } \rho(x) \text{ is defined,} \\ x & \text{otherwise} \end{cases}$$

$$(e+f)\rho \triangleq e\rho + f\rho$$

$$(s(e_1,\dots,e_n))\rho \triangleq s(e_1\rho,\dots,e_n\rho) \qquad (\mu x.e)\rho \triangleq \mu x.e(\rho\backslash x) \enspace .$$

(In the last case, $\rho\backslash x$ is the substitution obtained from ρ by removing its value at x, if any.) The above definition captures at once the two main kinds of substitutions we need:

- we let θ range over substitutions of variables from X only;
- we write $e\{x \leftarrow f\}$ for substitutions of an x-expression e where the recursion variable $x \in R$ is substituted for f, resulting in a closed expression.

In order to interpret expressions, we use algebras ordered by complete lattices:

▶ **Definition 3.3.** *A* complete algebra *is an algebra partially ordered on a complete lattice.*

▶ **Definition 3.4.** *Let A be a complete algebra. To every expression e, we associate a monotone function $[e]\_ : ((X \cup R) \to A) \to A$, by structural induction on e:*

$$[0]_\rho \triangleq \bot \qquad [x]_\rho \triangleq \rho(x)$$

$$[e+f]_\rho \triangleq [e]_\rho \vee [f]_\rho \qquad [\mu x.e]_\rho \triangleq \mu a.[e]_{\rho,x\mapsto a} \enspace .$$

$$[s(e_1,\dots,e_n)]_\rho \triangleq s_A([e_1]_\rho,\dots,[e_n]_\rho)$$

Given a valuation $\sigma \colon X \to A$ and a closed expression e, we write $\widehat{\sigma}(e)$ for $[e]_{\sigma,R\mapsto\bot}$, which we call the interpretation of e under σ.

In $[\mu x.e]_\rho$, $a \mapsto [e]_{\rho,x\mapsto a}$ is a monotone function by induction; hence we can take its least fixed point in A. Moreover, since taking least fixed points is a monotone operation, $\rho \mapsto \mu a.[e]_{\rho,x\mapsto a}$ is also monotone, as required. We overload $\widehat{\sigma}$: when σ is a valuation into a complete algebra, $\widehat{\sigma}$ can be applied to terms and closed expressions; this is safe thanks to the following lemma:

▶ **Lemma 3.5.** *For all valuations $\sigma \colon X \to A$ into a complete algebra A, the following diagram commutes.*

$$\begin{tikzcd} \mathrm{Term} \arrow[r, hook] \arrow[dr, "\widehat{\sigma}"'] & \mathrm{Expr} \arrow[d, "\widehat{\sigma}"] \\ & A \end{tikzcd}$$

In the sequel we just call *expressions* the closed expressions. Conversely, we systematically call *open expressions* the expressions which are not assumed to be closed.

3.2 Models

We distinguish terms and expressions because we will interpret expressions as languages of terms quotiented by a set E of equations. A *(term) equation* is a pair of terms written $u = v$. Variables in equations are universally quantified: $u = v$ is *satisfied* in an algebra A if for all valuations $\sigma \colon X \to A$ in that algebra, we have $\widehat{\sigma}(u) = \widehat{\sigma}(v)$. A *model* of a set of equations is an algebra satisfying these equations.

▶ **Example 3.6.** *Monoids* are the class of models given by the signature $\left\{\cdot^{(2)}, 1^{(0)}\right\}$ and the equations $\{x \cdot 1 = x, 1 \cdot x = x, x \cdot (y \cdot z) = (x \cdot y) \cdot z\}$. If we add commutativity $(x \cdot y = y \cdot x)$ to the previous equations, we obtain *commutative monoids*. *Bimonoids* [4, 37] are given by the signature $\left\{\cdot^{(2)}, \|^{(2)}, 1^{(0)}\right\}$ together with monoid equations for $(\cdot, 1)$ and the commutative monoid equations for $(\|, 1)$. When E is empty, models are arbitrary algebras.

An equation $u = v$ is *linear* if every variable appears at most once on either side; it is *regular* if every variable appearing in u also appears in v and vice versa.

▶ **Theorem 3.7** ([45, Theorem 3],[19, 2])**.** *The class of models of a set E of equations is closed under powerset lifting if and only if all equations in E are linear and regular.*

▶ **Example 3.8.** Monoids, commutative monoids and bimonoids are defined by linear and regular equations, hence these classes of models are closed under powerset lifting.

In addition to the signature Σ and the variables X, we fix in the sequel a set E of linear and regular term equations. When used alone, *model* implicitly means model of E.

The equations in E generate an equational theory $\equiv_E$ on terms: the least congruence relation containing the closure of E under arbitrary term substitutions; it coincides with the set of equations satisfied in all models of E. We write $\mathcal{A}_E$ for the set of terms quotiented by $\equiv_E$, called *atoms*. Atoms form the free model of E over X: $\mathcal{A}_E$ is a model of E and every valuation $\sigma \colon X \to M$ into a model M of E extends to a unique homomorphism $\widehat{\sigma} \colon \mathcal{A}_E \to M$.

▶ **Example 3.9.** In the case of monoids, atoms are words with letters in X. For commutative monoids, atoms are finite multisets over X. For bimonoids, atoms are *series-parallel pomsets (partially ordered multiset)* [37, Lemma 8]. Finally, when E is empty, atoms are just terms.

Note that the $\widehat{\sigma}$ notation is overloaded again: when σ is a valuation into a model, $\widehat{\sigma}$ can be applied to both terms and atoms, which is safe thanks to the following lemma.

▶ **Lemma 3.10.** *For all valuations $\sigma \colon X \to M$ into a model M of E, the following diagram commutes:*

$$\begin{tikzcd}
\mathrm{Term} \arrow[d, two heads] \arrow[dr, "\widehat{\sigma}"] & \\
\mathcal{A}_E \arrow[r, "\widehat{\sigma}"'] & M
\end{tikzcd}$$

Given an atom a, we write $\underline{a}$ for some term in its equivalence class. We use this only in contexts where the specific choice of term is irrelevant, typically thanks to the above lemma.

Since E is linear and regular, $\mathcal{P}(\mathcal{A}_E)$ is a model of E by Theorem 3.7. Its elements are sets of atoms, which we call *languages*. The notion of language thus depends on E: e.g. for monoids, where atoms are words, languages are sets of words. We abbreviate $\mathcal{P}(\mathcal{A}_E)$ as $\mathcal{L}_E$.

▶ **Definition 3.11.** *A* continuous model of E *is a complete algebra which is a model of E and whose operations are continuous in each argument: for all symbols s of arity $n+1$, for all elements $a_1, \ldots, a_n$, and for all subsets L of elements, we have*

$$s(a_1, \ldots, \bigvee L, \ldots, a_n) = \bigvee_{a \in L} s(a_1, \ldots, a, \ldots, a_n) \ .$$

▶ **Proposition 3.12.** *For all models M of E, $\mathcal{P}(M)$ is a continuous model of E.*

In particular, $\mathcal{L}_E$ is a continuous model of E, which we call the *standard language model of E*. Expressions can be interpreted in this model by mapping variables to singleton languages.

▶ **Definition 3.13.** *The* standard language interpretation *of an expression e is the language defined by $[\![e]\!]_E \triangleq \widehat{\eta}(e)$ for $\eta\colon x \mapsto \{x\}$, where x in $\{x\}$ is the equivalence class of x in $\mathcal{A}_E$.*

In the case of monoids, when restricting fixpoint expressions to be of the shape $\mu x.1 + e \cdot x$, the function $[\![\_]\!]_E$ coincides with the standard language interpretation of regular expressions.

Given a valuation $\sigma\colon X \to M$ into a complete model M, we have defined its extension $\widehat{\sigma}$ to both expressions and atoms; we also define its extension to languages, as follows:

$$\begin{aligned} \overline{\sigma}\colon \mathcal{L}_E &\to M \\ L &\mapsto \bigvee_{a \in L} \widehat{\sigma}(a) \end{aligned}$$

We establish properties of such extensions to languages in [49, Appendix B]; in particular:

▶ **Proposition 3.14.** *For all valuations $\sigma\colon X \to M$ into a continuous model M, the function $\overline{\sigma}$ is a continuous homomorphism from $\mathcal{L}_E$ to M.*

The main result of this section is that all interpretations factor through the standard language interpretation:

▶ **Theorem 3.15.** *For all valuations $\sigma\colon X \to M$ into a continuous model M of E, and for all expressions e, we have $\widehat{\sigma}(e) = \overline{\sigma}\,[\![e]\!]_E$. In other words, the following diagram commutes:*

$$\begin{tikzcd} & \mathrm{Expr} \arrow[dl, "{[\![\_]\!]_E}"'] \arrow[dr, "\widehat{\sigma}"] & \\ \mathcal{L}_E \arrow[rr, "\overline{\sigma}"'] & & M \end{tikzcd}$$

Theorem 3.15 allows us to decompose $\widehat{\sigma}(e)$ using the language $[\![e]\!]_E$. This makes later proofs easier, since we can focus on atoms rather than expressions. We also obtain as a corollary the following characterisation of the inequational theory of continuous models:

▶ **Corollary 3.16.** *The standard language interpretation is sound and complete with respect to continuous models of E: for all expressions e, f,*

$$\mathcal{C}E \models e \leq f \iff [\![e]\!]_E \subseteq [\![f]\!]_E \quad .$$

3.3 Inequations and Horn sentences

An *inequation* is a pair of expressions, written $e \leq f$. A *Horn sentence* $H \to e \leq f$ is a set of inequations H called the *hypotheses* together with an inequation $e \leq f$ called the *conclusion*. Note that we allow Horn sentences with infinitely many hypotheses.

▶ Remark 3.17. In the present work it is more convenient to put the emphasis on inequations rather than on equations, but both are interderivable: an equation $e = f$ can be encoded as the conjunction of $e \leq f$ and $f \leq e$, and thanks to binary joins, an inequation $e \leq f$ can be encoded as the equation $e + f = f$.

▶ **Definition 3.18.** *Let H be a set of hypotheses, and let e, f be two expressions. A valuation $\sigma\colon X \to M$ into a complete algebra* satisfies *the inequation $e \leq f$ if $\widehat{\sigma}(e) \leq \widehat{\sigma}(f)$; it satisfies H if it satisfies all inequations in H. We write $M \models H \to e \leq f$ if all valuations $\sigma\colon X \to M$ satisfying H also satisfy $e \leq f$. We write $\mathcal{C}E \models H \to e \leq f$ if the above holds for all continuous models M of E. When H is empty, we write $\models e \leq f$ instead of $\models \emptyset \to e \leq f$.*

Note how the universal quantification on all valuations σ is shared by the hypotheses H and the conclusion $e \leq f$: we are interested in the universal Horn theory and when the signature contains a binary operation $\cdot$, we typically have $\mathcal{C}E \models x \cdot x \leq x \to x \cdot (x \cdot x) \leq x$, but not $\mathcal{C}E \models x \cdot x \leq x \to y \cdot (y \cdot y) \leq y$.

By Theorem 3.15 we can readily reduce the analysis of the Horn theory to the case where the left-hand side of the conclusion is a term rather than an arbitrary expression.

▶ **Corollary 3.19.** *For all continuous models M of E, for all hypotheses H, and for all expressions e, f, we have: $M \models H \to e \leq f \iff \forall a \in [\![e]\!]_E\,,\ M \models H \to \underline{a} \leq f$.*

4 Closed language interpretation

We fix in the next two sections a set H of hypotheses. We define an operator $H^*\colon \mathcal{L}_E \to \mathcal{L}_E$ which we use to alter the standard language interpretation to take H into account. We then prove soundness and completeness of this new interpretation, relatively to H:

▶ **Theorem 4.1.** *For all expressions e, f, we have:*

$$\mathcal{C}E \models H \to e \leq f \iff H^*\,[\![e]\!]_E \subseteq H^*\,[\![f]\!]_E \quad .$$

When H is empty, in which case H^* is the identity, we recover precisely Corollary 3.16.

4.1 Contexts

We need a notion of *context* to define the operator H^*. To this end, we fix a designated variable $- \in X$, which we call the *hole*. A *term context* is a term c with exactly one occurrence of the hole. Given such a term c, we write cb for the term obtained from c by replacing the hole with a term b; similarly, we write ce for the expression obtained from c by replacing the hole with an expression e. A *context* is an atom C such that $\underline{C}$ is a term context. This is well-defined: since E is linear and regular, all terms in the equivalence class of C have the same number of occurrences of a given variable. Given a context C and an atom a, we write Ca for the equivalence class of $\underline{C}\,\underline{a}$ (which is an atom). We extend this notation to languages: given a context C and a language L, we write CL for $\{Ca \mid a \in L\}$.

▶ **Example 4.2.** In monoids, where atoms are words, a context is just a pair (l, r) of words; the application of such a context to a word u is the composite word lur; its application to the language $\{a, b\}$ is the language $\{lar, lbr\}$. In commutative monoids, where atoms are multisets, a context is just a multiset m, and the application of such a context to a multiset n is the multiset union $m \cup n$. In bimonoids, the hole of a context may appear under arbitrary alternations of sequential $(\cdot)$ and parallel $(\|)$ compositions.

▶ **Lemma 4.3.** *For all contexts C and expressions e, we have $C\,[\![e]\!]_E = [\![\underline{C}e]\!]_E$.*

4.2 Closure under hypotheses

In the sequel we often omit the subscript E from $[\![e]\!]$ and $\mathcal{L}$ to alleviate notation. Recall the set H of hypotheses; we turn it into a monotone function $H\colon \mathcal{L} \to \mathcal{L}$ on languages by setting

$$H(L) \triangleq \bigcup \{C\,[\![e]\!] \mid C \text{ a context},\ (e \leq f) \in H,\ C\,[\![f]\!] \subseteq L\} \quad .$$

▶ **Definition 4.4.** *A language L is H-*closed *if $H(L) \subseteq L$. The H-*closure *of a language L is the least H-closed language $H^*(L)$ containing L.*

Equivalently, H^* is the least closure above the function H (cf. [49, Appendix G]). Also note that L is H-closed if and only if $H^*(L) = L$, and that H^* is the identity function when H is empty. Below we sometimes write H^*L for $H^*(L)$ and HL for $H(L)$.

▶ **Example 4.5.** For monoids, if $H = \{x \leq y + z\}$ and $L = \{y, ayb, azb\}$, then $HL = \{axb\}$ and $H^*L = L \cup \{axb\}$; if $H = \{xx \leq x\}$ and $L = \{xx, xyx\}$, then $HL = \{xxx, xxyx, xyxx\}$ and $H^*L = \{x^i, x^jyx^k \mid i \geq 2, j, k \geq 1\}$. For commutative monoids, if $H = \{c \leq xy, a \leq cd\}$ and $L = \{ax, by, dxy\}$[1], then $HL = \{dc\} = \{cd\}$, and $H^*L = L \cup \{a, cd\}$. Finally, in bimonoids, if $H = \{b \leq a \parallel c\}$ and $L = \{x \cdot (a \parallel b \parallel c)\}$, then $H^*L = L \cup \{x \cdot (b \parallel b)\}$.

The main point of H-closed languages is that they validate the inequations in H:

▶ **Lemma 4.6.** *For all $(e \leq f) \in H$, we have $H^* [\![e]\!] \subseteq H^* [\![f]\!]$.*

4.3 Soundness and completeness

To prove Theorem 4.1, we first show that $H^* [\![\_]\!]$ is an interpretation into a continuous model. We write $H^*\mathcal{L}$ for the set of H-closed languages, ordered by inclusion. In general, the function H^* does *not* preserve unions and is *not* a homomorphism from $\mathcal{L}$ to $\mathcal{L}$. Therefore its image $H^*\mathcal{L}$ is not a model *as is*. To get a model of closed languages, we need to adapt the operations and close them explicitly, as follows.

For all collections $(L_i)_{i \in I}$ of H-closed languages, and for all symbols s of arity n, set:

$$\bigvee_{i \in I} L_i \triangleq H^* \left(\bigcup_{i \in I} L_i \right) \qquad \begin{aligned} s_{H^*\mathcal{L}} \colon (H^*\mathcal{L})^n &\to H^*\mathcal{L} \\ (L_1, \dots, L_n) &\mapsto H^*(s_{\mathcal{L}}(L_1, \dots, L_n)) \end{aligned}$$

▶ **Proposition 4.7.** *The set $H^*\mathcal{L}$ of H-closed languages, equipped with the above operations, is a continuous model of E.*

It remains to show that $H^* [\![\_]\!]$ actually is an interpretation into that model:

▶ **Proposition 4.8.** *Let $\eta_H \colon X \to H^*\mathcal{L}$ be the valuation defined by $\eta_H(x) \triangleq H^* \{x\}$. For all expressions e we have $H^* [\![e]\!] = \widehat{\eta_H}\,(e)$.*

We can finally prove our main theorem.

Proof of Theorem 4.1. The left to right implication follows directly from Proposition 4.8 and the fact that our interpretation satisfies all hypotheses in H (Lemma 4.6).

For the converse implication, assume $H^* [\![e]\!] \subseteq H^* [\![f]\!]$ and pick a continuous model M; we have to show $M \models H \to e \leq f$. Define the following language:

$$F \triangleq \{a \in \mathcal{A} \mid M \models H \to \underline{a} \leq f\} \ .$$

By Corollary 3.19, we have that for all expressions g,

$$[\![g]\!] \subseteq F \iff M \models H \to g \leq f \ . \qquad (\star)$$

Thus it suffices to show $[\![e]\!] \subseteq F$. By assumption, this follows from $H^* [\![f]\!] \subseteq F$, which in turn follows from $[\![f]\!] \subseteq F$ and $H(F) \subseteq F$, by definition of H^*. We get $[\![f]\!] \subseteq F$ directly from $(\star)$, and it only remains to show $H(F) \subseteq F$.

[1] We write finite multisets as words where the ordering of the letters is irrelevant.

$$\frac{\forall (e' \leq f') \in H',\ Q \vdash H \to e'\theta \leq f'\theta}{Q \vdash H \to e\theta \leq f\theta}(H' \to e \leq f) \in Q,\ \theta \colon X \rightharpoonup \mathbb{F}$$

$$\frac{}{Q \vdash H \to e \leq f}(e \leq f) \in H$$

$$\frac{}{Q \vdash H \to e \leq e}e \in \mathbb{F} \qquad \frac{Q \vdash H \to e \leq f \qquad Q \vdash H \to f \leq g}{Q \vdash H \to e \leq g}$$

$$\frac{Q \vdash H \to e_1 \leq f_1 \qquad Q \vdash H \to e_2 \leq f_2}{Q \vdash H \to e_1 + e_2 \leq f_1 + f_2} \qquad \frac{\forall i,\ Q \vdash H \to e_i \leq f_i}{Q \vdash H \to s(e_1, \dots, e_n) \leq s(f_1, \dots, f_n)}s^{(n)} \in \Sigma$$

Figure 1 Deduction rules for the Horn theory of an axiomatisation Q over a fragment $\mathbb{F}$.

Fix a context C, a hypothesis $(e_0 \leq f_0) \in H$, and suppose $C[\![f_0]\!] \subseteq F$ (†). We have to show $C[\![e_0]\!] \subseteq F$. From Lemma 4.3 and ($\star$), we have that for all expressions g,

$$C[\![g]\!] \subseteq F \iff (\forall \sigma \text{ satisfying } H \text{ in } M,\ \widehat{\sigma}(\underline{C}\, g) \leq \widehat{\sigma}(f)) \ . \tag{$\star\star$}$$

Accordingly, fix a valuation $\sigma \colon X \to M$ satisfying H. Since we have $(e_0 \leq f_0) \in H$, we know $\widehat{\sigma}(e_0) \leq \widehat{\sigma}(f_0)$. We derive $\widehat{\sigma}(\underline{C}e_0) \leq \widehat{\sigma}(\underline{C}f_0) \leq \widehat{\sigma}(f)$ first by monotonicity and then using (†) with ($\star\star$). We conclude using ($\star\star$) again. ◀

5 Axiomatisations

In this section we study axiomatisations. We restrict to quasi-equational ones, which consist only of Horn sentences. In order to deal with standard examples from the literature, we study those axiomatisations on arbitrary fragments of our general syntax of expressions.

▶ **Definition 5.1.** *A* fragment *is a subset $\mathbb{F}$ of expressions containing* 0 *and the variables, closed under $+$, the operations from Σ, application of substitutions $\theta \colon X \rightharpoonup \mathbb{F}$, (closed) sub-expressions, and such that $e\{x \leftarrow f\}$ belongs to $\mathbb{F}$ whenever $\mu x.e$ and f belong to $\mathbb{F}$.*

▶ **Example 5.2.** The set of all (closed) expressions and the set of fixpoint-free expressions is are fragments. On monoids, the set of expressions where all fixpoint sub-expressions are of the form $\mu x.1 + e \cdot x$ is a fragment, which corresponds to regular expressions. On bimonoids, the set of expressions where all fixpoint sub-expressions are either of the form $\mu x.1 + e \cdot x$ or of the form $\mu x.1 + e \parallel x$ is a fragment, which corresponds precisely to bi-KA expressions [37].

▶ **Definition 5.3.** *A* (quasi-equational) axiomatisation over a fragment *is a set of Horn sentences expressed in that fragment. Given such an axiomatisation Q, its* Horn theory *is the judgement $Q \vdash H \to e \leq f$ inductively defined by the rules in Figure 1, where H is a set of hypotheses in the fragment.*

Regarding variables and quantifications, such a judgement should intuitively be understood as the following implication:

$$\text{“}(\forall \vec{x},\ Q) \Rightarrow \forall \vec{x},\ (H \Rightarrow e \leq f)\text{”} \ .$$

$$
\begin{array}{rcl}
0 \leq x & s(\vec{x}, 0, \vec{z}) \leq 0 & s^{(n+1)} \in \Sigma \\
x + x \leq x & s(\vec{x}, y + y', \vec{z}) \leq s(\vec{x}, y, \vec{z}) + s(\vec{x}, y', \vec{z}) & s^{(n+1)} \in \Sigma \\
x \leq x + y & u \leq v & (u = v) \in E \\
y \leq x + y & v \leq u & (u = v) \in E
\end{array}
$$

$$
\begin{array}{rl}
e\{x \leftarrow \mu x.e\} \leq \mu x.e & \mu x.e \in \mathbb{F} \\
e\{x \leftarrow f\} \leq f \to \mu x.e \leq f & \mu x.e, f \in \mathbb{F}
\end{array}
$$

Figure 2 Naive axiomatisation $N(E, \mathbb{F})$ on equations E over a fragment $\mathbb{F}$.

This is why we use a substitution in the first rule of Figure 1, which makes it possible to use any instance of some axiom in Q, but not in in the second one, which only makes it possible to use an hypothesis from H, *as is*. The next two rules enforce reflexivity and transitivity, while the last two rules ensure monotonicity of $+$ and the operations from Σ.

We establish the basic properties of these deduction rules in [49, Appendix D]. Since they preserve validity in all complete models, we get:

▶ **Theorem 5.4.** *Let Q be an axiomatisation and let M be a complete model. If $M \models Q$ and $Q \vdash H \to e \leq f$ then $M \models H \to e \leq f$.*

We provide a generic axiomatisation in Figure 2, which we call *naive*. Given a fragment $\mathbb{F}$, this axiomatisation $N(E, \mathbb{F})$ comprises axioms ensuring that we get a semilattice with $+$ and 0, that all operations from Σ distribute over finite joins, that the equations from E are satisfied, and that fixpoint expressions that are allowed in $\mathbb{F}$ are indeed least (pre)fixpoints.

As expected, this naive axiomatisation is sound for all continuous models:

▶ **Proposition 5.5.** *For all fragments $\mathbb{F}$, we have $\mathcal{C}E \models N(E, \mathbb{F})$.*

Combined with Theorem 5.4, we get that all Horn sentences derivable from $N(E, \mathbb{F})$ are valid in all continuous models of E.

In the case of trees, i.e., when E is empty, and when considering all expressions ($\mathbb{F} = \mathrm{Expr}$), Ésik has proved that the naive axiomatisation is complete for the empty set of hypotheses:

▶ **Theorem 5.6** ([17, Theorem 21 with Remark 27])**.** *For all expressions e, f we have*

$$N(\emptyset, \mathrm{Expr}) \vdash e \leq f \iff [\![e]\!]_\emptyset \subseteq [\![f]\!]_\emptyset \iff \mathcal{C}\emptyset \models e \leq f \ .$$

Unfortunately, this result does not generalise to arbitrary equations E (and even less to arbitrary hypotheses H, as those can be used to emulate equations).

▶ **Example 5.7.** Consider for instance monoids and regular expressions, hence Kleene algebra. A counter-model can be used to show that the naive axiomatisation cannot derive $a^* \cdot a \leq a^*$ (see [49, Counterexample D.4], where we provide a complete model of the naive axiomatisation which is not continuous and refutes this law, although $[\![a^* \cdot a]\!] \subseteq [\![a^*]\!]$). Instead, several axiomatisations have been proposed by Conway [13], which were later proved complete by Krob [35], Kozen [28] and Boffa [5, 6]. Amongst them we find *left-handed Kleene algebra* [32, 14, 24], which is close to the naive axiomatisation, except that the least fixpoint induction axiom from the naive axiomatisation, which can be reformulated here as $1 + e \cdot f \leq f \to e^* \leq f$, is strengthened into $h + e \cdot f \leq f \to e^* \cdot h \leq f$. The resulting ability to perform inductions on stars "under a context h on the right" is crucial for completeness.

We conclude this section by showing that there are nevertheless situations where the naive axiomatisation is enough in the presence of both equations and hypotheses: when some of the expressions are fixpoint-free. We fix a fragment $\mathbb{F}$ in the remainder of this section.

First we show that the naive axiomatisation proves memberships:

▶ **Proposition 5.8.** *For all expressions $e \in \mathbb{F}$ and for all $a \in [\![e]\!]_E$, we have $N(E, \mathbb{F}) \vdash \underline{a} \leq e$.*

We deduce that fixpoint-free expressions can be rewritten as finite sums of terms:

▶ **Lemma 5.9.** *For all fixpoint-free expressions e, $[\![e]\!]_E$ is finite and $N(E, \mathbb{F}) \vdash e = \sum_{a \in [\![e]\!]} \underline{a}$.*

It follows that the naive axiomatisation is complete for inequations whose left-hand side is fixpoint-free, when the right-hand sides of the hypotheses are also fixpoint-free.

▶ **Theorem 5.10.** *Let $H = \{e_i \leq f_i \mid i \in I\}$ be a set of hypotheses and let e, f be expressions. If e and all f_i are fixpoint-free, then we have*

$$N(E, \mathbb{F}) \vdash H \to e \leq f \iff H^* [\![e]\!]_E \subseteq H^* [\![f]\!]_E \quad .$$

6 Reductions

We have seen that the naive axiomatisation (Figures 1 and 2) is sound with respect to the interpretation $H^* [\![\_]\!]$ but not always complete. Nevertheless, there are complete axiomatisations in the literature. That is, for specific signatures Σ, sets of equations E, and sets of hypotheses H, there are axiomatisations Q, such that for all expressions e, f, we have

$$Q \vdash H \to e \leq f \quad \Longleftarrow \quad H^* [\![e]\!]_E \subseteq H^* [\![f]\!]_E \quad . \tag{†}$$

Following [41], we introduce *reductions* in this section to derive new completeness results from existing ones – typically, to derive that Q is complete for a given set H of hypotheses from the fact that Q is complete for the empty set of hypotheses.

In full generality, reductions are best presented at the fairly abstract level of *representations* [9], which we show how to instantiate in the context of the present paper.

▶ **Definition 6.1.** *A* representation *is tuple $(\mathbb{F}, M, [\_], \equiv)$ where $\mathbb{F}$ and M are two sets, $\equiv$ is an equivalence relation on $\mathbb{F}$, and $[\_]$ is a function from $\mathbb{F}$ to M. Such a representation is* complete *if $[e] = [f]$ implies $e \equiv f$ for all $e, f \in \mathbb{F}$.*

In our case, representations will always consist of a fragment $\mathbb{F}$, the set of H-closed languages, the interpretation $H^* [\![\_]\!]$, and the equivalence induced by an axiomatisation Q and the hypotheses H: $e \equiv f$ if $Q \vdash H \to e \leq f$ and $Q \vdash H \to f \leq e$. For these representations we write (Q, H); they are complete precisely when Q is complete for H in the sense of (†).

▶ **Definition 6.2.** *A representation $(\mathbb{F}, M, [\_], \equiv)$ reduces to a representation $(\mathbb{F}', M', [\_]', \equiv')$ if there exists three functions (r, i, h) typed as below such that for all expressions $e \in \mathbb{F}$ and $e', f' \in \mathbb{F}'$, 1) $e' \equiv' f'$ implies $i(e') \equiv i(f')$; 2) $i(r(e)) \equiv e$; 3) $h([e]) = [r(e)]'$.*

$$\begin{tikzcd}
\mathbb{F} \arrow[r, "{[_]}"] \arrow[d, shift left, "r"] & M \arrow[d, "h"] \\
\mathbb{F}' \arrow[r, "{[_]'}"] \arrow[u, shift left, "i"] & M'
\end{tikzcd}$$

As expected, reductions can be composed, and we have:

▶ **Theorem 6.3.** *If a representation reduces to a complete one, then the former is complete.*

Most of the tools provided in [41] for creating reductions and combining existing ones can be adapted to our general framework, which we do in [49, Appendix E]. We highlight one of them here, which we need for one of our examples: reductions generated by substitutions, which generalise the *homomorphic reductions* from [41, Proposition 3.2].

▶ **Proposition 6.4.** *Let (Q, H) and (Q', H') be representations as above over $\mathbb{F}$, such that Q is sound for $\mathcal{C}E$ and derives all axioms in $N(E, \mathbb{F})$. If $\theta \colon X \rightharpoonup \mathbb{F}$ is a substitution such that:*
1. *for all expressions $e, f \in \mathbb{F}$, $Q' \vdash H' \to e \leq f$ implies $Q \vdash H \to e \leq f$,*
2. *for all variables x in the domain of θ, $Q \vdash H \to x = \theta(x)$,*
3. *for all variables x in the domain of θ, $x \in [\![\theta(x)]\!]$,*
4. *and for all $(e \leq f) \in H$, $[\![e\theta]\!] \subseteq H'^{*}\,[\![f\theta]\!]$,*

then (Q, H) reduces to (Q', H').

Proof Sketch. We take $(\_)\theta$ for r, the identity for i and an inclusion for h. ◀

7 Examples

We finally show how various structures and results from the literature fit our framework, and we prove new completeness results for specific forms of hypotheses.

7.1 Kleene algebra: monoids

Our starting point was Kleene algebra with hypotheses [15, 41], which we capture by choosing the signature and equations of monoids for Σ and E: in that case atoms are just words, and languages are as usual. When it comes to axiomatisations, we chose the fragment where all fixpoints must have the shape $\mu x.1 + e \cdot x$, and we may pick any of the various axiomatisations proposed in the literature [13, 5, 35, 28, 6, 32, 14, 24], which are all sound, and complete for the empty set of hypotheses.

For hypotheses, our notion of language closure coincides with that from [15, 41], and the reductions proposed in the literature fit our definition and make it possible to reach KA with tests [29], with converse [18, 3, 10], with top [42, 43], with observations [41], or even NetKAT [1, 41] and synchronous KA [44, 48]. (Note that some of these reductions depend on the specific choice of axiomatisation for KA without hypotheses; e.g., the two items of [41, Proposition 3.8(iv)] respectively require left-handed KA and right-handed KA.)

More concretely, we recall here a simple example [41, Proposition 3.8(v)], which we will transfer to a new setting in the following section. This example shows that we can eliminate "transitivity" hypotheses of the shape $a \cdot a \leq a$, in KA.

▶ **Lemma 7.1.** *$(KA, \{a \cdot a \leq a\})$ reduces to $(KA, \emptyset)$.*

Proof. The substitution $a \mapsto a \cdot a^*$ satisfies the conditions of Proposition 6.4: the first one vanishes as $Q = Q' = \mathrm{KA}$ and $H' = \emptyset$; we can easily derive $a = a \cdot a^*$ from KA and the hypothesis $a \cdot a \leq a$; we have $a \in [\![a \cdot a^*]\!]$; and we have $[\![(a \cdot a^*) \cdot (a \cdot a^*)]\!] \subseteq [\![a \cdot a^*]\!]$. ◀

7.2 Bi-Kleene algebra: bimonoids

We also cover bi-KA with hypotheses [25, 23, 47]: we chose the signature and equations of bimonoids, atoms become series-parallel pomsets, and our notion of language closure coincides with that from [25, 23, 47].

Regarding syntax, there is a choice: while Struth and Laurence [37] prove completeness for bi-KA with two fixpoint operators (the sequential one from KA, and its parallel counterpart:

$e^{\|} \triangleq \mu x.1 + e \parallel x$), Kappe et al. [25, 23, 47] develop the theory of hypotheses and reductions only with the sequential one. This restriction is necessary for the reduction they exhibit to deal with the exchange law and prove completeness of concurrent KA [22, 26]; it is unclear how to extend this reduction in the presence of unbounded parallelism.

We cover both choices with the present framework, by selecting the appropriate fragment. If we choose bi-KA with the two fixpoint operators [37], for which hypotheses were not studied in the literature before, we can redo Lemma 7.1 with both monoid compositions, in a straightforward manner.

▶ **Lemma 7.2.** *$(biKA, \{a \cdot a \leq a\})$ and $(biKA, \{a \parallel a \leq a\})$ both reduce to $(biKA, \emptyset)$.*

Proof. We use Proposition 6.4 with the substitutions $a \mapsto a \cdot a^*$ and $a \mapsto a \parallel a^{\|}$. ◀

Modularity tools from [41], which we have adapted in [49, Appendix E], make it possible to deduce that (biKA, $\{a \cdot a \leq a, b \parallel b \leq b\}$) combinedly reduces to (biKA, $\emptyset$), when a and b are distinct variables (see details in [49, Appendix F.2]). In contrast, there cannot be a reduction from (biKA, $\{a \cdot a \leq a, a \parallel a \leq a\}$) to (biKA, $\emptyset$): this would require an expression for all non-empty series-parallel pomsets over $\{a\}$, which is not available in the considered fragment.

7.3 Regular tree languages

When E is empty, atoms are just terms, or ordered trees with nodes labelled in Σ and of appropriate arities. In that case, there is a Kleene theorem [12] ensuring that our expressions and finite tree automata capture the same class of tree languages: the regular ones (note that we need all fixpoint expressions in order to obtain this result, unlike with languages of words). Moreover, as mentioned in Section 5, Ésik [17] has proved that the naive axiomatisation $N(\emptyset, \mathrm{Expr})$ is complete for the empty set of hypotheses in that case (Theorem 5.6).

Hypotheses were not studied in the literature in that setting, and we show below how to eliminate the counterpart to *Hoare hypotheses*, from [21] in the context of KAT.

▶ **Lemma 7.3.** *Let e_0 be an expression. If both the signature Σ and the set X of variables are finite, then $(N(\emptyset, \mathrm{Expr}), \{e_0 \leq 0\})$ reduces to $(N(\emptyset, \mathrm{Expr}), \emptyset)$.*

Proof Sketch. Using the aforementioned Kleene theorem, let $\top$ be an expression for the language of all trees over Σ and X. Define the following expression:

$$e_0^\top := \mu x. e_0 + \sum_{s^{(n)} \in \Sigma,\ n \geq 1} s(x, \top, \dots, \top) + s(\top, x, \dots, \top) + \dots + s(\top, \top, \dots, x) \ .$$

The language $[\![e_0^\top]\!]$ consists of all trees containing a tree from $[\![e_0]\!]$ as a subtree, and the function $e \mapsto e + e_0^\top$ yields the desired reduction. We give more details in [49, Appendix F]; the key observations are that $e_0^\top$ is provably equivalent to 0 using the naive axioms and the hypothesis $e_0 \leq 0$, and that writing H for $\{e_0 \leq 0\}$, we have $H^* = L \mapsto L \cup [\![e_0^\top]\!]$. ◀

7.4 Commutative Kleene algebra: multisets

Commutative Kleene algebras (cKA) are obtained from KA by postulating that $(\cdot)$ is commutative. Pilling has proved their completeness w.r.t. the Parikh image interpretation of regular expressions [38, 13, 8]. It is tempting to deal with cKA by using plain KA with the hypotheses $C = \{xy \leq yx \mid x, y \in X\}$. Indeed those hypotheses imply that any two expressions commute, and C^* is the Parikh image function in that case. However, we have

$C^* [\![(ab)^*]\!] = \{w \in \{a,b\}^* \mid |w|_a = |w|_b\}$, so that C^* does not preserve regularity. Thus, there cannot be a reduction from (KA, C) to $(\mathrm{KA}, \emptyset)$, and we may hardly find reductions to (KA, C). Instead, we can impose the commutativity condition via the set E of equations, and work directly with commutative monoids, for which atoms are finite multisets.

Below we provide three new reductions; the first one is about Hoare hypotheses.

▶ **Lemma 7.4.** *Let e_0 be an expression. For finite X, $(\mathrm{cKA}, \{e_0 \leq 0\})$ reduces to $(\mathrm{cKA}, \emptyset)$.*

Proof sketch. Let $\top$ be an expression for the language of all multisets (for which we need X to be finite). The function $e \mapsto e + e_0 \cdot \top$ yields the desired reduction. ◀

The second one looks similar but has no counterpart in the literature: while it is relatively easy to obtain in the commutative case, we cannot obtain such a reduction in the non-commutative case (e.g., with $e_1 = a \cdot b$, regularity is not preserved).

▶ **Lemma 7.5.** *Let e_1 be an expression. $(\mathrm{cKA}, \{e_1 \leq 1\})$ reduces to $(\mathrm{cKA}, \emptyset)$.*

Proof sketch. The function $e \mapsto e \cdot e_1^*$ yields the desired reduction. ◀

The third one is a first step towards a counterpart to [41, Lemma 3.8(i)] for the commutative case, which is surprisingly much harder due to the lack of a Kleene theorem.

▶ **Lemma 7.6.** *$(\mathrm{cKA}, \{a \leq a \cdot a\})$ reduces to $(\mathrm{cKA}, \emptyset)$.*

Proof sketch. Let $H \triangleq \{a \leq a \cdot a\}$. For all languages L, we have

$$H^* L = L \cup \{a^k u \mid a^{k'} u \in L, 1 \leq k < k'\}.$$

In cKA, all expressions are provably equivalent to an expression of the form $e_0 + \cdots + e_n$, where e_i is of the form $w_0 w_1^* \ldots w_n^*$ and all w_i are distinct finite multisets [38, 13]. For H^*, $H^*(L \cup L') = H^*(L) \cup H^*(L')$ (cf. [41, Lemma 3.1] for a similar argument); hence we construct an e'_i for each e_i and define $r(e) = e'_0 + \cdots + e'_n$.

For e_i of the form $w_0 w_1^* \ldots w_n^*$, we treat the cases where $a \in w_0$, and $a \notin w_0$ separately. The latter is in the appendix. If $a \in w_0$, we use commutativity to obtain the equivalent expression $(u_0 a^{k_0})(u_1 a^{k_1})^* \cdots (u_m a^{k_m})^* v_0^* \cdots v_{m'}^*$ with each $k_i \geq 1$. Then we define

$$e'_i \triangleq (u_0 a a^{\leq k_0 - 1})(u_1 a^{\leq k_1})^* \cdots (u_m a^{\leq k_m})^* v_0^* \cdots v_{m'}^*,$$

where $a^{\leq k} \triangleq 1 + a + \cdots + a^k$.

We show that r yields a reduction in the appendix; here we provide intuition through an example. Let $e = aa(aa)^*$, whose language is $\{a^k \mid k \text{ even and } k \geq 2\}$ with H-closure $\{a^k \mid k \geq 1\}$. Using the naive axioms we derive that $e' \triangleq (a(1+a))(1+a+aa)^* = (a+aa)(1+a+aa)^*$, and then it is easy to see that the language interpretation of e' coincides with the H-closure. We finally derive from the naive axioms and the hypothesis that $e = e'$:

$$aa(aa)^* \leq (a+aa)(1+a+aa)^* \leq (aa+aa)(1+aa+aa)^* = aa(1+aa)^* = aa(aa)^*$$

◀

Let us finally illustrate how we can assemble the previous reductions in a modular way. Using the tools from [49, Appendix E], we can strengthen the first two examples above into:

▶ **Lemma 7.7.** *For all set of hypotheses H and all expressions e, the sets $(\mathrm{cKA}, H \cup \{e \leq 0\})$ and $(\mathrm{cKA}, H \cup \{e \leq 1\})$ both reduce to (cKA, H).*

Since reductions can be composed, we deduce the following corollary:

▶ **Corollary 7.8.** *For all e_0, e_1, $(\mathrm{cKA}, \{e_0 \leq 0, e_1 \leq 1, a \leq a \cdot a\})$ reduces to $(\mathrm{cKA}, \emptyset)$.*

8 Conclusion

We have developed a framework to study algebras ordered by complete lattices in the presence of hypotheses. This framework captures existing work on Kleene algebra and bi-Kleene algebra with hypotheses, and can be used to study new structures in the presence of hypotheses. In particular, we instantiated our framework to regular tree languages and commutative KA and proved new completeness results for particular hypotheses. These structures had not been studied with additional hypotheses in the literature.

We conjecture that a completeness result can be obtained for commutative KA and regular tree languages for the hypothesis $a \leq w$, where a is a letter and w a multiset or a tree respectively, for which we can use the developed framework and theory. In addition, we want to instantiate our framework to Kleene algebra with additional continuous symbols to prove a completeness result that captures common preliminary results in the completeness proofs of bi-KA, concurrent KA and identity-free Kleene lattices [37, 16]. Subsequently we can then study these algebras in the presence of hypotheses using the developed theory.

Finally, we want to investigate the possibility of varying the signature of algebras when providing reductions, where in this paper we have left it fixed.

References

1 Carolyn Jane Anderson, Nate Foster, Arjun Guha, Jean-Baptiste Jeannin, Dexter Kozen, Cole Schlesinger, and David Walker. NetKAT: semantic foundations for networks. In *Proc. POPL*, pages 113–126. ACM, 2014. doi:10.1145/2535838.2535862.

2 M.N. Bleicher, Hans Schneider, and R.L. Wilson. Permanence of identities on algebras. *Algebra Universalis*, 3:72–93, 1973. doi:10.1007/BF02945106.

3 Stephen L Bloom, Zoltán Ésik, and Gheorghe Stefanescu. Notes on equational theories of relations. *Algebra Universalis*, 33(1):98–126, 1995. doi:10.1007/BF01190768.

4 Stephen L. Bloom and Zoltán Ésik. Free shuffle algebras in language varieties. *Theoretical Computer Science*, 163(1):55–98, 1996. doi:10.1016/0304-3975(95)00230-8.

5 Maurice Boffa. Une remarque sur les systèmes complets d'identités rationnelles. *RAIRO-Theoretical Informatics and Applications*, 24(4):419–423, 1990. doi:10.1051/ITA/1990240404191.

6 Maurice Boffa. Une condition impliquant toutes les identités rationnelles. *Informatique Théorique et Applications*, 29(6):515–518, 1995. doi:10.1051/ITA/1995290605151.

7 Thomas Braibant and Damien Pous. An efficient Coq tactic for deciding Kleene algebras. In *Proc. 1st ITP*, volume 6172 of *LNCS*, pages 163–178. Springer, 2010. doi:10.1007/978-3-642-14052-5_13.

8 Paul Brunet. A note on commutative Kleene algebra. *CoRR*, abs/1910.14381, 2019. arXiv:1910.14381.

9 Paul Brunet. Representations - A meta-model for system analysis. In *Proc. RAMiCS*, LNCS, pages 77–94. Springer, 2026. doi:10.1007/978-3-032-22469-9_5.

10 Paul Brunet and Damien Pous. Kleene algebra with converse. In Peter Höfner, Peter Jipsen, Wolfram Kahl, and Martin Eric Müller, editors, *Proc. RAMiCS*, LNCS, pages 101–118. Springer, 2014. doi:10.1007/978-3-319-06251-8_7.

11 Ernie Cohen. Hypotheses in Kleene algebra. Technical report, Bellcore, Morristown, N.J., 1994. URL: http://www.researchgate.net/publication/2648968_Hypotheses_in_Kleene_Algebra.

12 Hubert Comon, Max Dauchet, Rémi Gilleron, Florent Jacquemard, Denis Lugiez, Christof Löding, Sophie Tison, and Marc Tommasi. *Tree Automata Techniques and Applications*. HAL, 2008. URL: https://inria.hal.science/hal-03367725.

13 John Horton Conway. *Regular algebra and finite machines.* Chapman and Hall, 1971. URL: `https://books.google.nl/books?id=1KAXc5TpEV8C`.

14 Anupam Das, Amina Doumane, and Damien Pous. Left-handed completeness for Kleene algebra, via cyclic proofs. In *Proc. LPAR*, volume 57 of *EPiC Series in Computing*, pages 271–289. Easychair, 2018. `doi:10.29007/hzq3`.

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

16 Amina Doumane and Damien Pous. Completeness for identity-free Kleene lattices. In *Proc. CONCUR*, volume 118 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 18:1–18:17. Schloss Dagstuhl, 2018. `doi:10.4230/LIPIcs.CONCUR.2018.18`.

17 Zoltán Ésik. Axiomatizing the equational theory of regular tree languages. *Journal of Logic and Algebraic Programming*, 79(2):189–213, 2010. `doi:10.1016/j.jlap.2009.10.001`.

18 Zoltán Ésik and László Bernátsky. Equational properties of Kleene algebras of relations with conversion. *Theoretical Computer Science*, 137(2):237–251, 1995. `doi:10.1016/0304-3975(94)00041-G`.

19 N.D. Gautam. The validity of equations of complex algebras. *Archiv für mathematische Logik und Grundlagenforschung*, 3(3):117–124, 1957. `doi:10.1007/BF01988052`.

20 Niels Bjørn Bugge Grathwohl, Dexter Kozen, and Konstantinos Mamouras. KAT + B! In *Proc. CSL-LiCS*. ACM, July 2014. `doi:10.1145/2603088.2603095`.

21 Chris Harden and Dexter Kozen. On the elimination of hypotheses in Kleene algebra with tests. Technical report, Cornell University, 2002. URL: `https://www.cs.cornell.edu/~kozen/Papers/elim.pdf`.

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

23 Tobias Kappé. *Concurrent Kleene Algebra: Completeness and Decidability.* PhD thesis, UCL, London, 2020.

24 Tobias Kappé. Completeness and the finite model property for Kleene algebra, reconsidered. In Roland Glück, Luigi Santocanale, and Michael Winter, editors, *Proc. RAMiCS*, volume 13896 of *LNCS*, pages 158–175. Springer, 2023. `doi:10.1007/978-3-031-28083-2_10`.

25 Tobias Kappé, Paul Brunet, Alexandra Silva, Jana Wagemaker, and Fabio Zanasi. Concurrent Kleene algebra with observations: From hypotheses to completeness. In *Proc. FoSSaCS*, volume 12077 of *LNCS*, pages 381–400. Springer, 2020. `doi:10.1007/978-3-030-45231-5_20`.

26 Tobias Kappé, Paul Brunet, Alexandra Silva, and Fabio Zanasi. Concurrent Kleene algebra: Free model and completeness. In Amal Ahmed, editor, *Proc. ESOP*, LNCS, pages 856–882. Springer, 2018. `doi:10.1007/978-3-319-89884-1_30`.

27 Stephen Cole Kleene. *Representation of events in nerve nets and finite automata*, volume 34. Princeton University Press Princeton, 1956. URL: `https://www.rand.org/content/dam/rand/pubs/research_memoranda/2008/RM704.pdf`.

28 Dexter Kozen. A completeness theorem for Kleene Algebras and the algebra of regular events. In *Proc. LiCS*, pages 214–225. IEEE Computer Society, 1991. `doi:10.1109/LICS.1991.151646`.

29 Dexter Kozen. Kleene algebra with tests. *ACM Trans. Program. Lang. Syst.*, 19(3):427–443, May 1997. `doi:10.1145/256167.256195`.

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

31 Dexter Kozen and Konstantinos Mamouras. Kleene algebra with equations. In Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias, editors, *Proc. Automata, Languages, and Programming*, pages 280–292. Springer, 2014. `doi:10.1007/978-3-662-43951-7_24`.

32 Dexter Kozen and Alexandra Silva. Left-handed completeness. In *Proc. RAMiCS*, volume 7560 of *LNCS*, pages 162–178. Springer, 2012. `doi:10.1007/978-3-642-33314-9_11`.

33 Dexter Kozen and Frederick Smith. Kleene algebra with tests: Completeness and decidability. In Dirk van Dalen and Marc Bezem, editors, *Proc. CSL*, LNCS, pages 244–259. Springer, 1996. `doi:10.1007/3-540-63172-0_43`.

34 Alexander Krauss and Tobias Nipkow. Proof pearl: Regular expression equivalence and relation algebra. *Journal of Automated Reasoning*, 49(1):95–106, 2012. `doi:10.1007/s10817-011-9223-4`.

35 Daniel Krob. Complete systems of B-rational identities. *Theoretical Computer Science*, 89(2):207–343, 1991. `doi:10.1016/0304-3975(91)90395-I`.

36 Stepan L. Kuznetsov. On the complexity of reasoning in Kleene algebra with commutativity conditions. In Erika Ábrahám, Clemens Dubslaff, and Silvia Lizeth Tapia Tarifa, editors, *Proc. ICTAC*, LNCS, pages 83–99. Springer, 2023. `doi:10.1007/978-3-031-47963-2_7`.

37 Michael R. Laurence and Georg Struth. Completeness theorems for bi-Kleene algebras and series-parallel rational pomset languages. In Peter Höfner, Peter Jipsen, Wolfram Kahl, and Martin Eric Müller, editors, *Proc. RAMiCS*, LNCS, pages 65–82. Springer, 2014. `doi:10.1007/978-3-319-06251-8_5`.

38 Donald L Pilling. *The algebra of operators for regular events*. PhD thesis, Cambridge, UK; Cambridge University, 1970.

39 Damien Pous. Kleene algebra with tests and coq tools for while programs. In Sandrine Blazy, Christine Paulin-Mohring, and David Pichardie, editors, *Proc. ITP*, LNCS, pages 180–196. Springer, 2013. `doi:10.1007/978-3-642-39634-2_15`.

40 Damien Pous, Jurriaan Rot, and Jana Wagemaker. On tools for completeness of Kleene algebra with hypotheses. In *Proc. RAMiCS*, volume 13027 of *LNCS*, pages 378–395. Springer, 2021. `doi:10.1007/978-3-030-88701-8_23`.

41 Damien Pous, Jurriaan Rot, and Jana Wagemaker. On tools for completeness of Kleene algebra with hypotheses. *Logical Methods in Computer Science*, 20(2), 2024. `doi:10.46298/LMCS-20(2:8)2024`.

42 Damien Pous and Jana Wagemaker. Completeness Theorems for Kleene Algebra with Top. In *Proc. CONCUR*, volume 243 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 26:1–26:18. Schloss Dagstuhl, 2022. `doi:10.4230/LIPIcs.CONCUR.2022.26`.

43 Damien Pous and Jana Wagemaker. Completeness theorems for Kleene algebra with tests and top. *Logical Methods in Computer Science*, 20(3), 2024. `doi:10.46298/LMCS-20(3:27)2024`.

44 Cristian Prisacariu. Synchronous Kleene algebra. *Journal of Logic and Algebraic Programming*, 79(7):608–635, 2010. `doi:10.1016/J.JLAP.2010.07.009`.

45 A. Shafaat. On varieties closed under the construction of power algebras. *Bulletin of the Australian Mathematical Society*, 11(2):213–218, 1974. `doi:10.1017/S000497270004380X`.

46 Alfred Tarski. A lattice-theoretical fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5(2), 1955.

47 Jana Wagemaker. *Extensions of (Concurrent) Kleene Algebra*. PhD thesis, Radboud University, 2022.

48 Jana Wagemaker, Marcello M. Bonsangue, Tobias Kappé, Jurriaan Rot, and Alexandra Silva. Completeness and incompleteness of synchronous Kleene algebra. In *Proc. MPC*, LNCS, pages 385–413. Springer, 2019. `doi:10.1007/978-3-030-33636-3_14`.

49 Lukas Mulder, Damien Pous, and Jana Wagemaker. Continuous algebras with hypotheses, 2026. Full version of this article, with proofs in appendices. URL: `https://hal.science/hal-05620463`.