BOSTON UNIVERSITY

COLLEGE OF ENGINEERING

Dissertation

SOME EXTENSIONS OF KLEENE ALGEBRA AND THEIR APPLICATIONS

by

CHENG ZHANG

B.A., Wheaton College, 2018

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2024

MS theses only: Prior to having this page signed by the readers, please have it reviewed by the Mugar Library staff. This does not apply to PhD dissertations since this page is completed through DocuSign. **Remove this comment in the final document** by commenting out `\approvalpagewithcomment` and uncommenting `\approvalpage` statements in the *prelim.tex* file.

Approved by

First Reader

First M. Last, PhD
Professor of Electrical and Computer Engineering

Second Reader

First M. Last
Associate Professor of …

Third Reader

First M. Last
Assistant Professor of …

Facilis descensus Averni;
Noctes atque dies patet atri janua Ditis;
Sed revocare gradum, superasque evadere ad auras,
Hoc opus, hic labor est. Virgil (from Don's thesis!)

Acknowledgments

Here go all your acknowledgments. You know, your advisor, funding agency, lab mates, etc., and of course your family.

As for me, I would like to thank Jonathan Polimeni for cleaning up old LaTeX style files and templates so that Engineering students would not have to suffer typesetting dissertations in MS Word. Also, I would like to thank IDS/ISS group (ECE) and CV/CNS lab graduates for their contributions and tweaks to this scheme over the years (after many frustrations when preparing their final document for BU library). In particular, I would like to thank Limor Martin who has helped with the transition to PDF-only dissertation format (no more printing hardcopies – hooray !!!)

The stylistic and aesthetic conventions implemented in this LaTeX thesis/dissertation format would not have been possible without the help from Brendan McDermot of Mugar library and Martha Wellman of CAS.

Finally, credit is due to Stephen Gildea for the MIT style file off which this current version is based, and Paolo Gaudiano for porting the MIT style to one compatible with BU requirements.

Janusz Konrad

Professor

ECE Department

SOME EXTENSIONS OF KLEENE ALGEBRA AND THEIR APPLICATIONS

CHENG ZHANG

Boston University, College of Engineering, 2024

Major Professors: First M. Last, PhD
Professor of Electrical and Computer Engineering
Secondary appointment

First M. Last, PhD
Professor of Computer Science

ABSTRACT

Have you ever wondered why this is called an *abstract*? Weird thing is that its legal to cite the abstract of a dissertation alone, apart from the rest of the manuscript.

Contents

List of Tables

List of Figures

List of Abbreviations

As per BU library instructions, the list of abbreviations must be in alphabetical order by the **abbreviation**, not by the explanation, or it will be returned to you for re-ordering. **This comment must be removed in the final document.**

CAD	.....	Computer-Aided Design
CO	.....	Cytochrome Oxidase
DOG	.....	Difference Of Gaussian (distributions)
FWHM	.....	Full-Width at Half Maximum
LGN	.....	Lateral Geniculate Nucleus
ODC	.....	Ocular Dominance Column
PDF	.....	Probability Distribution Function
$\mathbb{R}^2$	.....	the Real plane

Chapter 1

Introduction

1.1 A Brief History Of Kleene Algebra

Our Contributions

1.2 Technical Background

1.2.1 Constructing Models

Upper-triangular matrix model plays an important role in this paper. Given a KA $\mathcal{K}$, the square upper-triangular matrices over $\mathcal{K}$ of size n, denoted as $M_n(\mathcal{K})$, forms a Kleene Algebra, with matrix addition, matrix multiplication, and a star operation inductively defined as follows (**?**):

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix}^* \triangleq \begin{bmatrix} A^* & A^*BD^* \\ 0 & D^* \end{bmatrix},$$

where A, D are square block matrices, and B is a block matrix.

Corollary 1. *Projections of the diagonal element of upper-triangular matrices are homomorphisms:*

$$\pi_{n,n} : M_m(\mathcal{K}) \to \mathcal{K} \text{ where } n \leq m.$$

Proof. By induction on the input term. □

Corollary 2. *Given an interpretation into a matrix model $I : \mathsf{KA}(X) \to M_m(\mathcal{K})$ and a Kleene subalgebra $\mathcal{K}' \subseteq \mathcal{K}$, any diagonal projection $\pi_{n,n}$ satisfy the following equivalences:*

$$\forall a \in X, \pi_{n,n}(I(a)) \subseteq \mathcal{K}' \iff \forall e \in \mathsf{KA}(X), \pi_{n,n}(I(e)) \subseteq \mathcal{K}'.$$

Proof. Direct consequence of Corollary 1. □

Another important class of models are models bounded by a single element:

Theorem 1. *Given a KA $\mathcal{K}$, and an element $p \in \mathcal{K}$, if $p \geq 1$ and $p \cdot p \leq p$, then all the element in $\mathcal{K}$ that is smaller than p forms a Kleene Algebra.*

Proof. We need to show that $\{q \mid q \leq p\}$ is closed under all operation of KA. We will show the star case as an example. Given an element $q \leq p$, we need to show $q^* \leq p$. By induction rule

$$qp \leq pp \leq p \Longrightarrow q^*p \leq p \Longrightarrow q^* \leq q^*p \leq p.$$ □

We denote the Kleene algebra formed by all the elements in $\mathcal{K}$ that is less than p as $\mathcal{K}_p$.

Chapter 2

Kleene Algebra with Atomic Commutativity

2.1 Free KA with Atomic Commutativity

It is common to extend Kleene Algebra with additional equations to enrich the theory(?; ?; ?). In this paper we will consider atomic commutativity hypotheses, where the equations in the hypotheses are of the form $pq = qp$ with p and q being primitives.

A *commutable set* $(X, \sim)$ is a set with a reflexive symmetric relation $\sim : X \times X$ called *commuting relation*, we typically omit $\sim$ and just denote the commutable set as X. In this paper we only consider finite commutable sets.

We say a commutable set X is *discrete* if the relation $\sim$ is the identity relation. A homomorphism $h : X \to Y$ between two commutable set X and Y is a function that preserves the commuting relation:

$$x_1 \sim x_2 \Longrightarrow h(x_1) \sim h(x_2).$$

The carrier of a commutable set X can be considered as a discrete commutable set, and we denote this discrete commutable set as $X_{\sim}$. There is a canonical homomorphism:

$$[-]_{\sim} : X_{\sim} \to X$$
$$[x]_{\sim} \triangleq x.$$

We can construct the free KA a commutable set X by taking all the KA terms

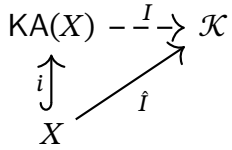
over X modulo the equalities provable from KA axioms plus the following equations $\{pq = qp \mid p \sim q\}$. Intuitively, the commuting relation of X specifies the atomic commutativity hypotheses in $\mathsf{KA}(X)$. Since the free KA over a set is just a free KA over discrete commutable set, we abuse the notation to denote the free KA over a commutable set X as $\mathsf{KA}(X)$.

Notice all Kleene Algebra form a commutable set, with the commuting relation defined as follows:

$$e_1 \sim e_2 \iff e_1 \cdot e_2 = e_2 \cdot e_1.$$

We can show that the free KA over commutable set enjoys similar universal property as free KA over set. We first prove the universal property without the uniqueness requirement:

Theorem 2. *For all commutable set X, a KA $\mathcal{K}$, and a commutable set homomorphism $\hat{I} : X \to \mathcal{K}$, then there is a KA interpretation $I : \mathsf{KA}(X) \to \mathcal{K}$, s.t. the following diagram commutes:*



Proof. Given the function $\hat{I}$, we can apply the standard technique to generate the homomorphism I by induction on the input:

$$\begin{aligned}
I(a) &\triangleq \hat{I}(a) && a \in X \\
I(1) &\triangleq 1_{\mathcal{K}} \\
I(0) &\triangleq 0_{\mathcal{K}} \\
I(e_1 + e_2) &\triangleq I(e_1) + I(e_2) \\
I(e_1 \cdot e_2) &\triangleq I(e_1) \cdot I(e_2) \\
I(e^*) &\triangleq I(e)^*
\end{aligned}$$

such a homomorphism exists, and makes the diagram commute:

$$I(a) = \hat{I}(a), \forall a \in X.$$

□

To prove uniqueness, we will prove a stronger theorem first.

Theorem 3. *Given two interpretation* $I, I' : \mathsf{KA}(X) \to \mathcal{K}$*,*

$$I(e) \geq I'(e) \iff \forall a \in X, I(a) \geq I'(a),$$

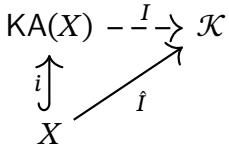
this result implies $I(e) = I'(e) \iff \forall a \in X, I(a) = I'(a)$*.*

Proof. By induction on the structure of e, and all KA operations preserve order. We show the star case as example: assume $I(e) \geq I'(e)$, we need to show $I(e^*) \geq I'(e^*)$. Since I is a homomorphism, and star preserves order:

$$I(e^*) = (I(e))^* \geq (I'(e))^* = I'(e^*).$$

□

Corollary 3 (Universal Property)**.** *For all commutable set* X*, a KA* $\mathcal{K}$*, and a commutable set homomorphism* $\hat{I} : X \to \mathcal{K}$*, then there is a unique KA interpretation* $I : \mathsf{KA}(X) \to \mathcal{K}$*, s.t. the following diagram commutes:*



Proof. By Theorem 2, I exists. By Theorem 3, if there exists another interpretation I' that makes the diagram commute, then

$$I(a) = I'(a), \forall a \in X \implies I(e) = I'(e).$$

□

As usual, we will use the notation I for both I and $\hat{I}$.

The words over a commutable set X are monoid terms modulo monoid equations plus the commutativity axioms $\{ab = ba \mid a \sim b\}$. We still use ϵ as the identity of the monoid and call it the empty word; and we use the same notation $\mathsf{Wrd}(X)$ for all the words over X. The language model over a commutable set X is the powerset of all words over X, with operation defined by Kozen (**?**), denoted as $\mathcal{L}_X$. The language interpretation is generated by the same action on primitives as in Kleene Algebra:

$$L : X \to \mathcal{L}_X$$
$$L(a) = \{a\}$$

Notation In the rest of the article, notations $\mathsf{KA}(X)$, $\mathsf{Wrd}(X)$, and $\mathcal{L}_X$ always refers to the commutative variant, where X is a commutable set. When we are referring to the non-commutative KA (word, language model, etc.), we will consider them as the KA (word, language model, etc.) over a discrete commutable set. As we have mentioned before, $2 \triangleq \{0, 1\}$ denotes the unique KA that only contains two distinct identities; this KA is also the free KA generated by the empty set $\mathsf{KA}(\varnothing)$. Finally, when given a finite set of terms $S \subseteq \mathsf{KA}(X)$, we will sometimes use S to denote the sum of all its elements $(\sum_{e \in S} e) \in \mathsf{KA}(e)$.

2.2 Word Inhabitant Problem

Given a commutable set, we allow a word in $\mathcal{L}_X$ to be implicitly coerced into $\mathsf{KA}(X)$, where we pick the multiplication operator in $\mathsf{KA}(X)$ as the monoidal multiplication.

Then given a word $w \in \mathcal{L}_X$ and a KA expression $e \in \mathsf{KA}(X)$, the word inhabitant problem is the following inequality:

$$w \leq e.$$

The problem is complete with language interpretation when:

$$w \in L(e) \iff w \leq e.$$

We will show that the word inhabitance problem is complete and decidable in Kleene Algebra with atomic commutativity hypotheses.

The core technique of this section is to construct a sound empty word predicate $E : \mathsf{KA}(X) \to 2$ and derivative operation $\delta_a : \mathsf{KA}(X) \to \mathsf{KA}(X)$.

2.2.1 Empty Word Predicate

In this section we will prove a stronger result than the soundness of empty word predicate:

$$\forall e \in \mathsf{KA}(X), e = E(e) + e',$$

where $E : \mathsf{KA}(X) \to 2$ is the empty word predicate on the free KA over any commutable set X, and $\epsilon \notin L(e')$. This result is obtained by decomposing using the following matrix model.

Theorem 4. *For any Kleene Algebra $\mathcal{K}$, matrix of the following shapes forms a Kleene Algebra:*

$$D_E(\mathcal{K}) \triangleq \{\begin{bmatrix} p & q \\ 0 & p+q \end{bmatrix} \mid p, q \in \mathcal{K}\}.$$

Proof. We will only need to show that matrix of this shape is closed under all the KA operations.

The identities and addition are easy to verify. So we will only focus on verifying the closure under multiplication and star operation.

The multiplication case:

$$\begin{bmatrix} p_1 & q_1 \\ 0 & p_1+q_1 \end{bmatrix}\begin{bmatrix} p_2 & q_2 \\ 0 & p_2+q_2 \end{bmatrix} = \begin{bmatrix} p_1p_2 & p_1q_2 + q_1(p_2+q_2) \\ 0 & (p_1+q_1)(p_2+q_2) \end{bmatrix}.$$

Since $p_1p_2 + p_1q_2 + q_1(p_2+q_2) = (p_1+q_1)(p_2+q_2)$, these matrices are closed under multiplication.

The star case:

$$\begin{bmatrix} p & q \\ 0 & p+q \end{bmatrix}^* = \begin{bmatrix} p^* & p^*q(p+q)^* \\ 0 & (p+q)^* \end{bmatrix}.$$

With a standard theorem of KA $(p+q)^* = p^*(qp^*)^*$, we are able to derive the closure under star operation:

$$\begin{aligned} p^* + p^*q(p+q)^* &= p^* + p^*qp^*(qp^*)^* \\ &= p^*(1 + qp^*(qp^*)^*) \\ &= p^*(qp^*)^* = (p+q)^* \end{aligned}$$

□

Given any commutable set X, consider the following matrix:

$$D_E(\mathsf{KA}(X)) \ni u_E \triangleq \begin{bmatrix} 1 & XX^* \\ 0 & X^* \end{bmatrix},$$

where X is a shorthand for the expression $(\sum_{x \in X} x)$. By simply unfolding the definition, we can verify that $u_E \cdot u_E = u_E$ and $u_E \geq 1$. Therefore, all the matrices less than u_E in $D_E(\mathsf{KA}(X))$ forms a Kleene Algebra. We denote this Kleene Algebra as $D_E(\mathsf{KA}(X))_{u_E}$.

In order to decompose an arbitrary expression, we will define an interpretation into $D_E(\mathsf{KA}(X))_{u_E}$ by lifting the following actions

$$\begin{aligned} &I_E : \mathsf{KA}(X) \to D_E(\mathsf{KA}(X))_{u_E} \\ &I_E(a) \triangleq \begin{bmatrix} 0 & a \\ 0 & a \end{bmatrix}. \end{aligned}$$

Because the projection $\pi_{2,2}$ is a homomorphism, then $\pi_{2,2} \circ I_E$ is an interpretation. Recall that interpretation is uniquely determined by the action on the primitives, and

$$\pi_{2,2} \circ I_E(a) = a, \forall a \in X.$$

Therefore, for all term in $e \in \mathsf{KA}(X)$, the $2,2$ component of $I_E(e)$ is exactly e itself:

$$\pi_{2,2} \circ I_E(e) = e.$$

Then we define the empty word predicate as follows:

$$E(e) \triangleq \pi_{1,1}(I_E(e)), \quad e' \triangleq \pi_{1,2}(I_E(e)).$$

By Corollary 2, and $\pi_{1,1}(I_E(a)) = 0 \in 2$ for all primitives a,

$$\forall e \in \mathsf{KA}(X), E(e) = \pi_{1,1}(I_E(e)) \in 2 \subseteq \mathsf{KA}(X).$$

Therefore, we can treat E as a homomorphism of the type $\mathsf{KA}(X) \to 2$.

Corollary 4 (empty word decomposition)**.** *All expression $e \in \mathsf{KA}(X)$ over a commutable set X can be decomposed in the following way:*

$$e = E(e) + e' \text{ where } \epsilon \notin L(e').$$

Proof. Recall that

$$\begin{bmatrix} E(e) & e' \\ 0 & e \end{bmatrix} \triangleq I_E(e).$$

Since $I_E(e) \in D_E(\mathsf{KA}(X))$, we have

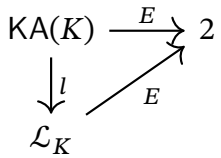
$$e = E(e) + e'.$$

Furthermore, since elements in $D_E(\mathsf{KA}(X))_{u_E}$ is bounded by u_E,

$$e' = \pi_{2,2}(I_E(e)) \leq XX^*.$$

Because $\epsilon \notin L(XX^*)$, and l is a homomorphism, we conclude $\epsilon \notin L(e) \subseteq L(XX^*)$. □

Corollary 5 (Soundness Of Empty Word Property)**.** *Let $E : \mathcal{L}_X \to 2$ the empty word predicate on the language over a commutable set $E(l) = \epsilon \in l$, then the following diagram commute*



Proof. We only need to prove that for all $e \in \mathsf{KA}(X)$,

$$E(e) = 1 \iff \epsilon \in l.$$

We show this by case analysis on $E(e)$:

- If $E(e) = 1$, then

$$L(e) = L(E(e)) \cup L(e') = \{\epsilon\} \cup L(e') \ni \epsilon.$$

- If $E(e) = 0$, recall that $\epsilon \notin L(e')$,

$$L(e) = L(E(e)) \cup L(e') = L(e') \not\ni \epsilon.$$

□

2.2.2 Derivative

Similar to the last section, the derivative operation will also be defined by a decomposition: for all $e \in \mathsf{KA}(X)$ and $a \in X$,

$$e = a \cdot \delta_a(e) + \rho_a(e),$$

where the language interpretation for $a \cdot \delta_a(e)$ and $\rho_a(e)$ are disjoint. This result will imply the soundness of derivative.

Theorem 5. *Given a KA $\mathcal{K}$ and an element $t \in \mathcal{K}$, the following matrices form a KA:*

$$D_t(\mathcal{K}) = \{\begin{bmatrix} a & b & c \\ 0 & d & 0 \\ 0 & 0 & d \end{bmatrix} \mid d = a + b + tc, at = ta\}$$

Proof. We need to show that these matrices are closed under KA operations. The closure under identities and addition are trivial, we only show the multiplication case and the star case.

The multiplication case:

$$\begin{bmatrix} p_1 & q_1 & r_1 \\ 0 & s_1 & 0 \\ 0 & 0 & s_1 \end{bmatrix} \begin{bmatrix} p_2 & q_2 & r_2 \\ 0 & s_2 & 0 \\ 0 & 0 & s_2 \end{bmatrix}$$
$$= \begin{bmatrix} p_1p_2 & p_1q_2 + q_1s_2 & p_1r_2 + r_1s_2 \\ 0 & s_1s_2 & 0 \\ 0 & 0 & s_1s_2 \end{bmatrix}$$

We verify that the equation is preserved:

$$\begin{aligned} s_1 s_2 &= (p_1 + q_1 + pr_1) \cdot s_2 \\ &= p_1 s_2 + q_1 s_2 + pr_1 s_2 \\ &= p_1(p_2 + q_2 + pr_2) + q_1 s_2 + pr_1 s_2 \\ &= p_1 p_2 + (p_1 q_2 + q_1 s_2) + p_1 pr_2 + pr_1 s_2 \\ &= p_1 p_2 + (p_1 q_2 + q_1 s_2) + p(p_1 r_2 + r_1 s_2) \end{aligned}$$

The last step uses the commutativity of t and p_1. Then we verify the commutativity condition:

$$(p_1 p_2)t = p_1 t p_2 = t(p_1 p_2).$$

Hence, $D_t(\mathcal{K})$ is closed under multiplication.

The star case:

$$\begin{bmatrix} p & q & r \\ 0 & s & 0 \\ 0 & 0 & s \end{bmatrix}^* = \begin{bmatrix} p^* & p^* q s^* & p^* r s^* \\ 0 & s^* & 0 \\ 0 & 0 & s^* \end{bmatrix}$$

the equation is preserved:

$$\begin{aligned} s^* &= (p + q + tc)^* \\ &= p^*((q + pc)p^*)^* \\ &= p^*(1 + (q + pc)p^*((q + pc)p^*)^*) \\ &= p^*(1 + (q + pc)s^*) \\ &= p^* + p^* q s^* + p^* t r s^* \\ &= p^* + p^* q s^* + t p^* r s^* \end{aligned}$$

The last line is by standard KA theorem:

$$pt = tp \Longrightarrow p^* t = t p^*.$$

The commutativity condition $p^* t = t p^*$ is also implied by the above theorem. Therefore, $D_t(\mathcal{K})$ is closes under star operations. □

Given a commutable set X, and an element $a \in X$, we can partition the rest of

the elements in X by whether they commute with a:

$$X_{\sim a} \triangleq \{b \mid b \sim a, b \neq a\}, \quad X_{\nsim a} = \{b \mid b \nsim a\}.$$

Since a commutes with every element of $X_{\sim a}$, a commutes with $X_{\sim a}$: $X_{\sim a} \cdot a = a \cdot X_{\sim a}$, then by standard theorem of KA:

$$X_{\sim a}^* \cdot a = a \cdot X_{\sim a}^*.$$

Consider the following matrix:

$$D_a(\mathsf{KA}(X)) \ni u_a = \begin{bmatrix} X_{\sim a}^* & X_{\sim a}^* X_{\nsim a} X^* & X^* \\ 0 & X^* & 0 \\ 0 & 0 & X^* \end{bmatrix}.$$

It is easy to verify that $u_a \geq 1$ and $u_a \cdot u_a \leq u_a$. Therefore, the elements under u_a forms a KA: $D_a(\mathsf{KA}(X))_{u_a}$. The purpose of model $D_a(\mathsf{KA}(X))_{u_a}$ is clear when we look at the language interpretation for each of the component, let

$$\begin{bmatrix} p & q & r \\ 0 & s & 0 \\ 0 & 0 & s \end{bmatrix} \in D_a(\mathsf{KA}(X))_{u_a}$$

then

- $L(p) \leq L(X_{\sim a}^*)$ contains only words with symbols that commutes with primitive a, but is not a.
- $L(q) \leq L(X_{\sim a}^* X_{\nsim a} X^*)$ contains words that starts with arbitrary number of primitives that commutes with a, then a primitive that does not commute with a, followed by arbitrary primitives.

Both $L(p)$ and $L(q)$ do not contain words of the form $a \cdot w$ for any word $w \in \mathsf{Wrd}(X)$;

by the property of $D_a(\mathsf{KA}(X))$:

$$L(s) = L(p) + L(q) + a \cdot L(r).$$

Thus, $L(r)$ will be the language derivative of $L(s)$ with respect to primitive a,

To apply this decomposition on an arbitrary expression, we define an interpretation by lifting the following action on primitives:

$$I_a : X \to D_a(\mathsf{KA}(X))_{u_a}$$

$$I_a(b) \triangleq \begin{cases} \begin{bmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & b \end{bmatrix} & b \in X_{\sim a} \\ \begin{bmatrix} 0 & b & 0 \\ 0 & b & 0 \\ 0 & 0 & b \end{bmatrix} & b \in X_{\nsim a} \\ \begin{bmatrix} 0 & 0 & 1 \\ 0 & b & 0 \\ 0 & 0 & b \end{bmatrix} & b = a \end{cases}$$

Again, since $\pi_{3,3}$ and $\pi_{2,2}$ are homomorphisms, $\pi_{3,3} \circ I_a$ and $\pi_{2,2} \circ I_a$ are interpretations. Because interpretations are uniquely determined by the action on the primitives, and

$$\forall b \in X, \pi_{3,3} \circ I_a(b) = \pi_{2,2} \circ I_a(b) = b,$$

then $\pi_{3,3} \circ I_a$ and $\pi_{2,2} \circ I_a$ are both identity homomorphisms. This means the $2,2$ and

$3,3$ component of $I_a(e)$ are exactly e for all $e \in \mathsf{KA}(X)$. Let

$$\begin{bmatrix} p & q & r \\ 0 & e & 0 \\ 0 & 0 & e \end{bmatrix} \triangleq I_a(e) \in D_a(\mathsf{KA}(X))_{u_a}.$$

We can define the derivative δ_a and residual ρ_a as follows:

$$\delta_a(e) \triangleq r, \quad \rho_a(e) \triangleq p + q.$$

Then the following corollary can be derived simply from the definition of $D_a(\mathsf{KA}(X))_{u_a}$.

Corollary 6 (decomposition)**.** *For all expressions $e \in \mathsf{KA}(X)$, primitives $a \in X$, and word $w \in \mathcal{L}_X$,*

$$e = \rho_a(e) + a \cdot \delta_a(e) \text{ and } a \cdot w \notin L(\rho_a(e)).$$

Theorem 6 (Soundness Property)**.** *For a primitive a in a commutable set X, let the derivative on language δ_a defined as $\delta_a(l) \triangleq \{s \mid a \cdot w \in l\}$, the following diagram commute:*

$$\begin{array}{ccc} \mathsf{KA}(X) & \xrightarrow{\delta_a} & \mathsf{KA}(X) \\ \downarrow{\scriptstyle L} & & \downarrow{\scriptstyle L} \\ \mathcal{L}_X & \xrightarrow{\delta_a} & \mathcal{L}_X \end{array}$$

Proof. Given any word $w \in \mathcal{L}_X$ and in $e \in \mathsf{KA}(X)$:

$$\begin{aligned} & s \in \delta_a(L(e)) \\ \iff & a \cdot w \in L(e) && \text{by definition of language } \delta_a \\ \iff & a \cdot w \in L(\rho_a(e) + a \cdot \delta_a(e)) && \text{Corollary 6} \\ \iff & a \cdot w \in L(a \cdot \delta_a(e)) && a \cdot w \notin L(\rho_a(e)) \\ \iff & s \in L(\delta_a(e)). \end{aligned}$$

Thus, for all $e \in \mathsf{KA}(X)$, $\delta_a(L(e)) = L(\delta_a(e))$, we have reached our conclusion. □

Finally, we prove a Galois connection that the derivative is expected to satisfy.

Lemma 1 (Basic Algebraic Properties)**.** *Following basic algebraic properties are true,*

for all primitive a *and expressions* e, e'*:*

$$\begin{aligned}
\delta_a(ae) &= e \\
\rho_a(ae) &= 0 \\
\delta_a(\rho_a(e)) &= 0 \\
e \geq e' &\Longrightarrow \delta_a(e) \geq \delta_a(e') \\
e \geq e' &\Longrightarrow \rho_a(e) \geq \rho_a(e')
\end{aligned}$$

Proof. We first compute $I_a(ae)$:

$$I_a(ae) = I_a(a) \cdot I_a(e) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix} \begin{bmatrix} p & q & r \\ 0 & e & 0 \\ 0 & 0 & e \end{bmatrix},$$

for some expressions $p, q, r \in \mathsf{KA}(X)$. Then

$$I_a(ae) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & a & 0 \\ 0 & 0 & a \end{bmatrix} \begin{bmatrix} p & q & r \\ 0 & e & 0 \\ 0 & 0 & e \end{bmatrix} = \begin{bmatrix} 0 & 0 & e \\ 0 & ae & 0 \\ 0 & 0 & ae \end{bmatrix}.$$

Therefore, we obtain the conclusion $\delta_a(ae) = e$ and $\rho_a(ae) = 0$.

Notice that $\rho(e) \leq X^*_{\sim a} + X^*_{\sim a} X_{\nsim a} X^*$, therefore

$$\begin{aligned}
I_a(\rho(e)) &\leq I_a(X^*_{\sim a} + X^*_{\sim a} X_{\nsim a} X^*) \\
&= \begin{bmatrix} X^*_{\sim a} & X^*_{\sim a} X_{\nsim a} X^* & 0 \\ 0 & X^*_{\sim a} + X^*_{\sim a} X_{\nsim a} X^* & 0 \\ 0 & 0 & X^*_{\sim a} + X^*_{\sim a} X_{\nsim a} X^* \end{bmatrix}
\end{aligned}$$

Therefore $I_a(\rho(e)) \leq 0$, and since 0 is the smallest element, We obtain the conclusion $I_a(\rho(e)) = 0$.

The monotonicity can be derived from the monotonicity of I_a. When $e \geq e'$, we have $I_a(e) \geq I_a(e')$. Recall that the ordering on matrices are component order, since $\delta_a(e)$ and $\rho_a(e)$ are either component of $I_a(e)$ or the sum of components of $I_a(e)$, therefore $\delta_a(e) \geq \delta_a(e')$ and $\rho_a(e) \geq \rho_a(e')$. □

Theorem 7 (Galois Connection)**.** *Given a commutative set* X*, for all expression* $e, e' \in \mathsf{KA}(X)$ *and primitive* $a \in X$*,*

$$ae \leq e' \Longleftrightarrow e \leq \delta_a(e').$$

Proof. We first show $ae \leq e' \iff e \leq \delta_a(e')$: $\implies$ direction can be proved by applying δ_a to both sides:

$$ae \leq e' \implies \delta_a(ae) \leq \delta_a(e') \implies e \leq \delta_a(e').$$

$\impliedby$ direction proven by multiplying a on both sides:

$$e \leq \delta_a(e') \implies ae \leq a\delta_a(e') \implies ae \leq a\delta_a(e') \leq e'.$$

□

2.2.3 Decidability And Completeness

In this section, we prove the completeness and decidability of the word inhabitance problem, by explicitly define an algorithm to check for word inhabitance.

Theorem 8 (Decidability and Completeness)**.** *Given a word $w \in \mathsf{Wrd}(X)$ and an expression $e \in \mathsf{KA}(X)$, we can define the following algorithm to test for inhabitants:*

$$\begin{aligned} &i : \mathsf{Wrd}(X) \times \mathsf{KA}(X) \to 2 \\ &i(\epsilon, e) \triangleq E(e) \\ &i(a \cdot w, e) \triangleq i(w, \delta_a(e)) \end{aligned}$$

Such an algorithm will always terminate, and it is sound:

$$w \in L(e) \iff i(w, e) = 1 \iff w \leq e.$$

Proof. The algorithm i will terminate because both δ_a and E can be computed by computing the interpretation I_a and I_E.

We first show $w \in L(e) \iff i(w, e)$ by induction on w.

- If $w = \epsilon$, then $\epsilon \in L(e) \iff E(e) = 1$ by soundness of E.
- If $w = a \cdot w'$ then:

$$\begin{aligned} a \cdot w' \in L(e) &\iff w' \in \delta_a(L(e)) && \text{definition} \\ &\iff w' \in L(\delta_a(e)) && \text{soundness of } \delta_a \\ &\iff i(w, \delta_a(e)) && \text{induction hypothesis} \end{aligned}$$

We then show $i(w, e) = 1 \iff w \leq e$ by induction on w.

- If $w = \epsilon$, then $i(\epsilon, e) = E(e)$. When $E(e) = 1$, then $1 = E(e) \leq e$; When $E(e) = 0$, then $1 \not\leq e$, because $1 \leq e$ is not true in the language interpretation.
- If $w = a \cdot w'$ then:

$$\begin{aligned} a \cdot w' \leq e &\iff w' \leq \delta_a(e) && \text{Theorem 7} \\ &\iff i(w, \delta_a(e)) && \text{induction hypothesis} \end{aligned}$$

□

2.2.4 Fundamental Theorem

Fundamental theorem is an important soundness condition for the definition of derivative and empty word predicate, it also exhibits a strong connection between KA and automata (**?**; **?**). Because of the significance of the fundamental theorem, we decide to prove it for KA with atomic commutativity, despite it is not used in the rest of the paper.

In order to show the fundamental theorem for KA with atomic commutativity, we will establish the relation between derivative and empty word predicate in KA with their counterparts in KA with commutativity, and with this relation, we can show that fundamental theorem KA implies the fundamental theorem in KA with commutativity

Recall that for all commutable set X, we can construct a discrete commutable set $X_\asymp$ by replacing the commuting relation in X with the identity relation. Notice that $\mathsf{KA}(X_\asymp)$ is a free KA, and by **??**, derivative and empty word predicate is unique on free KAs. Since our definition of E and δ_a are sound, therefore they are exactly the conventional E and δ_a when applied to a term in the free KA. Therefore, the

fundamental theorem holds for $\mathsf{KA}(X_\sim)$:

$$\forall e_\sim \in \mathsf{KA}(X_\sim), e_\sim = E(e_\sim) + \sum_{a \in X} a \cdot \delta_a(e_\sim).$$

Finally, there is a canonical KA homomorphism from $\mathsf{KA}(X_\sim)$ to $\mathsf{KA}(X)$, by lifting the following action on the primitives:

$$[a]_\sim \triangleq a.$$

This KA homomorphism imposes the commutativity of X to the input expression, and this homomorphism is surjective.

Lemma 2. *Consider a Kleene Algebra $\mathcal{K}$ and an element $t \in \mathcal{K}$, there is a homomorphism:*

$$h : D_t(\mathcal{K}) \to M_2(\mathcal{K})$$
$$h(\begin{bmatrix} p & q & r \\ 0 & s & 0 \\ 0 & 0 & s \end{bmatrix}) = \begin{bmatrix} p & r \\ 0 & s \end{bmatrix}$$

Proof. Perseverance of identities and addition is trivial, we will only check for perseverance of multiplication and star

The multiplication case:

$$\begin{aligned} & h(\begin{bmatrix} p_1 & q_1 & r_1 \\ 0 & s_1 & 0 \\ 0 & 0 & s_1 \end{bmatrix} \begin{bmatrix} p_2 & q_2 & r_2 \\ 0 & s_2 & 0 \\ 0 & 0 & s_2 \end{bmatrix}) \\ = & h(\begin{bmatrix} p_1 p_2 & p_1 q_2 + q_1 s_2 & p_1 r_2 + r_1 s_2 \\ 0 & s_1 s_2 & 0 \\ 0 & 0 & s_1 s_2 \end{bmatrix}) \\ = & \begin{bmatrix} p_1 p_2 & p_1 r_2 + r_1 s_2 \\ 0 & s_1 s_2 \end{bmatrix} \\ = & h(\begin{bmatrix} p_1 & q_1 & r_1 \\ 0 & s_1 & 0 \\ 0 & 0 & s_1 \end{bmatrix}) \cdot h(\begin{bmatrix} p_2 & q_2 & r_2 \\ 0 & s_2 & 0 \\ 0 & 0 & s_2 \end{bmatrix}). \end{aligned}$$

The star case:

$$h(\begin{bmatrix} p & q & r \\ 0 & s & 0 \\ 0 & 0 & s \end{bmatrix}^*) = h(\begin{bmatrix} p^* & p^*qs^* & p^*rs^* \\ 0 & s^* & 0 \\ 0 & 0 & s^* \end{bmatrix})$$
$$= \begin{bmatrix} p^* & p^*rs^* \\ 0 & s^* \end{bmatrix} = h(\begin{bmatrix} p & q & r \\ 0 & s & 0 \\ 0 & 0 & s \end{bmatrix})^*.$$

□

Since the derivative $\delta_a(e)$ is defined as the $1,3$ component of $I_a(e)$, after we apply above homomorphism h to $I_a(e)$, the derivative $\delta_a(e)$ becomes the $1,2$ component of the matrix:

$$\forall e \in \mathsf{KA}(X), \delta_a(e) = \pi_{1,2}(h(I_a(e))).$$

Lemma 3. *Let X be a commutable set, for all non-commutativity expressions $e_\asymp \in \mathsf{KA}(X_\asymp)$:*

$$E([e_\asymp]_\sim) = [E(e_\asymp)]_\sim, \quad \delta_a([e_\asymp]_\sim) \geq [\delta_a(e_\asymp)]_\sim.$$

Proof. Consider the following interpretations

$$I_E \circ [-]_\sim \text{ and } [-]_\sim \circ I_E : \mathsf{KA}(X_\asymp) \to \mathsf{KA}(X).$$

Their actions coincide on the primitives:

$$\forall a \in X_\asymp, I_E([a]_\sim) = \begin{bmatrix} 0 & a \\ 0 & a \end{bmatrix} = [E(a)]_\sim.$$

Therefore, by Theorem 3,

$$\forall e_\asymp \in \mathsf{KA}(X_\asymp), I_E([e_\asymp]_\sim) = [I_E(e_\asymp)]_\sim.$$

Since E is a component of I_E,

$$E([e_\asymp]_\sim) = [E(e_\asymp)]_\sim.$$

The same can be done for derivatives. We consider the following interpretations:

$$h \circ I_a \circ [-]_\sim \text{ and } [-]_\sim \circ h \circ I_a : \mathsf{KA}(X_\asymp) \to \mathsf{KA}(X).$$

Given a primitive b,

- if $b = a$, then

$$\forall b \in X_{\asymp}, h(I_a([b]_\sim)) = \begin{bmatrix} 0 & 1 \\ 0 & b \end{bmatrix} = [h(I_a(b))]_\sim;$$

- if $b \neq a$, then

$$h \circ I_a([b]_\sim) = \begin{cases} \begin{bmatrix} 0 & 0 \\ 0 & b \end{bmatrix} & b \asymp a \text{ in } X \\ \begin{bmatrix} b & 0 \\ 0 & b \end{bmatrix} & b \sim a \text{ in } X \end{cases}$$

  both of which are greater than

$$[h(I_a(b))]_\sim = \begin{bmatrix} 0 & 0 \\ 0 & b \end{bmatrix} \quad \text{because } b \asymp a \text{ in } X_{\asymp}.$$

Therefore, for all primitive $b \in X_{\asymp}$, $h(I_a([b]_\sim)) \geq [h(I_a(b))]_\sim$. By Theorem 3

$$\forall e_{\asymp} \in \mathsf{KA}(X_{\asymp}), h \circ I_a([e_{\asymp}]_\sim) \geq [h \circ I_a(e_{\asymp})]_\sim.$$

Since δ_a is a component of $h \circ I_a$,

$$\delta_a([e_{\asymp}]_\sim) \geq [\delta_a(e_{\asymp})]_\sim. \qquad \square$$

The above lemma state that the derivative for commutative expression in $\mathsf{KA}(X)$ is always larger than their non-commutativity counterparts in $\mathsf{KA}(X_{\asymp})$. Since $\mathsf{KA}(X_{\asymp})$ is a free KA, we can easily derive the fundamental theorem for $\mathsf{KA}(X)$ using its connection with $\mathsf{KA}(X_{\asymp})$.

Theorem 9 (Fundamental Thoerem)**.** *For all $e \in \mathsf{KA}(X)$, the following equality holds:*

$$e = E(e) + \sum_{a \in X} a \cdot \delta_a(e)$$

Proof. We first show $e \geq E(e) + \sum_{a \in X} a \cdot \delta_a(e)$, which is a direct consequence of

decompositions in corollaries 4 and 6:

$$e \geq E(e),$$
$$e \geq a \cdot \delta_a(e), \forall a \in X.$$

We then show $e \leq E(e) + \sum_{a \in X} a \cdot \delta_a(e)$. Since $[-]_\sim$ is a surjective homomorphism, we consider $e_\asymp \in \mathsf{KA}(X_\asymp)$ s.t. $[e_\asymp]_\sim = e$. Because $\mathsf{KA}(X_\asymp)$ is a free KA, fundamental theorem holds for $e_\asymp$:

$$e_\asymp \leq E(e_\asymp) + \sum_{a \in X} a \cdot \delta_a(e_\asymp).$$

We can apply the homomorphism $[-]_\sim$ to both sides:

$$e \leq [E(e_\asymp)]_\sim + \sum_{a \in X} a \cdot [\delta_a(e_\asymp)]_\sim.$$

By Lemma 3,

$$[E(e_\asymp)]_\sim = E([e_\asymp]_\sim) = E(e),$$
$$[\delta_a(e_\asymp)]_\sim \leq \delta_a([e_\asymp]_\sim) = \delta_a(e).$$

Thus, obtain the desired inequality:

$$e \leq [E(e_\asymp)]_\sim + \sum_{a \in X} a \cdot [\delta_a(e_\asymp)]_\sim \leq E(e) + \sum_{a \in X} a \cdot \delta_a(e).$$

□

2.3 Undecidability

In this section, we will show the undecidability result for general Kleene Algebra equalities with atomic commutativity hypotheses. The undecidability result is obtained by using a proof to simulate the execution of a two-counter machine. From there, we can encode state reachability of terminating two-counter machines into an KA inequality. Enabling us to carry out a diagonal argument, similar to the proof of undecidability of halting problem.

2.3.1 Encoding Two-Counter Machines

Counter machine is a well-studied machine (**?**; **?**; **?**), and it can simulate any Turing machine (**?**, Theorem 14.1-1) with just two counters. In this paper, we only consider two-counter machines. A two-counter machine $M \triangleq (S, \hat{s}, \iota)$ consists of a finite set of *state* S, a start state $\hat{s} \in S$, and each state is equipped with an instruction $\iota : S \to I_S$, where instructions I_S is defined as follows:

$$\begin{aligned} I_S \triangleq & \{\texttt{Inc}(s, q) \mid s \in \{1, 2\}, q \in Q\} \\ & \cup \{\texttt{Dec}(s, q) \mid s \in \{1, 2\}, q \in Q\} \\ & \cup \{\texttt{If}(s, q_1, q_2) \mid s \in \{1, 2\}, q_1, q_2 \in Q\} \\ & \cup \{\texttt{Halt}\}. \end{aligned}$$

Each instruction has a semantics, we define $S_\perp \triangleq S + \{\perp\}$:

$$\begin{aligned} [\![i]\!] & : \mathbb{N} \times \mathbb{N} \to S_\perp \times \mathbb{N} \times \mathbb{N} \\ [\![\texttt{Inc}(1, s)]\!](n, m) & \triangleq (s, n + 1, m) \\ [\![\texttt{Inc}(2, s)]\!](n, m) & \triangleq (s, n, m + 1) \\ [\![\texttt{Dec}(1, s)]\!](n, m) & \triangleq (s, \max(n - 1, 0), m) \\ [\![\texttt{Dec}(2, s)]\!](n, m) & \triangleq (s, n, \max(m - 1, 0)) \\ [\![\texttt{If}(1, q_1, q_2)]\!](n, m) & \triangleq \begin{cases} (s_1, n, m) & \text{if } n = 0 \\ (s_2, n, m) & \text{if } n \neq 0 \end{cases} \\ [\![\texttt{If}(2, s_1, s_2)]\!](n, m) & \triangleq \begin{cases} (s_1, n, m) & \text{if } m = 0 \\ (s_2, n, m) & \text{if } m \neq 0 \end{cases} \\ [\![\texttt{Halt}]\!](n, m) & \triangleq (\perp, n, m). \end{aligned}$$

From the semantics of the instruction, we can define a transition relation for any

machine M:

$$R_M \in (S \times \mathbb{N} \times \mathbb{N}) \times (S_\perp \times \mathbb{N} \times \mathbb{N})$$
$$R_M \triangleq \{((s, m, n), [\![\iota(s)]\!](m, n)) \mid s \in S; m, n \in \mathbb{N}\}.$$

Note that R_M is a functional relation, that is for all input $(s, m, n) \in (S \times \mathbb{N} \times \mathbb{N})$ there exists a unique element in $S_\perp \times \mathbb{N} \times \mathbb{N}$ relating to it. We call the elements in $S_\perp \times \mathbb{N} \times \mathbb{N}$ *configurations* of the machine M. Let R_M^* be the reflexive transition closure for R_M, and we write $c \to^* c'$ if $(c, c') \in R_M^*$. We say a state $s \in S$ is *reachable* from input (n, m) when there exists $n', m' \in \mathbb{N}$, s.t. $(\hat{s}, m, n) \to^* (s, n', m')$.

Finally, we can consider a machine as a partial function, we say that $M(m, n)$ returns (m', n') when $(\hat{s}, m, n) \to^* (\perp, n', m')$. Since complex data structure can be encoded as a pair of numbers using the classical Gödel numbers, we will abuse the notation to say $M(i)$ returns o for input i and output o of arbitrary type, not just pairs of numbers.

Given a two-counter machine with finite state set S, We define the set $\Sigma = S + \{\perp, a, b\}$ and the following commutable set $\ddot{\Sigma}$:

- the carrier is $\langle\Sigma| \cup |\Sigma\rangle$, where

$$\langle\Sigma| \triangleq \{\sigma_l \mid \sigma \in \Sigma\} \text{ and } |\Sigma\rangle \triangleq \{\sigma_r \mid \sigma \in \Sigma\};$$

- the commuting relation is $\langle\sigma| \sim |\sigma'\rangle$, where $\sigma, \sigma' \in \Sigma$.

We call primitives in $\langle\Sigma|$ *left primitives* and primitives in $|\Sigma\rangle$ *right primitives*. This definition of commutativity is similar to BiKA (**?**), however instead of using an underlying KA with two homomorphisms, we simply impose the commutativity onto the primitives. We consider Σ as a discrete commutable set, therefore we can define the free Kleene algebra $\mathsf{KA}(\Sigma)$ and the free KA with atomic commutative $\mathsf{KA}(\ddot{\Sigma})$.

There are three function we can define from Σ to $\mathsf{Wrd}(\ddot{\Sigma})$:

$$\langle - | : \Sigma \to \mathsf{Wrd}(\ddot{\Sigma}) \qquad | - \rangle : \Sigma \to \mathsf{Wrd}(\ddot{\Sigma}) \qquad \langle - \rangle : \Sigma \to \mathsf{Wrd}(\ddot{\Sigma})$$
$$\langle \sigma | \triangleq \sigma_l \qquad | \sigma \rangle \triangleq \sigma_r \qquad \langle \sigma \rangle \triangleq \sigma_l \cdot \sigma_r.$$

These maps can be lifted to monoidal homomorphism on words $\mathsf{Wrd}(\ddot{\Sigma}) \to \mathsf{Wrd}(\ddot{\Sigma})$: $\langle w|$ is the word w with all primitives replaced by corresponding left primitives and $|w\rangle$ is the word w with all primitives replaced by corresponding right primitives.

By composing $\langle -|, |-\rangle, \langle - \rangle$ with the natural monoidal embedding $\mathsf{Wrd}(\ddot{\Sigma}) \to \mathsf{KA}(\ddot{\Sigma})$, where the monoidal operation of $\mathsf{KA}(\ddot{\Sigma})$ is multiplication with identity 1, we can obtain functions in $\Sigma \to \mathsf{KA}(\ddot{\Sigma})$. Similarly, these maps can be lifted to KA homomorphisms $\mathsf{KA}(\Sigma) \to \mathsf{KA}(\ddot{\Sigma})$:

- $\langle e|$ and $|e\rangle$ will replace all the primitives in e with their respective left primitives or right primitives.
- $\langle e \rangle$ will produce two expression with *matching* left and right primitives.

We will abbreviate the multiplication $\langle e_1| \cdot |e_2\rangle$ as $\langle e_1 | e_2 \rangle$.

Example 1. *Consider $a \in \Sigma$, then $\langle a^* \rangle \in \mathsf{KA}(\ddot{\Sigma})$ and*

$$L(\langle a^* \rangle) = \{\langle a^n | a^n \rangle \mid n \in \mathbb{N}\}.$$

More generally, given an expression $e \in \mathsf{KA}(\Sigma)$, then $\langle e \rangle \in \mathsf{KA}(\ddot{\Sigma})$ and

$$L(\langle e \rangle) = \{\langle w | w \rangle \mid w \in L(e)\}.$$

For a word in $\mathsf{Wrd}(\ddot{\Sigma})$ we can always canonically separate it into its left and right components, this separation gives a normal form for all the words in $\mathsf{Wrd}(\ddot{\Sigma})$.

Definition 1. *For a word w, the* left component *$\langle w_l|$ is the word formed by all the left primitives in its original order, and the* right component *$|w_r\rangle$ is the word formed by all the right primitives in its original order. The concatenation of the left*

component and the right component is equal to the original word. Therefore, for all word $w, w' \in \mathsf{Wrd}(\ddot{\Sigma})$,

$$w = w' \iff w_r = w'_r \wedge w_l = w'_l.$$

Example 2. *Consider the following word*

$$w \triangleq \langle s|s'\rangle \cdot \langle a^m b^n | a^{m+1} b^n\rangle,$$

then its left component $\langle w_l| = \langle s a^m b^n|$ *and the right component is* $|w_r\rangle = |s' a^{m+1} b^n\rangle$. *The concatenation of right and left component is equal to the original word:*

$$\langle w_l|w_r\rangle = \langle s a^m b^n | s' a^{m+1} b^n\rangle = \langle s|s'\rangle \cdot \langle a^m b^n | a^{m+1} b^n\rangle.$$

We first show couple useful lemmas about $\mathsf{KA}(\Sigma)$ and $\mathsf{KA}(\ddot{\Sigma})$:

Lemma 4. *We first prove several lemmas that will be useful in our derivation. For all* $e, e_1, e_2, e'_1, e'_2 \in \mathsf{KA}(\Sigma)$

$$\langle e_1|e_2\rangle = |e_2\rangle \cdot \langle e_1|; \tag{2.1}$$

$$\langle e_1|e_2\rangle \leq \langle e'_1|e'_2\rangle \iff e_1 \leq e'_1 \wedge e_2 \leq e'_2; \tag{2.2}$$

$$\langle e^*\rangle \leq \langle e^*|e^*\rangle. \tag{2.3}$$

Proof. $\langle e_1|e_2\rangle = |e_2\rangle \cdot \langle e_1|$ can be derived by induction on structure of e_1, the only non-trivial case is the star case, which can be proven using the induction rule.

The equivalence

$$\langle e_1|e_2\rangle \leq \langle e'_1|e'_2\rangle \iff e_1 \leq e'_1 \wedge e_2 \leq e'_2$$

can be derived as follows: The $\Longleftarrow$ part can be shown by multiplication preserves order. The $\Longrightarrow$ part can be shown by looking at the language interpretation: if $L(e_1) \not\leq L(e'_1)$ or $L(e_2) \not\leq L(e'_2)$, then we can derive

$$L(\langle e_1|e_2\rangle) \not\leq L(\langle e'_1|e'_2\rangle)$$

$\langle e^* \rangle \leq \langle e^* | e^* \rangle$ can be shown by induction rule, because

$$\langle e^* | e^* \rangle \geq 1,$$
$$\langle e^* | e^* \rangle \geq \langle e e^* | e e^* \rangle = \langle e | e \rangle \cdot \langle e^* | e^* \rangle = \langle e \rangle \cdot \langle e^* | e^* \rangle,$$

by induction rule $\langle e^* | e^* \rangle \geq \langle e^* \rangle$. □

We can encode components of the machine as Kleene Algebra terms.

Definition 2. *For simplicity, we will implicitly coerce all the configuration into an expression in* $\mathsf{KA}(\Sigma)$ *in the following way: for* $s \in S_\perp$*,*

$$(s, m, n) \mapsto s a^m b^n.$$

We interpret each instruction $i \in I_S$ *as an element* $[i] \in \mathsf{KA}(\ddot{\Sigma})$ *as follows.*

$$\begin{aligned}
[\mathtt{Inc}(1, s)] &\triangleq |s\rangle\langle a^*\rangle|a\rangle\langle b^*\rangle \\
[\mathtt{Inc}(2, s)] &\triangleq |s\rangle\langle a^* b^*\rangle|b\rangle \\
[\mathtt{Dec}(1, s)] &\triangleq |s\rangle\langle a^*\rangle\langle a|\langle b^*\rangle + |s\rangle\langle b^*\rangle \\
[\mathtt{Dec}(2, s)] &\triangleq |s\rangle\langle a^* b^*\rangle\langle b| + |s\rangle\langle a^*\rangle \\
[\mathtt{If}(1, s_1, s_2)] &\triangleq |s_1\rangle\langle b^*\rangle + |s_2\rangle\langle a^+\rangle\langle b^*\rangle \\
[\mathtt{If}(2, s_1, s_2)] &\triangleq |s_1\rangle\langle a^*\rangle + |s_2\rangle\langle a^*\rangle\langle b\rangle^+ \\
[\mathtt{Halt}] &\triangleq |\perp\rangle\langle a^* b^*\rangle.
\end{aligned}$$

The transition relation is encoded as $R_M \in \mathsf{KA}(\ddot{\Sigma})$*:*

$$R_M \triangleq \sum_{s \in S} \langle s| \cdot [\iota(s)].$$

The reason for such encoding is apparent when we look the language interpretation:

Corollary 7 (Language Soundness)**.** *Given a machine* $M \triangleq (S, \hat{s}, \iota)$*,* $c \in S \times \mathbb{N} \times \mathbb{N}$*, and* $c' \in S_\perp \times \mathbb{N} \times \mathbb{N}$*, the following equivalence holds*

$$[\![i]\!](m, n) = c' \iff \langle a^m b^n | c' \rangle \in L([i]); \tag{2.4}$$
$$c \to c' \iff \langle c | c' \rangle \in L(R_M). \tag{2.5}$$

Proof. The equivalence

$$[\![i]\!](m,n) = c' \iff \langle a^m b^n | c' \rangle \in L([i])$$

can be shown by looking at each case of instruction i, and explicitly compute the language model for each instruction.

Therefore,

$$\begin{aligned}
&(s,m,n) \to (s',m',n') \\
&\iff (s',m',n') = [\![\iota(s)]\!](m,n) \\
&\iff \langle a^m b^n | s' a^{m'} b^{n'} \rangle \in L([\iota(s)]) \\
&\iff \langle s a^m b^n | s' a^{m'} b^{n'} \rangle \in L(\langle s| \cdot [\iota(s)]) \\
&\iff \langle s a^m b^n | s' a^{m'} b^{n'} \rangle \in L(R_M).
\end{aligned}$$

□

Corollary 8. *Because the transition relation of any machine $M = (S, \hat{s}, \iota)$ is functional, therefore for all word $w \in L(Sa^*b^*)$, there exists a word $w' \in L(S_\perp a^* b^*)$ s.t. $\langle w | w' \rangle \in L(R_M)$.*

2.3.2 From Reachability to Undecidability

Our undecidability result relies on an equivalence between provability of a KA inequality and state reachability in a certain kind of machine. In order to obtain such equivalence, we will start with the provability of a single transition.

We will first define two useful terms. For a machine $M \triangleq (S, \hat{s}, \iota)$, and a subset of states $S' \subseteq S$, we can define all the configuration for S' including termination:

$$C_{S'} \in \mathsf{KA}(\Sigma)$$

$$C_{S'} \triangleq S'_\perp a^* b^*.$$

And the term N_S representing left-right configuration mismatch:

$$
\begin{aligned}
&N_S \in \mathsf{KA}(\ddot{\Sigma}) \\
&N_S \triangleq \sum_{s \neq s' \in S} \langle s|s'\rangle(\langle a+b| + |a+b\rangle)^* && \text{state mismatch} \\
&\quad + \langle S\rangle\langle a^*\rangle(\langle a|^+ + |a\rangle^+)\langle b^*|b^*\rangle && \text{counter } a \text{ mismatch} \\
&\quad + \langle S\rangle\langle a^*b^*\rangle(\langle b|^+ + |b\rangle^+) && \text{counter } b \text{ mismatch}
\end{aligned}
$$

And we can use these terms to bound the encoding of transition R_M:

Lemma 5. *For all instructions $i \in I_S$ and transition $R_M \in \mathsf{KA}(\ddot{\Sigma})$, let S be the state set of M:*

$$[i] \leq \langle a^*b^*|C_S\rangle = \langle a^*b^*|S_\perp \cdot a^*b^*\rangle; \tag{2.6}$$

$$R_M \leq \langle C_S \mid C_S\rangle = \langle S_\perp \cdot a^*b^*|S_\perp \cdot a^*b^*\rangle. \tag{2.7}$$

Proof. The first inequality can be proven by looking at each case of i, and apply Inequality (2.3) when necessary.

The second inequality can be proven by unfolding the definition of R_M:

$$
\begin{aligned}
R_M &= \sum_{s\in S}\langle s| \cdot [\iota(s)] \\
&\leq \sum_{s\in S}\langle s| \cdot \langle a^*b^*|S_\perp \cdot a^*b^*\rangle && \text{by Inequality (2.6)} \\
&\leq \langle S \cdot a^*b^*|S_\perp \cdot a^*b^*\rangle \leq \langle C_S \mid C_S\rangle.
\end{aligned}
$$

□

With the above tools in place, we will establish the connection of provability and a single transition.

Theorem 10 (provability of single transition)**.** *In any machine M, $(s,m,n) \to (s',m',n')$ if and only if the following inequality is provable:*

$$|sa^mb^n\rangle R_M \leq \langle C_S\rangle \cdot |s'a^{m'}b^{n'}\rangle + N_S \cdot |C_S\rangle.$$

Proof. To prove the $\Longrightarrow$ direction, we will first unfold definition of the left-hand side:

$$\begin{aligned}
&|sa^m b^n\rangle R_M \\
&= |sa^m b^n\rangle \cdot (\sum_{s_1 \in S} \langle s_1| \cdot [\iota(s_1)]) \\
&= \sum_{s_1 \in S} |sa^m b^n\rangle \cdot \langle s_1| \cdot [\iota(s_1)] \\
&= \langle s|sa^m b^n\rangle \cdot [\iota(s)] + \sum_{s_1 \neq s} \langle s_1|sa^m b^n\rangle \cdot [\iota(s_1)].
\end{aligned}$$

It suffices to show the following inequalities:

$$\begin{aligned}
\langle s|sa^m b^n\rangle \cdot [\iota(s)] &\leq \langle C_S\rangle|s' a^{m'} b^{n'}\rangle + N_S|C_S\rangle; \\
\sum_{s_1 \neq s} \langle s_1|sa^m b^n\rangle \cdot [\iota(s_1)] &\leq \langle C_S\rangle|s' a^{m'} b^{n'}\rangle + N_S|C_S\rangle.
\end{aligned}$$

We show the second inequality first:

$$\begin{aligned}
&\sum_{s_1 \neq s} \langle s_1|sa^m b^n\rangle \cdot [\iota(s_1)] \\
&\leq \sum_{s_1 \neq s} \langle s_1|sa^m b^n\rangle \cdot \langle a^* b^*|C_S\rangle \\
&\leq \sum_{s \neq s' \in S} \langle s|s'\rangle(\langle a+b| + |a+b\rangle)^*|C_S\rangle \\
&\leq N_S|C_S\rangle \leq \langle C_S\rangle|s' a^{m'} b^{n'}\rangle + N_S|C_S\rangle,
\end{aligned}$$

To show the first inequality:

$$\langle s|sa^m b^n\rangle \cdot [\iota(s)] \leq \langle C_S\rangle|s' a^{m'} b^{n'}\rangle + N_S|C_S\rangle.$$

We need to look at all the cases for $\iota(s)$, we show the $\texttt{If}(1, s)$ case as examples:

- If $\iota(s) = \texttt{If}(1, s_1', s_2')$ and $m = 0$, then $(s, 0, n) \to (s_1', 0, n)$:

$$\begin{aligned}
&\langle s|sb^n\rangle \cdot [\iota(s)] \\
&= \langle s|sb^n\rangle \cdot (|s_1'\rangle\langle b^*\rangle + |s_2'\rangle\langle a^+\rangle\langle b^*\rangle) \\
&= \langle s|sb^n\rangle \cdot |s_1'\rangle \cdot (\sum_{i\le n}\langle b^i\rangle + \langle b^n\rangle \cdot \langle b^+\rangle) \\
&\quad + \langle s|sb^n\rangle \cdot |s_2'\rangle \cdot \langle a^+\rangle\langle b^*\rangle \\
&= (\sum_{i<n}\langle sb^i|sb^n\rangle|s_1'b^i\rangle) + \langle sb^n|sb^n\rangle \cdot |s_1'b^n\rangle \\
&\quad + \langle sb^nb|sb^n\rangle|s_1'b^nb\rangle\langle b^*\rangle \\
&\quad + \langle sa|sb^n\rangle \cdot |s_2'a\rangle \cdot \langle a^*b^*\rangle
\end{aligned}$$

  Notice:

$$\begin{aligned}
\sum_{i<n}\langle sb^i|sb^n\rangle|s_1'b^i\rangle &\le \langle S\rangle\langle a^*b^*\rangle|b^+\rangle \cdot |C_S\rangle \le N_S \cdot |C_S\rangle; \\
\langle sb^n|sb^n\rangle \cdot |s_1'b^n\rangle &\le \langle C_S\rangle \cdot |s_1'b^n\rangle; \\
\langle sb^nb|sb^n\rangle|s_1'b^nb\rangle\langle b^*\rangle &\le \langle sb^nb|sb^n\rangle|s_1'b^nb\rangle\langle b^*|b^*\rangle \\
&\le \langle sb^nbb^*|sb^n\rangle|s_1'b^nbb^*\rangle \\
&\le \langle S\rangle\langle a^*b^*\rangle\langle b^+| \cdot |C_S\rangle \le N_S \cdot |C_S\rangle; \\
\langle sa|sb^n\rangle \cdot |s_2'a\rangle \cdot \langle a^*b^*\rangle &\le \langle sa|sb^n\rangle \cdot |s_2'a\rangle \cdot \langle a^*|a^*\rangle\langle b^*|b^*\rangle \\
&\le \langle saa^*b^*|sb^n\rangle \cdot |s_2'ab^*a^*\rangle \\
&\le \langle S\rangle\langle a^*\rangle\langle a|^+\langle b^*|b^*\rangle \cdot |C_S\rangle \\
&\le N_S \cdot |C_S\rangle.
\end{aligned}$$

  Thus, we obtain the inequality we desire:

$$|sa^mb^n\rangle \cdot \langle s| \cdot [\iota(s)] \le \langle C_S\rangle \cdot |s_1'b^n\rangle + N_S|C_S\rangle.$$

- If $\iota(s) = \texttt{If}(1, s_1', s_2')$ and $m \neq 0$, then $(s, m, n) \to (s_2', m, n)$, and the proof is

similar to above:

$$
\begin{aligned}
&\langle s|sa^m b^n\rangle \cdot [\iota(s)] \\
&= \langle s|sa^m b^n\rangle \cdot (|s_1'\rangle\langle b^*\rangle + |s_2'\rangle\langle a^+ b^*\rangle) \\
&= \langle s|sa^m b^n\rangle \cdot |s_1'\rangle\langle b^*\rangle \\
&\qquad + \langle s|sa^m b^n\rangle \cdot |s_2'\rangle \cdot (\sum_{i\le m, j\le n} \langle a^i b^j\rangle + \langle a^m a^+ b^n b^+\rangle) \\
&= \langle s|sa^m b^n\rangle \cdot |s_1'\rangle\langle b^*\rangle \\
&\qquad + (\sum_{i<m, j\le n} \langle sa^i b^j|sa^m b^n\rangle \cdot |s_2' a^i b^j\rangle) \\
&\qquad + (\sum_{i=m, j\le n} \langle sa^i b^j|sa^m b^n\rangle \cdot |s_2' a^i b^j\rangle) \\
&\qquad + \langle sa^m b^n|sa^m b^n\rangle \cdot |s_2' a^m b^n\rangle \\
&\qquad + \langle sa^m a^+ b^n b^+|sa^m b^n\rangle \cdot |s_1' a^m a^+ b^n b^+\rangle
\end{aligned}
$$

Then we obtain the following inequality

$$
\begin{aligned}
\langle s|sa^m b^n\rangle \cdot |s_1'\rangle\langle b^*\rangle &\le \langle sb^*|sa^m b^n\rangle \cdot |s_1' b^*\rangle \\
&\le \langle S\rangle\langle a^*\rangle|a\rangle^+\langle b^*|b^*\rangle \cdot |C_S\rangle \\
&\le N_S|C_S\rangle; \\
(\sum_{i<m, j\le n} \langle sa^i b^j|sa^m b^n\rangle \cdot |s_2' a^i b^j\rangle) &\le \langle S\rangle\langle a^*\rangle|a\rangle^+\langle b^*|b^*\rangle \cdot |C_S\rangle \\
&\le N_S|C_S\rangle; \\
(\sum_{i=m, j\le n} \langle sa^i b^j|sa^m b^n\rangle \cdot |s_2' a^i b^j\rangle) &\le \langle S\rangle\langle a^* b^*|b\rangle^+ \cdot |C_S\rangle \\
&\le N_S|C_S\rangle; \\
\langle sa^m b^n|sa^m b^n\rangle \cdot |s_2' a^m b^n\rangle &\le \langle C_S\rangle \cdot |s_2' a^m b^n\rangle; \\
\langle sa^m a^+ b^n b^+|sa^m b^n\rangle \cdot |s_1' a^m a^+ b^n b^+\rangle &\le \langle S\rangle\langle a^*\rangle|a\rangle^+\langle b^*|b^*\rangle \cdot |C_S\rangle \\
&\le N_S|C_S\rangle
\end{aligned}
$$

Thus, we obtain the inequality we desire:

$$
|sa^m b^n\rangle \cdot \langle s| \cdot [\iota(s)] \le \langle C_S\rangle \cdot |s_2' a^m b^n\rangle + N_S|C_S\rangle.
$$

The $\Longleftarrow$ direction can be shown by looking at the language model. If the inequality holds, then the language interpretation holds:

$$|sa^m b^n\rangle L(R_M) \subseteq L(\langle C_S\rangle) \cdot |s' a^{m'} b^{n'}\rangle + L(N_S) \cdot L(|C_S\rangle).$$

By Corollary 8, there exists a word $w \in L(R_M)$ s.t. its left component is $\langle sa^m b^n|$; we write the right component of w as $|w_r\rangle$. By Theorem 8, $w \leq R_M$; and by Inequalities (2.2) and (2.7),

$$\langle sa^m b^n|w_r\rangle \leq R_M \leq \langle C_S|C_S\rangle \Longrightarrow w_r \leq C_S.$$

Therefore, $w_r \in L(C_S)$ and by the definition of $L(N_S) \cdot L(|C_S\rangle)$,

$$|sa^m b^n\rangle \cdot \langle sa^m b^n|w_r\rangle = \langle sa^m b^n|sa^m b^n\rangle \cdot |w_r\rangle \notin L(N_S) \cdot L(|C_S\rangle).$$

Because of the language inclusion

$$|sa^m b^n\rangle L(R_M) \subseteq L(\langle C_S\rangle) \cdot |s' a^{m'} b^{n'}\rangle + L(N_S) \cdot L(|C_S\rangle),$$

we have

$$\langle sa^m b^n|sa^m b^n\rangle \cdot |w_r\rangle \in L(\langle C_S\rangle) \cdot |s' a^{m'} b^{n'}\rangle$$

therefore $w_r = s' a^{m'} b^{n'}$, and by Equivalence (2.5):

$$w = \langle sa^m b^n|s' a^{m'} b^{n'}\rangle \in R_M \Longrightarrow (s, m, n) \to (s', m', n').$$

□

We can further extend our previous result to establish a connection between provability and state reachability in certain type of machines. This result will be the core of our diagonal argument.

Theorem 11. *For a machine $M \triangleq (S, \hat{s}, \iota)$ and an input (m, n), if the set of reachable configurations from the input is finite, then a set $S' \subseteq S$ contains all the reachable states from input (m, n) if and only if the following inequality is provable:*

$$|\hat{s}a^n b^m\rangle R^*_M \leq \langle C^*_S\rangle|C_{S'}\rangle + \langle C^*_S\rangle N_S \langle C^*_S|C^*_S\rangle. \tag{2.8}$$

Proof. The $\Longrightarrow$ direction. We define the following term:

$$\mathsf{KA}(\Sigma) \ni C_r \triangleq \sum\{c_r \mid c_r \text{ is reachable}\}.$$

C_r is well-defined because there are only finitely many reachable configurations. By Theorem 10, assume $c_r \to c_r'$, we have the following inequality:

$$\begin{aligned} |c_r\rangle R_M &\leq \langle C_S|c_r'\rangle + N_S|C_S\rangle \\ &\leq \langle C_S|C_r\rangle + N_S|C_S\rangle \qquad\qquad c_r' \text{ is reachable} \end{aligned}$$

Since the above inequality is true for every reachable configuration c_r, then the following inequality is true:

$$|C_r\rangle R_M = \sum_{c_r} |c_r\rangle R_M \leq \langle C_S|C_r\rangle + N_S|C_S\rangle.$$

To show provability of Inequality (2.8), we will prove a stronger inequality:

$$|\hat{s}a^n b^m\rangle R_M^* \leq \langle C_S^*|C_r\rangle + \langle C_S^*\rangle N_S\langle C_S^*|C_S^*\rangle.$$

With Theorem 10 and inequality (2.7), we can derive the following two inequality:

$$\begin{aligned} \langle C_S^*|C_r\rangle R_M &\leq \langle C_S^*\rangle\langle C_S|C_r\rangle + \langle C_S^*\rangle N_S|C_S\rangle \\ &\leq \langle C_S^*|C_r\rangle + \langle C_S^*\rangle N_S\langle C_S^*|C_S^*\rangle \\ \langle C_S^*\rangle N_S\langle C_S^*|C_S^*\rangle R_M &\leq \langle C_S^*\rangle N_S\langle C_S^*|C_S^*\rangle\langle C_S|C_S\rangle \\ &\leq \langle C_S^*\rangle N_S\langle C_S^*|C_S^*\rangle \\ &\leq \langle C_S^*|C_r\rangle + \langle C_S^*\rangle N_S\langle C_S^*|C_S^*\rangle. \end{aligned}$$

Therefore,

$$\begin{aligned} &(\langle C_S^*|C_r\rangle + \langle C_S^*\rangle N_S\langle C_S^*|C_S^*\rangle)R_M \\ &\qquad \leq \langle C_S^*|C_r\rangle + \langle C_S^*\rangle N_S\langle C_S^*|C_S^*\rangle; \\ &|\hat{s}a^n b^m\rangle \leq |C_r\rangle \\ &\qquad \leq \langle C_S^*|C_r\rangle + \langle C_S^*\rangle N_S\langle C_S^*|C_S^*\rangle. \end{aligned}$$

By induction rule, we have the desired inequality:

$$\begin{aligned}|\hat{s}a^n b^m\rangle R_M^* &\leq \langle C_S^*\rangle|C_r\rangle + \langle C_S^*\rangle N_S\langle C_S^*|C_S^*\rangle \\ &\leq \langle C_S^*\rangle|C_{S'}\rangle + \langle C_S^*\rangle N_S\langle C_S^*|C_S^*\rangle.\end{aligned}$$

The $\Longleftarrow$ direction. We will first prove a small lemma: consider a configuration c_r that is reachable from input (m, n), we will show there exists a word $w \in L(\langle C_S^*\rangle)$ s.t.

$$w|c_r\rangle \in L(|\hat{s}a^n b^m\rangle)L(R_M)^*.$$

The theorem above is shown by induction on the number of steps to reach c_r:

- If c_r is reached in 0 steps, then

  $$c_r = (\hat{s}, m, n).$$

  In this case, $w = \epsilon$ and

  $$|c_r\rangle = |\hat{s}a^n b^m\rangle \in L(|\hat{s}a^n b^m\rangle)L(R_M)^*.$$

- If c_r is reached in $n+1$ steps, we find the configuration c_r' that is reached in n steps:

  $$(\hat{s}, m, n) \to^* c_r' \to c_r.$$

  By induction hypothesis, there is $w \in L(\langle C_S^*\rangle)$, s.t.

  $$w|c_r'\rangle \in L(|\hat{s}a^n b^m\rangle)L(R_M)^*.$$

  By Corollary 7,

  $$c_r' \to c_r \Longrightarrow \langle c_r'|c_r\rangle \in L(R_M).$$

  We have obtained our desired word:

  $$w|c_r'\rangle \cdot \langle c_r'|c_r\rangle = w\langle c_r'|c_r'\rangle|c_r\rangle \in L(|\hat{s}a^n b^m\rangle)L(R_M)^*.$$

Consider $w \in L(\langle C_S^*\rangle)$ s.t. $w|c_r\rangle \in L(|\hat{s}a^n b^m\rangle)L(R_M)^*$, by definition

$$w|c_r\rangle \notin L(\langle C_S^*\rangle N_S\langle C_S^*|C_S^*\rangle).$$

Because Inequality (2.8) holds, we have the following language inclusion:

$$L(|\hat{s}a^n b^m\rangle)L(R_M)^* \subseteq L(\langle C_S^*|)L(|C_{S'}\rangle) + L(\langle C_S^*|N_S\langle C_S^*|C_S^*\rangle).$$

Therefore,

$$w|c_r\rangle \in L(\langle C_S^*|) \cdot L(|C_{S'}\rangle).$$

Finally, by unfolding the definition of $L(\langle C_S^*|) \cdot L(|C_{S'}\rangle)$, we obtain $s' \in S$. □

Corollary 9. *For a machine $M \triangleq (S, \hat{s}, \iota)$ that always halt regardless of the input, it will have finitely many reachable states from any input. Therefore, for any input (m, n), S' contains all the reachable state from $(\hat{s}, m, n)$ if and only if inequality (2.8) is provable.*

We can then construct our diagonal argument. Assume that the equational theory of KA with atomic commutativity is decidable, there is a machine $P(S', M, M')$ that decides whether inequality (2.8) is provable when given machine $M \triangleq (S, \hat{s}, \iota)$ with the encoding of M' as input, and a subset of states $S' \subseteq S$.

Fix two distinct states s_1, s_2, we define the diagonal machine $D(M)$ as follows: let $M \triangleq (S, \hat{s}, \iota)$ and $S' \triangleq S \setminus \{s_1\}$,

- if $P(S', M, M)$ returns true, then we will go to state s_1 and returns true;
- if $P(S', M, M)$ returns false, then we will go to state s_2 and returns false.

Hence, state s_1 is reachable if and only if $P(S', M, M)$ returns true.

Then we employ the standard technique to feed the diagonal machine to itself. since we assumed equalities in KA with atomic commutativity is decidable, therefore D always terminates; hence we can enumerate all the possible the output of $D(D)$:

- If $D(D)$ returns true, then $P(S', D, D)$ returns true. By Corollary 9, $S' \triangleq S \setminus \{s_1\}$ contains all the reachable states of $D(D)$. However, by definition of D, s_1 is reachable when $P(S', D, D)$ is true, and $s_1 \notin S'$. Therefore, we obtain a contradiction.

- If $D(D)$ returns false, this means that $P(S', D, D)$ is false. By Corollary 9, $S' \triangleq S \setminus \{s_1\}$ do not contain all the reachable state. However, by definition of D, s_1 is not reachable in this case, and S' contains every state other than s_1. Hence, S' has to contain all the reachable state of $D(D)$. We got a contradiction again.

Therefore, our assumption that KA with atomic commutativity is decidable has to be false.

Corollary 10 (Incompleteness)**.** *There exists some commutable set X and two expression $e_1, e_2 \in \mathsf{KA}(X)$ s.t. $L(e_1) \subseteq L(e_2)$ but $e_1 \not\leq e_2$. In other words, there exists inequalities in the language interpretation that is not derivable using the theory.*

Proof. Assume that the language interpretation is complete, that is for all expression e_1, e_2

$$L(e_1) \subseteq L(e_2) \iff e_1 \leq e_2.$$

By definition of $\mathsf{KA}(X)$, $e_1 = e_2$ if and only if it can be proven using the theory of KA plus the commutativity in X, therefore deciding general equality is recursively-enumerable, by enumerating the proof.

However, since word inhabitance is decidable, language inclusion is co-recursively-enumerable, since we can simply check whether all words in $L(e_1)$ is in $L(e_2)$.

If the language inclusion is equivalent to inequalities in the theory, then the problem of inequalities in the theory is both recursively-enumerable and co-recursively-enumerable, hence decidable. This result contradicts our undecidability result for general inequality in KA with atomic commutativity hypotheses. Therefore, our assumption is false, and language interpretation is incomplete. □

2.4 Conclusion And Open Problem

In this paper we have shown that the word inhabitance problem in KA with commutativity hypotheses is decidable and complete, yet the general equalities are neither decidable nor complete. We believe this is the first known KA extension where the word inhabitance problem is decidable, yet the general equality is not.

Our method to show the decidability of word inhabitance problem involves using the matrix model to decompose a word into several components, which we believe is a novel technique in defining the empty word predicate and derivative in extensions of Kleene Algebra. This technique also yields straight-forward proof of soundness and the fundamental theorem.

However, there are still several important open problems: Several theorems leading to the undecidability result requires introspection on each case of the instructions, which leads to very long and tedious proof. We suspect some of these proofs, like the proof for Theorem 10, can be simplified by establishing more connection between the language interpretation and the free models. The exact complexity of KA with atomic commutativity is still unknown. In particular, we do not know whether the problem of deciding general equalities are RE-complete.

References

Debreuve, E., Barlaud, M., Aubert, G., Laurette, I., and Darcourt, J. (2001). Space-time segmentation using level set active contours applied to myocardial gated SPECT. *IEEE Trans. Med. Imag.*, 20(7):643–659.

Lamport, L. (1985). *LaTeX—A Document Preparation System—User's Guide and Reference Manual*. Addison-Wesley.

Chapter 3

Important Details

The use of Type 1 fonts and font embedding into the document are both dependent on a specific Latex installation and even on operating system. There is a good chance that it will work with no problem for you. However, should your thesis PDF be returned, please consider the following remedies discovered by students over many years.

3.1 Type 1 fonts

All Boston University thesis and dissertation submissions must use only Type 1 fonts to assure high-quality rendering. Type 3 fonts are not acceptable.

For some students adding the following two lines in "thesis.tex" preamble has worked:

```
\usepackage[T1]{fontenc}
\usepackagepslatex
```

The easiest way to check if fonts are embedded well and of what type, is to use Adobe Acrobat's Preflight – it shows exactly where the Type 3 fonts are in the thesis. You can learn more here: https://community.adobe.com/t5/acrobat/figure-out-where-a-specific-font-is-used-in-a-pdf/m-p/10880057?page=1#M238035

If you don't have Adobe Acrobat (BU students get it for free), you can quickly check which fonts have which type by looking into Files >> Properties >> Fonts, but

it doesn't tell where the text with a specific font type is.

Linux/Unix: If you are using LaTeX or Unix, the problem is that, by default, LaTeX uses Type 3 fonts. Since most users have a tendency to use the default settings, then Type 3 fonts will be used by default. You can try to change the first line in the preamble in "thesis.tex" to:

```
\documentstyle[12pt,times,letterpaper]{report}
```

since then Times fonts will be used (which are not Type 3). If there are mathematical formulas in the text, it is better to use:

```
\documentstyle[12pt,times,mathptm,letterpaper]{report}
```

3.2 Font embedding

All fonts must be embedded into the final PDF file. If they are not, sometimes equations may look strange or may not show up at all for several pages. This is often due to unembedded font problem. Should you have a font-embedding issue, this page may prove useful:

https://www.karlrupp.net/2016/01/embed-all-fonts-in-pdfs-latex-pdflatex

For those using Overleaf, this page might help: https://www.overleaf.com/learn/latex/Questions/My_submission_was_rejected_by_the_journal_because_%22Font_XYZ_is_not_embedded%22._What_can_I_do%3F

Chapter 4

Conclusions

4.1 Summary of the thesis

Time to get philosophical and wordy.

Important: In the list of references at the end of thesis, abbreviated journal and conference titles aren't allowed. Either you must put the full title in each item, or create a List of Abbreviations at the beginning of the references, with the abbreviations in one column on the left (arranged in alphabetical order), and the corresponding full title in a second column on the right. Some abbreviations, such as IEEE, SIGMOD, ACM, have become standardized and accepted by librarians, so those should not be spelled out in full.

Appendix A

Proof of xyz

This is the appendix.

References

Debreuve, E., Barlaud, M., Aubert, G., Laurette, I., and Darcourt, J. (2001). Space-time segmentation using level set active contours applied to myocardial gated SPECT. *IEEE Trans. Med. Imag.*, 20(7):643–659.

Lamport, L. (1985). *LaTeX—A Document Preparation System—User's Guide and Reference Manual*. Addison-Wesley.

CURRICULUM VITAE

Joe Graduate

Basically, this needs to be worked out by each individual, however the same format, margins, typeface, and type size must be used as in the rest of the dissertation.