

Kleene Algebra with Tests: Completeness and Decidability

Dexter Kozen Frederick Smith
kozen@cs.cornell.edu fms@cs.cornell.edu

Computer Science Department
Cornell University
Ithaca, NY 14853-7501, USA

Abstract. Kleene algebras with tests provide a rigorous framework for equational specification and verification. They have been used successfully in basic safety analysis, source-to-source program transformation, and concurrency control. We prove the completeness of the equational theory of Kleene algebra with tests and *-continuous Kleene algebra with tests over language-theoretic and relational models. We also show decidability. Cohen’s reduction of Kleene algebra with hypotheses of the form $r = 0$ to Kleene algebra without hypotheses is simplified and extended to handle Kleene algebras with tests.

1 Introduction

A *Kleene algebra with tests* is an algebraic structure consisting of a Kleene algebra with an embedded Boolean subalgebra. This formalism provides a rigorous framework for equational specification and verification of programs. It has been applied successfully to problems in basic safety analysis, source-to-source program transformation, and concurrency control [3, 4, 5, 17].

Kleene algebra dates back to a 1956 paper of S. C. Kleene [12] and was developed extensively in a 1971 monograph of Conway [7]. It has appeared in one form or another in relational algebra [20, 25], semantics and logics of programs [13, 23], automata and formal language theory [18], and the design and analysis of algorithms [1, 11]. See [16] for an introduction and a comprehensive list of citations.

Kleene algebra forms an essential component of Propositional Dynamic Logic (PDL) [8], in which it is mixed with modal logic to give a theoretically appealing and practical system for reasoning about computation at the propositional level. Syntactically, PDL is a two-sorted logic consisting of *programs* and *propositions* defined by mutual induction. A basic operator in PDL is the *test operator* $?$, by which a program $\varphi?$ can be formed from any proposition φ . Intuitively, $\varphi?$ acts as a guard that succeeds with no side effects in states satisfying φ and fails or aborts in states not satisfying φ . Tests are used to manipulate flow of control, and are needed to model conventional programming constructs such as conditionals and **while** loops.

From a practical standpoint, many simple program manipulations such as loop unwinding and basic safety analysis do not require the full power of PDL, but can be carried out in a purely equational subsystem using the axioms of Kleene algebra. However, tests are an essential ingredient for modeling real programs. This motivates the definition of *Kleene algebra with tests* (KAT), an equational system introduced in [17]. In that paper, the utility of KAT was illustrated by giving a purely equational proof of the following classical result: every **while** program can be simulated by a **while** program with at most one **while** loop [10, 19].

E. Cohen has taken a slightly different approach in which tests are defined to be elements b satisfying the condition $b \leq 1$. He has given several practical examples of the use of Kleene algebra with conditions in program verification, such as lazy caching and concurrency control [4, 5]. He has shown that Kleene algebra with extra conditions of the form $r = 0$ reduces to Kleene algebra without extra conditions [3], and is therefore decidable. He has also given a direct proof that *-continuous Kleene algebra in the presence of extra commutativity conditions of the form $pq = qp$, even for atomic p and q , is undecidable (see [17]), although with a little extra work this result can be shown to follow from a 1979 result of Berstel [2] (see also [9]).

The proof in [17] only needed extra commutativity conditions of the form $bp = pb$, where b is a test. But as shown in that paper, this equation is equivalent to $bp\bar{b} + \bar{b}pb = 0$. Thus if Cohen's reduction of Kleene algebra with extra conditions $r = 0$ to Kleene algebra without extra conditions could be carried over to Kleene algebra with tests, then one could effectively get rid of the conditions in the proof of [17]. We show that this is indeed the case.

The following are the main results of this paper.

1. A Kleene algebra with tests is called **-continuous* if its Kleene algebra satisfies the *-continuity axiom (7) below. The system KAT with this additional axiom is called KAT*. We show that the equational theories of KAT and KAT* coincide.
2. We show that KAT is complete over relational models. This implies decidability of the equational theory by an essentially trivial reduction to Propositional Dynamic Logic (PDL). In [6], we show by different methods that the problem is *PSPACE*-complete, thus of the same complexity as Kleene algebra.
3. We show that the equational theory of Kleene algebra with tests admits free language-theoretic models consisting of regular sets of "guarded strings". This result is analogous to the completeness result of [16], which states that the regular sets over a finite alphabet Σ form the free Kleene algebra on generators Σ .
4. As mentioned above, Cohen [3] shows that Kleene algebra with extra conditions $r = 0$ reduces efficiently to Kleene algebra without conditions. We simplify Cohen's construction and generalize it to handle Kleene algebra with tests.

2 Kleene Algebra with Tests

A *Kleene algebra with tests* [17] is a Kleene algebra with an embedded Boolean subalgebra. Formally, it is a two-sorted structure

$$(\mathcal{K}, \mathcal{B}, +, \cdot, *, \bar{}, 0, 1)$$

where $\bar{}$ is a unary operator defined only on \mathcal{B} , such that

- $\mathcal{B} \subseteq \mathcal{K}$,
- $(\mathcal{K}, +, \cdot, *, 0, 1)$ is a Kleene algebra, and
- $(\mathcal{B}, +, \cdot, \bar{}, 0, 1)$ is a Boolean algebra.

The elements of \mathcal{B} are called *tests*. We reserve the letters p, q, r, s for arbitrary elements of \mathcal{K} and a, b, c for tests. In PDL, a test would be written $b?$, but since we are using different symbols for tests we omit the $?$.

As is customary, we omit the \cdot , writing pq instead of $p \cdot q$. The precedence of the operators is $\bar{} > * > \cdot > +$. Thus $p + qr^*$ should be parsed $p + (q(r^*))$.

2.1 Kleene Algebra

There have been many competing axiomatizations of Kleene algebra. The formulation we adopt here (KA) is from [16]. Succinctly put, a *Kleene algebra* is an idempotent semiring under $+, \cdot, 0, 1$ satisfying the additional properties

$$1 + pp^* = p^* \tag{1}$$

$$1 + p^*p = p^* \tag{2}$$

$$q + pr \leq r \rightarrow p^*q \leq r \tag{3}$$

$$q + rp \leq r \rightarrow qp^* \leq r \tag{4}$$

where \leq refers to the natural partial order on \mathcal{K} :

$$p \leq q \stackrel{\text{def}}{\iff} p + q = q .$$

The operation $+$ gives the supremum with respect to the natural order \leq . Instead of (3) and (4), we might take the equivalent axioms

$$pr \leq r \rightarrow p^*r \leq r \tag{5}$$

$$rp \leq r \rightarrow rp^* \leq r . \tag{6}$$

Typical models include the family of regular sets over a finite alphabet, the family of binary relations on a set, and the family of $n \times n$ matrices over another Kleene algebra.

A Kleene algebra is said to be **-continuous* if it satisfies the infinitary condition

$$pq^*r = \sup_{n \geq 0} pq^n r \tag{7}$$

where

$$q^0 \stackrel{\text{def}}{=} 1 \quad q^{n+1} \stackrel{\text{def}}{=} qq^n$$

and the supremum is with respect to the natural order \leq .

In the presence of the other axioms, the $*$ -continuity condition (7) implies (3–6), and is strictly stronger in the sense that there exist Kleene algebras that are not $*$ -continuous [14].

The main result of [16] says that all true identities between regular expressions, interpreted as regular sets of strings, are derivable from the axioms of Kleene algebra [16], and only such identities are derivable. In other words, the algebra of regular sets of strings over the finite alphabet Σ is the free Kleene algebra on generators Σ . It is also the free $*$ -continuous Kleene algebra on generators Σ ; i.e., the equational theory of the Kleene algebras and the $*$ -continuous Kleene algebras coincide.

Two useful identities of Kleene algebra are

$$p^*(qp^*)^* = (p+q)^* \tag{8}$$

$$p(qp)^* = (pq)^*p. \tag{9}$$

All the operators are monotone with respect to \leq . In other words, if $p \leq q$, then $pr \leq qr$, $p+r \leq q+r$, and $p^* \leq q^*$ for any r .

See [16] for a more thorough introduction.

2.2 The Boolean Subalgebra

The Boolean subalgebra \mathcal{B} admits a Boolean negation operator $\bar{}$ defined only on \mathcal{B} . Join and meet are given by the Kleene algebra operators $+$ and \cdot , respectively. \mathcal{B} satisfies the axioms of Boolean algebra in addition to the Kleene algebra axioms given above.

2.3 The Language of Kleene Algebra with Tests

Let Σ and \mathcal{B} be disjoint finite sets of symbols. Elements of Σ are called *primitive actions* and elements of \mathcal{B} are called *primitive tests*. *Terms* and *Boolean terms* are defined inductively:

- any primitive action p is a term
- any primitive test b is a Boolean term
- 0 and 1 are Boolean terms
- if p and q are terms, then so are $p+q$, pq , and p^* (suitably parenthesized if necessary)
- if b and c are Boolean terms, then so are $b+c$, bc , and \bar{b} (suitably parenthesized if necessary)
- any Boolean term is a term.

The set of all terms over Σ and \mathbf{B} is denoted $T_{\Sigma, \mathbf{B}}$. The set of all Boolean terms over \mathbf{B} is denoted $T_{\mathbf{B}}$.

An *interpretation* over a Kleene algebra with tests \mathcal{K} is any homomorphism (function commuting with the distinguished operations and constants) defined on $T_{\Sigma, \mathbf{B}}$ and taking values in \mathcal{K} such that the Boolean terms are mapped to elements of the distinguished Boolean subalgebra.

If \mathcal{K} is a Kleene algebra with tests and I is an interpretation over \mathcal{K} , we write $\mathcal{K}, I \models \varphi$ if the formula φ holds in \mathcal{K} under the interpretation I according to the usual semantics of first-order logic. We write $\text{KAT} \models \varphi$ (respectively, $\text{KAT}^* \models \varphi$) if the formula φ is a logical consequence of the axioms of KAT (respectively, KAT^*). In this paper the only formulas we consider are equations or equational implications (universal Horn formulas).

3 A Language-Theoretic Model

Let Σ and \mathbf{B} be disjoint finite sets of symbols. Our language-theoretic model of Kleene algebras with tests is based on the idea of *guarded strings* over Σ and \mathbf{B} . We obtain a guarded string from a string $x \in \Sigma^*$ by inserting *atoms* interstitially among the symbols of x . An *atom* is a Boolean expression representing an atom (minimal nonzero element) of the free Boolean algebra on generators \mathbf{B} .

Formally, an *atom* of $\mathbf{B} = \{b_1, \dots, b_k\}$ is a string of literals $c_1 c_2 \dots c_k$, where each $c_i \in \{b_i, \bar{b}_i\}$. This assumes an arbitrary but fixed order $b_1 < b_2 < \dots < b_k$ on \mathbf{B} ; for technical reasons, we require the literals in an atom to occur in this order. There are exactly 2^k atoms. We denote atoms of \mathbf{B} by $\alpha, \beta, \alpha_0, \dots$. The set of all atoms of \mathbf{B} is denoted $1_{\mathcal{G}}$ (this notation is chosen because $1_{\mathcal{G}}$ will turn out to be the multiplicative identity of our language-theoretic model \mathcal{G}).

If $b \in \mathbf{B}$ and α is an atom of \mathbf{B} , we write $\alpha \leq b$ if b occurs positively in α and $\alpha \leq \bar{b}$ if b occurs negatively in α . This notation is consistent with the natural order in the free Boolean algebra generated by \mathbf{B} .

Intuitively, the symbols of Σ can be thought of as instructions and atoms as conditions that must be satisfied at some point in the computation. If $\alpha \leq c_i$, then α asserts that c_i holds (and \bar{c}_i fails) at that point in the computation.

Definition 1. A *guarded string* over Σ and \mathbf{B} is any element of $(1_{\mathcal{G}}\Sigma)^*1_{\mathcal{G}}$, i.e., any string

$$\alpha_0 p_1 \alpha_1 p_2 \dots p_n \alpha_n, \quad n \geq 0,$$

where each α_i is an atom of \mathbf{B} and each $p_i \in \Sigma$. Note that a guarded string begins and ends with an atom. In the case $n = 0$, a guarded string is just a single atom.

The set of all guarded strings over Σ and \mathbf{B} is denoted $\text{GS}_{\Sigma, \mathbf{B}}$, or just GS when Σ and \mathbf{B} are understood.

Let $\bar{\mathbf{B}} = \{\bar{b} \mid b \in \mathbf{B}\}$. We denote strings in $(\Sigma \cup \mathbf{B} \cup \bar{\mathbf{B}})^*$, including guarded strings, by the letters x, y, z, x_1, \dots .

The analog of concatenation for guarded strings is *coalesced product* (\diamond).

Definition 2. The *coalesced product* operation \diamond is a *partial* binary operation on GS defined as follows:

$$x\alpha \diamond \beta y \stackrel{\text{def}}{=} \begin{cases} x\alpha y, & \text{if } \alpha = \beta \\ \text{undefined}, & \text{otherwise.} \end{cases}$$

In other words, if the terminal atom of the first string is the same as the initial atom of the second string, then the two strings can be *coalesced*. This is like concatenation, except that we combine the two intermediate atoms into one.

If $A, B \subseteq \text{GS}$, define

$$A \diamond B \stackrel{\text{def}}{=} \{x \diamond y \mid x \in A, y \in B\}.$$

Thus $A \diamond B$ consists of all existing coalesced products of guarded strings in A with guarded strings in B .

Whereas the operation \diamond is partial when applied to guarded strings, it is total when applied to *sets* of guarded strings. Note that if there are no existing coalesced products of strings from A and B , then $A \diamond B = \emptyset$. It is not difficult to show that \diamond is associative, that it distributes over union, and that it has two-sided identity $1_{\mathcal{G}}$.

We now define a language-theoretic model $\mathcal{G} = \mathcal{G}_{\Sigma, \mathbf{B}}$ based on guarded strings. The elements of \mathcal{G} will be the regular sets of guarded strings over Σ and \mathbf{B} (although we have not yet defined *regular* in this context). We will also give a standard interpretation of terms in $T_{\Sigma, \mathbf{B}}$ over \mathcal{G} analogous to the standard interpretation of regular expressions as regular sets.

For $A \subseteq \text{GS}$, define inductively

$$A^0 \stackrel{\text{def}}{=} 1_{\mathcal{G}} \quad A^{n+1} \stackrel{\text{def}}{=} A \diamond A^n.$$

The asterate operation for sets of guarded strings is defined by

$$A^* \stackrel{\text{def}}{=} \bigcup_{n \geq 0} A^n.$$

Let $\bar{}$ denote set complementation in $1_{\mathcal{G}}$. That is, if $A \subseteq 1_{\mathcal{G}}$, then $\bar{A} = 1_{\mathcal{G}} - A$. Consider the structure

$$\mathcal{P}_{\Sigma, \mathbf{B}} = (2^{\text{GS}}, 2^{1_{\mathcal{G}}}, \cup, \diamond, *, \bar{}, \emptyset, 1_{\mathcal{G}}).$$

We write \mathcal{P} for $\mathcal{P}_{\Sigma, \mathbf{B}}$ when Σ and \mathbf{B} are understood. It is quite straightforward to verify that \mathcal{P} is a $*$ -continuous Kleene algebra with tests, i.e. is a model of KAT^* . The Boolean algebra axioms hold for $2^{1_{\mathcal{G}}}$ because it is a set-theoretic Boolean algebra.

The $*$ -continuity condition follows immediately from the definition of $*$ and the distributivity of coalesced product over infinite union. We have that

$$A \diamond B^* \diamond C = A \diamond \left(\bigcup_{n \geq 0} B^n \right) \diamond C = \bigcup_{n \geq 0} A \diamond B^n \diamond C.$$

Both of these expressions denote the set

$$\{x \diamond y \diamond z \mid x \in A, z \in C, \exists n y \in B^n\}.$$

For $p \in \Sigma$ and $b \in \mathbf{B}$, define

$$\begin{aligned} G(p) &\stackrel{\text{def}}{=} \{\alpha p \beta \mid \alpha, \beta \in 1_{\mathcal{G}}\} \\ G(b) &\stackrel{\text{def}}{=} \{\alpha \in 1_{\mathcal{G}} \mid \alpha \leq b\}. \end{aligned} \tag{10}$$

The structure $\mathcal{G} = \mathcal{G}_{\Sigma, \mathbf{B}}$ is defined to be the subalgebra of \mathcal{P} generated by the elements $G(p)$ for $p \in \Sigma$ and $G(b)$ for $b \in \mathbf{B}$. Elements of \mathcal{G} are called *regular sets*.

3.1 Standard Interpretation

The map G defined on primitive actions and primitive tests in (10) extends uniquely by induction to a homomorphism $G : T_{\Sigma, \mathbf{B}} \rightarrow \mathcal{G}$:

$$\begin{aligned} G(p + q) &= G(p) \cup G(q) & G(pq) &= G(p) \diamond G(q) \\ G(1) &= 1_{\mathcal{G}} & G(\bar{b}) &= 1_{\mathcal{G}} - G(b) \\ G(0) &= \emptyset & G(p^*) &= G(p)^*. \end{aligned}$$

The map G is called the *standard interpretation* over \mathcal{G} .

4 Relational Models

Relational Kleene algebras with tests are interesting because they closely model our intuition about programs. In a relational model, the elements of \mathcal{K} are binary relations and \cdot is interpreted as relational composition. Elements of the Boolean subalgebra are subsets of the identity relation.

Formally, a *relational Kleene algebra with tests* on a set X is any structure

$$(\mathcal{K}, \mathbf{B}, \cup, \circ, *, \bar{}, \emptyset, \iota)$$

such that

$$(\mathcal{K}, \cup, \circ, *, \emptyset, \iota)$$

is a relational Kleene algebra, i.e. \mathcal{K} is a family of binary relations on X , \circ is ordinary relational composition, $*$ is reflexive transitive closure, and ι is the identity relation on X ; and

$$(\mathbf{B}, \cup, \circ, \bar{}, \emptyset, \iota)$$

is a Boolean algebra of subsets of ι (not necessarily the whole powerset).

All relational Kleene algebras with tests are $*$ -continuous. We write $\text{REL} \models \varphi$ if the formula φ holds in all relational Kleene algebras in the usual sense of first-order logic.

5 Completeness of KAT* under the Standard Interpretation

In this section we prove that an equation $p = q$ is a theorem of *-continuous Kleene algebra with tests iff it holds under the standard interpretation over $\mathcal{G}_{\Sigma, \mathbf{B}}$, where Σ and \mathbf{B} contain all primitive action and test symbols, respectively, appearing in p and q . We will later strengthen this result in §7 by removing the assumption of *-continuity.

Theorem 3. *Let $p, q \in T_{\Sigma, \mathbf{B}}$. Then*

$$\text{KAT}^* \models p = q \iff G(p) = G(q) .$$

*Equivalently, $\mathcal{G}_{\Sigma, \mathbf{B}}$ is the free *-continuous Kleene algebra with tests on generators Σ and \mathbf{B} .*

The forward implication is easy, since \mathcal{G} is a *-continuous Kleene algebra. The converse is a consequence of the following lemma.

Lemma 4. *For any *-continuous Kleene algebra with tests \mathcal{K} , interpretation $I : T_{\Sigma, \mathbf{B}} \rightarrow \mathcal{K}$, and $p, q, r \in T_{\Sigma, \mathbf{B}}$,*

$$I(pqr) = \sup_{x \in G(q)} I(pxr)$$

where the supremum is with respect to the natural order in \mathcal{K} . In particular,

$$I(q) = \sup_{x \in G(q)} I(x) .$$

This result is analogous to the same result for Kleene algebras [15, Lemma 7.1, p. 35] and the proof is similar. Note that the *-continuity axiom is a special case.

Proof of Lemma 4. We proceed by induction on the structure of q . The basis consists of cases for primitive tests, primitive actions, 0 and 1. We argue the case for primitive actions and primitive tests explicitly.

For a primitive action $q \in \Sigma$, recall that

$$G(q) = \{\alpha q \beta \mid \alpha, \beta \in 1_{\mathcal{G}}\} .$$

Then

$$\begin{aligned} I(pqr) &= I(p)I(1)I(q)I(1)I(r) \\ &= \sup\{I(p)I(\alpha)I(q)I(\beta)I(r) \mid \alpha, \beta \in 1_{\mathcal{G}}\} \\ &= \sup\{I(p\alpha q \beta r) \mid \alpha, \beta \in 1_{\mathcal{G}}\} \\ &= \sup\{I(pxr) \mid x \in G(q)\} . \end{aligned}$$

Finite distributivity was used in the second step.

For a primitive test $b \in B$, recall that

$$G(b) = \{\alpha \mid \alpha \leq b\}.$$

Then

$$\begin{aligned} I(pbr) &= I(p)I(b)I(r) \\ &= \sup\{I(p)I(\alpha)I(r) \mid \alpha \leq b\} \\ &= \sup\{I(p\alpha r) \mid \alpha \leq b\} \\ &= \sup\{I(px r) \mid x \in G(b)\}. \end{aligned}$$

Again, finite distributivity was used in the second step.

The induction step consists of cases for $+$, \cdot , $*$, and $\bar{}$. The cases other than \cdot and $\bar{}$ are the same as in [15, Lemma 7.1, p. 35].

For the case \cdot , recall that

$$G(qq') = G(q) \diamond G(q') = \{y\alpha z \mid y\alpha \in G(q), \alpha z \in G(q')\}.$$

Applying the induction hypothesis twice,

$$\begin{aligned} I(pqq'r) &= \sup\{I(pqvr) \mid v \in G(q')\} \\ &= \sup\{\sup\{I(puvr) \mid u \in G(q)\} \mid v \in G(q')\} \\ &= \sup\{I(puvr) \mid u \in G(q), v \in G(q')\}. \end{aligned}$$

The last step follows from a purely lattice-theoretic argument: if all the suprema in question on the left hand side exist, then the supremum on the right hand side exists and the two sides are equal.

Now

$$\begin{aligned} &\sup\{I(puvr) \mid u \in G(q), v \in G(q')\} \\ &= \sup\{I(py\alpha\beta zr) \mid y\alpha \in G(q), \beta z \in G(q')\} \\ &= \sup\{I(py\alpha\alpha zr) \mid y\alpha \in G(q), \alpha z \in G(q')\} \tag{11} \\ &= \sup\{I(py\alpha zr) \mid y\alpha \in G(q), \alpha z \in G(q')\} \\ &= \sup\{I(px r) \mid x \in G(qq')\}. \end{aligned}$$

The justification for step (11) is that if $\alpha \neq \beta$, then the product in \mathcal{K} is 0 and does not contribute to the supremum.

For the case $\bar{}$, recall that

$$G(\bar{b}) = 1_{\mathcal{G}} - G(b) = \{\alpha \mid \alpha \not\leq b\} = \{\alpha \mid \alpha \leq \bar{b}\}.$$

Then

$$I(p\bar{b}r) = \sup\{I(p\alpha r) \mid \alpha \leq \bar{b}\} = \sup\{I(p\alpha r) \mid \alpha \in G(\bar{b})\}.$$

Proof of Theorem 3. If $\text{KAT}^* \models p = q$ then $G(p) = G(q)$, since \mathcal{G} is a $*$ -continuous Kleene algebra with tests. Conversely, if $G(p) = G(q)$, then by Lemma 4, for any $*$ -continuous Kleene algebra with tests \mathcal{K} and any interpretation I over \mathcal{K} , $I(p) = I(q)$. Therefore $\text{KAT}^* \models p = q$.

6 Completeness over Relational Models

In this section we establish completeness over relational models. It will suffice to construct a relational model isomorphic to \mathcal{G} . This construction is similar to a construction of Pratt [22] for regular sets.

For A any set of guarded strings, define

$$h(A) \stackrel{\text{def}}{=} \{(x, x \diamond y) \mid x \in \text{GS}, y \in A\} .$$

Lemma 5. *The language-theoretic model \mathcal{P} and its submodel \mathcal{G} are isomorphic to relational models.*

Proof. We show that the function $h : \mathcal{P} \rightarrow 2^{\text{GS} \times \text{GS}}$ defined above embeds \mathcal{P} isomorphically onto a subalgebra of the Kleene algebra of all binary relations on GS.

It is straightforward to verify that h is a homomorphism. We present the case for \diamond as an example.

$$\begin{aligned} h(A \diamond B) &= \{(z, z \diamond p \diamond q) \mid z \in \text{GS}, p \in A, q \in B\} \\ &= \{(z, z \diamond p) \mid z \in \text{GS}, p \in A\} \\ &\quad \circ \{(z \diamond p, z \diamond p \diamond q) \mid z \in \text{GS}, p \in A, q \in B\} \\ &= \{(z, z \diamond p) \mid z \in \text{GS}, p \in A\} \circ \{(y, y \diamond q) \mid y \in \text{GS}, q \in B\} \\ &= h(A) \circ h(B) . \end{aligned}$$

The function h is injective, since A is uniquely recoverable from $h(A)$:

$$A = \{y \mid \exists \alpha (\alpha, y) \in h(A)\} .$$

The submodel \mathcal{G} is perforce isomorphic to a relational model on GS, namely the image of \mathcal{G} under h .

The following theorem establishes the completeness of KAT^* over relational models.

Theorem 6. *Let REL denote the class of all relational Kleene algebras with tests. Let $p, q \in T_{\Sigma, \mathbf{B}}$. The following are equivalent:*

- (i) $\text{KAT}^* \models p = q$
- (ii) $G(p) = G(q)$
- (iii) $\text{REL} \models p = q$.

Proof. The equivalence of (i) and (ii) was proved in Theorem 3. Since all relational models are *-continuous Kleene algebras with tests, (i) implies (iii). Finally, (iii) implies (ii) by Lemma 5.

7 Completeness of KAT

In this section we show that the equational theories of the Kleene algebras with tests and the *-continuous Kleene algebras with tests coincide by showing that every term p can be transformed into a KAT-equivalent term \widehat{p} such that $G(\widehat{p})$, the set of guarded strings represented by \widehat{p} , is the same as $R(\widehat{p})$, the set of strings represented by \widehat{p} under the ordinary interpretation of regular expressions. The Boolean algebra axioms are not needed in equivalence proofs involving such terms, so we can apply the completeness result of [16] directly.

Consider the set $\overline{\mathbf{B}} = \{\overline{b} \mid b \in \mathbf{B}\}$, the set of negated atomic tests. We can view $\overline{\mathbf{B}}$ as a separate set of primitive symbols disjoint from \mathbf{B} and Σ . Using the DeMorgan laws and the law $\overline{\overline{b}} = b$ of Boolean algebra, every term p can be transformed to a KAT-equivalent term p' in which $\overline{}$ is applied only to primitive test symbols, thus we can view p' as a regular expression over the alphabet $\Sigma \cup \mathbf{B} \cup \overline{\mathbf{B}}$. As such, it represents a set of strings

$$R(p') \subseteq (\Sigma \cup \mathbf{B} \cup \overline{\mathbf{B}})^*$$

under the standard interpretation R of regular expressions as regular sets.

In general, the sets $R(p')$ and $G(p')$ may differ. For example, $R(q) = \{q\}$ for primitive action q , but $G(q) = \{\alpha q \beta \mid \alpha, \beta \in 1_{\mathcal{G}}\}$.

Our main task will be to show how to further transform p' to another KAT-equivalent string \widehat{p} such that all elements of $R(\widehat{p})$ are guarded strings and $R(\widehat{p}) = G(\widehat{p})$. We can then use the completeness result of [16], since p and q will be KAT-equivalent iff \widehat{p} and \widehat{q} are equivalent as regular expressions over $\Sigma \cup \mathbf{B} \cup \overline{\mathbf{B}}$, i.e., if they can be proved equivalent in pure Kleene algebra.

In our inductive proof, it will be helpful to maintain terms in the following special form. Call a term *externally guarded* if it is of the form α or $\alpha q \beta$, where α and β are atoms of \mathbf{B} . Define the *coalesced product* of two such terms as follows:

$$r\alpha \diamond \beta s \stackrel{\text{def}}{=} \begin{cases} r\alpha s, & \text{if } \alpha = \beta \\ 0, & \text{if } \alpha \neq \beta. \end{cases}$$

(Here we must distinguish between a guarded string as a guarded string and a guarded string as a term, since coalesced product is undefined for incompatible pairs of guarded strings.)

For any two externally guarded terms q and r ,

$$G(q \diamond r) = G(q) \diamond G(r),$$

and $q \diamond r$ is externally guarded.

If $\sum_i q_i$ and $\sum_j r_j$ are sums of zero or more externally guarded terms, define

$$\left(\sum_i q_i\right) \diamond \left(\sum_j r_j\right) \stackrel{\text{def}}{=} \sum_{i,j} q_i \diamond r_j.$$

For any two sums q and r of externally guarded terms,

$$G(q \diamond r) = G(q) \diamond G(r),$$

and $q \diamond r$ is a sum of externally guarded terms.

Lemma 7. *For every term p , there is a term \widehat{p} such that*

- (i) $\text{KAT} \models p = \widehat{p}$
- (ii) $R(\widehat{p}) = G(\widehat{p})$
- (iii) \widehat{p} is a sum of zero or more externally guarded terms.

Proof. As argued above, we can assume without loss of generality that all occurrences of $\bar{}$ in p are applied to primitive tests only, thus we may view p as a term over the alphabet $\Sigma \cup \mathbf{B} \cup \overline{\mathbf{B}}$.

We define \widehat{p} by induction on the structure of p . For the basis, take

$$\begin{aligned} \widehat{p} &\stackrel{\text{def}}{=} \sum_{\alpha, \beta \in 1_{\mathcal{G}}} \alpha p \beta, & p \in \Sigma & & \widehat{1} &\stackrel{\text{def}}{=} \sum_{\alpha \in 1_{\mathcal{G}}} \alpha \\ \widehat{b} &\stackrel{\text{def}}{=} \sum_{\alpha \leq b} \alpha, & b \in \mathbf{B} \cup \overline{\mathbf{B}} & & \widehat{0} &\stackrel{\text{def}}{=} 0. \end{aligned}$$

In each of these cases, it is straightforward to verify (i), (ii), and (iii).

For the induction step, suppose we have terms p and q satisfying (ii) and (iii). We take

$$\widehat{p+q} \stackrel{\text{def}}{=} p+q \quad \widehat{p \diamond q} \stackrel{\text{def}}{=} p \diamond q.$$

These constructions are easily shown to satisfy (i), (ii), and (iii).

It remains to construct p^* . We proceed by induction on the number of externally guarded terms in the sum p .

For the basis, we define

$$\begin{aligned} \widehat{0}^* &\stackrel{\text{def}}{=} \widehat{1} \\ \widehat{\alpha}^* &\stackrel{\text{def}}{=} \widehat{1} \\ (\widehat{\alpha q \beta})^* &\stackrel{\text{def}}{=} \widehat{1} + \alpha q \beta, & \alpha \neq \beta & \quad (12) \end{aligned}$$

$$(\widehat{\alpha q \alpha})^* \stackrel{\text{def}}{=} \widehat{1} + \alpha q (\alpha q)^* \alpha. \quad (13)$$

For the induction step, let $p = q + r$, where r is an externally guarded term and q is a sum of externally guarded terms, one fewer in number than in p . By the induction hypothesis, we can construct $q' = \widehat{q}^*$ with the desired properties. Suppose the initial atom of the externally guarded term r is α . Then $\text{KAT} \models r = \alpha r$. Moreover, the expression $(r q' \alpha)^*$ is KAT-equivalent to $(r \diamond q' \diamond \alpha)^*$, which by distributivity can be put into a form in which (12) or (13) applies, yielding a term q'' satisfying (ii) and (iii).

Reasoning in KAT,

$$\begin{aligned} p^* &= (q + r)^* \\ &= q^*(r q^*)^* && \text{by (8)} \\ &= q'(r q')^* \\ &= q' + q' r q' (r q')^* && \text{by (1) and distributivity} \\ &= q' + q' r q' (\alpha r q')^* \\ &= q' + q' (r q' \alpha)^* r q' && \text{by (9)} \\ &= q' + q' q'' r q' \\ &= q' + q' \diamond q'' \diamond r \diamond q', \end{aligned}$$

which is of the desired form.

Theorem 8.

$$\text{KAT} \models p = q \iff G(p) = G(q) .$$

*In other words, the equational theories of the Kleene algebras with tests and the *-continuous Kleene algebras with tests coincide.*

Proof. The forward implication is immediate, since \mathcal{G} is a Kleene algebra with tests.

For the reverse implication, suppose $G(p) = G(q)$. By Lemma 7(i) and Theorem 3, $G(\widehat{p}) = G(\widehat{q})$. By Lemma 7(ii), $R(\widehat{p}) = R(\widehat{q})$. By the completeness result of [16], $\text{KA} \models \widehat{p} = \widehat{q}$. Combining this with Lemma 7(i), we have $\text{KAT} \models p = q$.

Since we have shown that the equational theories of the Kleene algebras with tests and the *-continuous Kleene algebras with tests coincide, we can henceforth write $\models p = q$ unambiguously in place of $\text{KAT}^* \models p = q$ or $\text{KAT} \models p = q$.

8 Eliminating Hypotheses $r = \mathbf{0}$

Cohen [3] shows that in Kleene algebra, any equational implication of the form $r = \mathbf{0} \rightarrow p = q$ reduces efficiently to a single equation. In this section we simplify Cohen's proof and extend it to handle Kleene algebras with tests.

Let $p, q, r \in T_{\Sigma, \mathbf{B}}$. Let u be the *universal expression* $(p_1 + \dots + p_m)^*$, where $\Sigma = \{p_1, \dots, p_m\}$. Under the standard interpretation over the language-theoretic model \mathcal{G} , the term u represents the set of all guarded strings.

The main property of the universal expression is that for any $x \in T_{\Sigma, \mathbf{B}}$, $\models x \leq u$. This can be shown easily in two steps: first, $\models x \leq x'$, where x' is obtained from x by deleting all Boolean symbols; this holds because $\models b \leq 1$ for all Boolean expressions b . Then, $\models x' \leq u$ by ordinary Kleene algebra.

Theorem 9. *The following are equivalent:*

- (i) $\text{KAT} \models r = \mathbf{0} \rightarrow p = q$
- (ii) $\text{KAT}^* \models r = \mathbf{0} \rightarrow p = q$
- (iii) $\models p + uru = q + uru$.

Note that the equivalence of (i) and (ii) does not follow immediately from Theorem 8, since they are not equations but equational implications.

Proof. We first define a congruence on the set $T_{\Sigma, \mathbf{B}}$ of terms in the language of Kleene algebra with tests. For $s, t \in T_{\Sigma, \mathbf{B}}$, define

$$s \equiv t \stackrel{\text{def}}{\iff} \models s + uru = t + uru .$$

The relation \equiv is an equivalence relation. We show that it is a *-continuous Kleene algebra congruence.

If $s = t$ is a theorem of KAT, then $s \equiv t$, since $\vDash s = t$ implies $\vDash s + uru = t + uru$.

To show \equiv is a congruence with respect to $+$, we need to show that $s \equiv t$ implies $s + w \equiv t + w$. But this says only that $\vDash s + uru = t + uru$ implies $\vDash s + w + uru = t + w + uru$, which is immediately apparent.

To show \equiv is a congruence with respect to \cdot , we need to show that $s \equiv t$ implies $sw \equiv tw$ and $ws \equiv wt$. We establish the former; the latter follows by symmetry.

$$\begin{aligned} & \vDash s + uru = t + uru \\ & \Rightarrow \vDash sw + uruw = tw + uruw \\ & \Rightarrow \vDash sw + uruw + uru = tw + uruw + uru \\ & \Rightarrow \vDash sw + uru = tw + uru . \end{aligned}$$

To show \equiv is a congruence with respect to * , we need to show that $s \equiv t$ implies $s^* \equiv t^*$.

$$\begin{aligned} & \vDash s + uru = t + uru \\ & \Rightarrow \vDash (s + uru)^* = (t + uru)^* \\ & \Rightarrow \vDash s^*(urus^*)^* = t^*(urut^*)^* \\ & \Rightarrow \vDash s^*(1 + urus^*(urus^*)^*) = t^*(1 + urut^*(urut^*)^*) \\ & \Rightarrow \vDash s^* + s^*urus^*(urus^*)^* + uru = t^* + t^*urut^*(urut^*)^* + uru \\ & \Rightarrow \vDash s^* + uru = t^* + uru . \end{aligned}$$

To show \equiv is a congruence with respect to $\bar{}$, we need to show that for Boolean terms b, c , if $b \equiv c$ then $\bar{b} \equiv \bar{c}$. This case follows from previous results. If $b \equiv c$, then $b + \bar{c} \equiv c + \bar{c} \equiv 1$, thus $\bar{c}\bar{b} \equiv (b + \bar{c})\bar{b} \equiv \bar{b}$. By symmetry, $\bar{c}\bar{b} \equiv \bar{c}$, therefore $\bar{b} \equiv \bar{c}$.

Finally, to show that \equiv respects * -continuity (7), we need only show that if $st^n v + y \equiv y$ for all n , then $st^* v + y \equiv y$:

$$\begin{aligned} & \vDash (st^n v + y) + uru = y + uru \text{ for all } n \\ & \Rightarrow \vDash st^n v + (y + uru) = y + uru \text{ for all } n \\ & \Rightarrow \vDash st^* v + (y + uru) = y + uru \\ & \Rightarrow \vDash (st^* v + y) + uru = y + uru . \end{aligned} \tag{14}$$

The crucial step (14) follows from the fact that if $st^n v \leq y + uru$ for all n in all * -continuous Kleene algebras, then $st^* v \leq y + uru$ in all * -continuous Kleene algebras.

Since \equiv is a KAT * congruence on $T_{\Sigma, \mathbf{B}}$, we can form the quotient $T_{\Sigma, \mathbf{B}}/\equiv$ and canonical interpretation $s \mapsto [s]$, where $[s]$ denotes the \equiv -congruence class of s , and this structure is a * -continuous Kleene algebra with tests. The equation $r = 0$ is satisfied under this interpretation, since

$$\vDash r + uru = uru = 0 + uru ,$$

so $r \equiv 0$.

Now we are ready to prove the equivalence of the three conditions in the statement of the theorem.

(i) \Rightarrow (ii) Any formula true in all Kleene algebras with tests is certainly true in all $*$ -continuous Kleene algebras with tests.

(ii) \Rightarrow (iii) If $\text{KAT}^* \models r = 0 \rightarrow p = q$, then since $T_{\Sigma, \mathbf{B}}/\equiv$ is a $*$ -continuous Kleene algebra with tests and $T_{\Sigma, \mathbf{B}}/\equiv, [\] \models r = 0$, we have $T_{\Sigma, \mathbf{B}}/\equiv, [\] \models p = q$. By definition, $p \equiv q$, which is what we wanted to show.

(iii) \Rightarrow (i) Suppose $\models p + uru = q + uru$. Let \mathcal{K} be an arbitrary Kleene algebra with tests and let I be an arbitrary interpretation over \mathcal{K} such that $\mathcal{K}, I \models r = 0$. Then $\mathcal{K}, I \models p = p + uru = q + uru = q$. Since \mathcal{K} and I were arbitrary, $\text{KAT} \models r = 0 \rightarrow p = q$.

9 Decidability

Once we have Theorem 6, the decidability of the equational theory of Kleene algebra with tests follows almost immediately from a simple reduction to Propositional Dynamic Logic (PDL). Any term in the language of KAT is a program of PDL (after replacing Boolean terms b with PDL tests $b?$), and it is known that two such terms p and q represent the same binary relation in all relational structures iff

$$\text{PDL} \models \langle p \rangle c \leftrightarrow \langle q \rangle c ,$$

where c is a new primitive proposition symbol [8]. By Theorems 6 and 8, this is tantamount to deciding KAT-equivalence.

PDL is known to be exponential time complete [8, 21], thus the equational theory of KAT is decidable in no more than exponential time. It is at least *PSPACE*-hard, since the equational theory of Kleene algebras is [24].

It can be shown by different methods that the equational theory of KAT is *PSPACE*-complete [6].

Acknowledgements

Ernie Cohen provided valuable comments. The support of the National Science Foundation under grant CCR-9317320 is gratefully acknowledged. The second author is supported on a National Science Foundation Graduate Fellowship.

References

1. A. V. Aho, J. E. Hopcroft, and J. D. Ullman. *The Design and Analysis of Computer Algorithms*. Addison-Wesley, 1975.
2. J. Berstel. *Transductions and Context-free Languages*. Teubner, 1979.
3. E. Cohen. Hypotheses in Kleene algebra.
<ftp://ftp.bellcore.com/pub/ernie/research/homepage.html>, April 1994.

4. E. Cohen. Lazy caching.
ftp://ftp.bellcore.com/pub/ernie/research/homepage.html, 1994.
5. E. Cohen. Using Kleene algebra to reason about concurrency control.
ftp://ftp.bellcore.com/pub/ernie/research/homepage.html, 1994.
6. E. Cohen, D. Kozen, and F. Smith. The complexity of Kleene algebra with tests.
Tech. Rep. TR96-1598, Cornell University, July 1996.
7. J. H. Conway. *Regular Algebra and Finite Machines*. Chapman and Hall, 1971.
8. M. J. Fischer and R. E. Ladner. Propositional dynamic logic of regular programs.
J. Comput. Syst. Sci., 18(2):194–211, 1979.
9. A. Gibbons and W. Rytter. On the decidability of some problems about rational subsets of free partially commutative monoids. *Theor. Comput. Sci.*, 48:329–337, 1986.
10. D. Harel. On folk theorems. *Comm. Assoc. Comput. Mach.*, 23(7):379–389, July 1980.
11. K. Iwano and K. Steiglitz. A semiring on convex polygons and zero-sum cycle problems. *SIAM J. Comput.*, 19(5):883–901, 1990.
12. S. C. Kleene. Representation of events in nerve nets and finite automata. In Shannon and McCarthy, editors, *Automata Studies*, pages 3–41. Princeton University Press, 1956.
13. D. Kozen. On induction vs. *-continuity. In Kozen, editor, *Proc. Workshop on Logic of Programs*, volume 131 of *Lect. Notes in Comput. Sci.*, pages 167–176. Springer, 1981.
14. D. Kozen. On Kleene algebras and closed semirings. In Rovan, editor, *Proc. Math. Found. Comput. Sci.*, volume 452 of *Lect. Notes in Comput. Sci.*, pages 26–47. Springer, 1990.
15. D. Kozen. *The Design and Analysis of Algorithms*. Springer-Verlag, 1991.
16. D. Kozen. A completeness theorem for Kleene algebras and the algebra of regular events. *Infor. and Comput.*, 110(2):366–390, May 1994.
17. D. Kozen. Kleene algebra with tests and commutativity conditions. In T. Margaria and B. Steffen, editors, *Proc. Second Int. Workshop Tools and Algorithms for the Construction and Analysis of Systems (TACAS'96)*, volume 1055 of *Lect. Notes in Comput. Sci.*, pages 14–33. Springer, March 1996.
18. W. Kuich and A. Salomaa. *Semirings, Automata, and Languages*. Springer, 1986.
19. G. Mirkowska. *Algorithmic Logic and its Applications*. PhD thesis, University of Warsaw, 1972. In Polish.
20. K. C. Ng. *Relation Algebras with Transitive Closure*. PhD thesis, University of California, Berkeley, 1984.
21. V. R. Pratt. Models of program logics. In *Proc. 20th Symp. Found. Comput. Sci.*, pages 115–122. IEEE, 1979.
22. V. R. Pratt. Dynamic algebras and the nature of induction. In *Proc. 12th Symp. Theory of Comput.*, pages 22–28. ACM, 1980.
23. V. R. Pratt. Dynamic algebras as a well-behaved fragment of relation algebras. In D. Pigozzi, editor, *Proc. Conf. on Algebra and Computer Science*, volume 425 of *Lect. Notes in Comput. Sci.*, pages 77–110. Springer, June 1988.
24. L. J. Stockmeyer and A. R. Meyer. Word problems requiring exponential time. In *Proc. 5th Symp. Theory of Computing*, pages 1–9. ACM, 1973.
25. A. Tarski. On the calculus of relations. *J. Symb. Logic*, 6(3):65–106, 1941.