Pushing the Limits of Kleene Algebra

Sergey Goncharov

FAU Erlangen-Nürnberg

RAMICS-AiML 2024, 19-22 August 2024

Overview

What I will talk about:

- Compositionality
- Modularity
- Genericity
- **Design**
- \blacksquare Semantics

What I wont talk about:

- **Efficiency**
- Optimization ш.
- Computation Complexity

|≡ Goncharov, ["Shades of Iteration: From Elgot to Kleene",](#page-0-0) 2023, WADT 2022

 \equiv Goncharov and Uustalu, ["A Unifying Categorical View of](#page-0-0) [Nondeterministic Iteration and Tests",](#page-0-0) 2024, CONCUR 2024

Historical Overture

Regular Events

- **EXPLORED STAT** $e \mapsto e^*$
- \blacksquare Kleene theorem
	- Syntax for finite state machines
	- \blacktriangleright Algebraic equational reasoning

 $(b + a(ab * a)b)*(1 + aa)$

Regular expressions over Σ:

$$
e, e_1, e_2 \coloneqq (a \in \Sigma) | 0 | 1 | e_1 + e_2 | e_1 e_2 | e^*
$$

Language interpretation: п.

> $\llbracket 0 \rrbracket = \{\}\$ $\llbracket e_1 e_2 \rrbracket = \{ xy \mid x \in \llbracket e_1 \rrbracket, y \in \llbracket e_2 \rrbracket \}$ $\llbracket 1 \rrbracket = \{ \epsilon \}$ $\llbracket e_1 + e_2 \rrbracket = \llbracket e_1 \rrbracket \cup \llbracket e_2 \rrbracket$ $\llbracket e^* \rrbracket = \{\epsilon\} \cup \llbracket e \rrbracket \cup \llbracket ee \rrbracket \cup \ldots$

Language $L \subseteq \Sigma^*$ is regular iff $L = [\![e]\!]$ for some regular expression e
with $\lceil a \rceil = a$ for $a \in \Sigma$ with $\llbracket a \rrbracket = a$ for $a \in \Sigma$

å Other interpretations

▶ Yes, e.g. relational one!

 \odot Complete reasoning system for regular expressions

$$
(ab)^{\ast}=\mathbf{1}+a(ba)^{\ast}b
$$

$$
(ab)^{\ast}=\mathbf{1}+a(ba)^{\ast}b
$$

$$
1 + a(ba)^*b = 1 + a(1 + (ba)(ba)^*)b
$$

$$
(ab)^{\ast}=\mathbf{1}+a(ba)^{\ast}b
$$

$$
1 + a(ba)*b = 1 + a(1 + (ba)(ba)*b)
$$

= 1 + a1b + a(ba)(ba)*b

$$
(ab)^{\ast}=\mathbf{1}+a(ba)^{\ast}b
$$

$$
1 + a(ba)*b = 1 + a(1 + (ba)(ba)*b = 1 + a1b + a(ba)(ba)*b = 1 + ab + (ab)a(ba)*b
$$

$$
(ab)^{\ast}=\mathbf{1}+a(ba)^{\ast}b
$$

$$
1 + a(ba)*b = 1 + a(1 + (ba)(ba)*b)= 1 + a1b + a(ba)(ba)*b= 1 + ab + (ab)a(ba)*b= 1 + (ab)(a(ba)*b)
$$

```
(ab)^* = 1 + a(ba)^*b
```

$$
1 + a(ba)*b = 1 + a(1 + (ba)(ba)*b)
$$

= 1 + a1b + a(ba)(ba)*b
= 1 + ab + (ab)a(ba)*b
= 1 + (ab)(1 + a(ba)*b)

```
(ab)^* = 1 + a(ba)^*b
```
is true, because 1 $+$ $a(ba)^*b$ is a fixpoint of the map that defines $(ab)^*$

$$
1 + a(ba)*b = 1 + a(1 + (ba)(ba)*b)
$$

= 1 + a1b + a(ba)(ba)*b
= 1 + ab + (ab)a(ba)*b
= 1 + (ab)(1 + a(ba)*b)

This only works because $x \mapsto 1 + abx$ is guarded

 \blacksquare *x* \mapsto 1 + (*a* + 1)*x* is un-guarded and has infinitely many fixpoints

Salomaa's Complete Axiomatization

 \blacktriangleright plus rule:

$$
\frac{v = e + uv \qquad u \text{ guarded}}{v = u^*e}
$$

standalone/salomaa-photo.

No Finite Equational Axiomatization

Redko˚ noticed that

All identities (power identities)

$$
e^* = (e^k)^* (1 + e + \ldots + e^{k-1})
$$

are sound

- Any finite set of sound equations entails only finitely many of them
- \blacksquare Hence, no finite axiomatizability (even on one-letter alphabet)

So,

- Ω How to choose infinite set of non-obvious axioms of iteration?
- $(?)$ How would we know that this choice is correct?

[˚]V. N. Redko, On defining relations for the algebra of regular events, 1964

Conway's Monograph

Conway˚ came up with various insights:

 \blacksquare Power identities do not suffice. e.g. they do not imply

$$
(e+u)^* = ((e+u)(u + (eu^*)^{n-2}e))^*
$$

$$
(1 + (e+u)\sum_{i=0}^{n-2} (eu^*)^i)
$$

- Made several conjectures on potential complete axiomatization
- Observed that algebraic laws of regular expressions transfer to matrices of regular expressions

 \hat{Y} \Rightarrow Bridge between algebra and automata (represented by matrices)

[˚]J. H. Conway, Regular Algebra and Finite Machines, 1971

Matrices of Regular Expressions

 $(n \times n)$ -matrices of regular expressions support same operations. For $n = 2$:

"
$$
1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} a+a' & b+b' \\ c+c' & d+d' \end{bmatrix}
$$

\n"o" is
$$
0 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{bmatrix}
$$

Idea for $A^*: I + A + A^2 + ...$ $\widehat{\mathsf{Y}}$ Key insight: there is closed form for A* as matrix of regular expressions **Intuition**: in $\begin{bmatrix} e_{11} & e_{12} \ e_{21} & e_{22} \end{bmatrix} = A^*$, e_{ij} represents language of 2-state

automaton where *i* – initial, *j* – final

Automata are triples

$$
A \in \{0, 1\}^n
$$
, $B \in \mathcal{E}^{n \times n}$, $C \in \{0, 1\}^n$

 ϵ - certain class of regular expression

start	a,b					
expressions	\n $\left[\begin{array}{c} 1 \\ 0 \end{array}\right]^T$ \n	\n $\left[\begin{array}{c} a \\ 0 \end{array}\right]^T$ \n	\n $\left[\begin{array}{c} 1 \\ 0 \end{array}\right]^T$ \n	\n $\left[\begin{array}{c} 0 \\ a \end{array}\right]^T$ \n	\n $\left[\begin{array}{c} 1 \\ 0 \end{array}\right]^T$ \n	\n $\left[\begin{array}{c} (a+b)b^*a)^* \\ (b^*a(a+b))^*a & b^*(a(a+b))^* \end{array}\right]^T$ \n
\n $\left[\begin{array}{c} (a+b)b^*a)^* \\ (a+b)b^*a)^* \end{array}\right]$ \n						
\n $\left[\begin{array}{c} (a+b)b^*a)^* \\ (a+b)b^*a)^* \end{array}\right]$ \n						

Kleene Algebra

5. show that all ensuing transitions $A^\top B^*C = \hat{A}^\top \hat{B}^* \hat{C}$ are provable

■ Corollary: Language interpretation = free Kleene algebra

Not tailored to language model – complete also over relational model \blacksquare

Algebraic, i.e. closed under substitution, in contrast to Salomaa's rule

$$
\frac{w = u + ew \quad e \text{ guarded}}{w = e^*u}
$$

All fixpoints are least (pre-)fixpoints

▶ in Salomaa's system: particular fixpoints are unique fixpoints

 \blacksquare Induction rules

$$
\frac{u + ew \leq w}{e^*u \leq w} \qquad \qquad \frac{u + we \leq w}{ue^* \leq w}
$$

encompass infinitely many identities, critical for completeness

Tests for Control

- Another reading: Algebra elements $=$ programs
	- \triangleright 0 divergence and/or deadlock, 1 neutral program, etc.
- Kleene algebra with tests (KAT) adds control via tests:
	- ▶ Kleene sub-algebra *B*
	- \blacktriangleright *B* is Boolean algebra under $(0, 1, \ldots)$
- This enables encodings:
	- ▶ Branching (if *b* then *p* else *q*) as b ; $p + \overline{b}$; *q* \triangleright Looping (while *b* do *p*) as $(b;p)^{*}; \overline{b}$ ▶ Hoare triples $\{a\} p \{b\}$ as $a; p; b = a; p$
	-

Example:

while *b* do $p =$ if *b* then *p* else (while *b* do *p*)

Kleene Algebra Today

- Regular expressions
- Algebraic language of finite state machines and beyond
- Relational semantics of programs
- Relational reasoning and verification, e.g. via dynamic logic
- Plenty of extensions: ш.
	- \triangleright modal \Rightarrow modal Kleene algebra (Struth et al.)
	- \triangleright stateful \Rightarrow KAT + B! (Grathwohl, Kozen, Mamouras)
	- \triangleright concurrent \Rightarrow concurrent Kleene algebra (Hoare et al.)
	- \triangleright nominal \Rightarrow nominal Kleene algebra (Kozen et al.)
	- \triangleright differential equations \Rightarrow differential dynamic logic (Platzer et al.)
	- \blacktriangleright network primitives \Rightarrow NetKAT (Foster et al.)
	- \blacktriangleright etc., etc., etc.

 \blacksquare decidability and completeness (most famously w.r.t. language interpretation and relational interpretation)

Pushing Limits

Assumming programs raise <mark>exceptions:</mark> raise e_i $=$ "raise exception e_i ",

 r aise e_1 = raise e_1 ; o = o = raise e_2 ; o = raise e_2

- So, we cannot have more than one exception
	- \blacktriangleright ... unless we discard the law

$$
p; o = o
$$

Scenario II: Branching Time

are famously non-bisimular, failing Kleene algebra law

p; $(q + r) = p$; $q + p$; *r*

Scenario III: Divergence

\blacksquare Identity

$$
(p+1)^* = p^*
$$

is provable in Kleene algebra, because p^* is a least fixpoint

Alternatively:

$$
1^* = 1
$$

 \blacksquare Hence deadlock = divergence

What is generic core of Kleene iteration?

- Core reasoning principles
- Robustness under adding features (e.g. exceptions) п.
- Generic completeness argument
- Compatibility with classical program semantics
	- \Rightarrow Soundness of while-loop encoding

What is generic core of Kleene iteration?

- Core reasoning principles
- Robustness under adding features (e.g. exceptions)
- Generic completeness argument ш.
- Compatibility with classical program semantics
	- \Rightarrow Soundness of while-loop encoding

Categorifying Iteration

From Algebras to Categories

Categories \approx many-sorted monoids: ш.

$$
1_A: A \to A \quad \text{(unit)} \qquad \qquad \frac{p: A \to B \qquad q: B \to C}{p; q: A \to C} \qquad \text{(multiplication)}
$$

 \triangleright Objects *A*, *B*, ... – sorts, Morphisms *p*: *A* \rightarrow *B* – programs ▶ **Fact:** monoid = single-object category

Kleene-Kozen categories – additionaly

$$
O_{A,B}: A \to B \qquad \frac{p:A \to B}{p+q:A \to B} \qquad \frac{p:A \to A}{p^*: A \to A}
$$

subject to Kleene algebra laws

- ▶ **Fact:** Kleene algebra = single-object Kleene-Kozen category
- ▶ **Example:** Category of relations = relational interpretation
- \blacksquare Tests = particular morphisms $b: A \rightarrow A$

Coproducts and Elgot Iteration

 \heartsuit Quasi-equational axiomatizations little

[˚]S. Bloom, **Z. Ésik**, Iteration Theories, 1993

Coproducts and Elgot Iteration

Quasi-equational axiomatizations little explored

[˚]S. Bloom, **Z. Ésik**, Iteration Theories, 1993

- Given Elgot iteration operator, fix carrier of exceptions *E*
- Exception-raising morphisms $A \rightarrow B \oplus E$ themeselves form a category
- Elgot iteration and its laws carry over m.
	- ▶ This fails for Kleene-Kozen categories
- Elgot iteration's laws are thus stable under exception monad transformer
- Similarly: state, reading, writing, adjoining process algebra actions

 \odot Can we formulate uniform Conway iteration via familiar while-loops

■ Call morphisms of the form $d: A \rightarrow A \oplus A$ decisions

- \blacktriangleright In particular: ff left injection, tt right injection
- We then can express if-then-else:

$$
\frac{d: A \to A \oplus A \qquad p: A \to B \qquad q: A \to B}{\text{if } d \text{ then } p \text{ else } q: A \to B}
$$

▶ In particular: $\text{-} d$ = if *d* then ff else tt, $(d || e)$ = if *d* then tt else *e* ■ Various expected laws are entailed, but some are not, e.g.

$$
d \mid \mid \mathsf{tt} \neq \mathsf{tt}
$$

Uniform Conway While-Operator

Theorem˚**:** if the class of decisions is large enough, uniform Conway iteration is equivalent to while-loops

Axioms:

while
$$
d \, d \, o \, p = \text{if } d \, \text{then } p
$$
; (while $d \, d \, o \, p$) else 1

\nwhile $(d \, || \, e) \, d \, o \, p = (\text{while } d \, d \, o \, p)$; while $e \, d \, o \, (p; \text{while } d \, d \, o \, p)$

\nwhile $(d \, \text{of } e \, || \, \text{tt})$) $d \, o \, p = \text{while } d \, d \, o$ (if $e \, \text{then } p \, \text{else } p$)

Uniformity Rule:

 u ; if *d* then *p*; tt else ff = if *e* then *q*; *u*; tt else *v*; ff *u*; while *d* do $p = ($ while *e* do *q* $)$; *v*

where *u*, *v* come from a selected class of programs

[˚]S. Goncharov, Shades of Iteration: From Elgot to Kleene, 2023

Tests and Decisions

In presence of non-determinism, decisisons $d: A \rightarrow A \oplus A$ decompose:

$$
d = b; \text{tt} + \bar{b}; \text{ff} \qquad (b, \bar{b}: A \to A)
$$

■ Test-based 'if' and 'while':

Axioms:

while *b* do $p =$ if *b* then p ; (while *b* do p) else 1

while $(b \vee c)$ do $p = ($ while *b* do *p* $)$; while *c* do $(p;$ while *b* do *p* $)$

Uniformity:

$$
u; b; p = c; q; u \qquad u; \bar{b} = \bar{c}; v
$$

u; while b do p = (while c do q); v

Alternative axiomatization: idempotent semiring, and

$$
p^* = 1 + p; p^* \qquad (p + q)^* = p^*; (q; p^*)^*
$$

$$
1^* = 1 \qquad \frac{u; p = q; u}{u; p^* = q^*; u}
$$

- This is true for Kleene-Kozen categories \Rightarrow Kleene algebra
- Removing $1^* = 1$ yields may-diverge Kleene algebras, $(-)^*$ is no longer least fixpoint
- Uniformity is postulated for all *u*

Restricting Uniformity

$$
u; p = q; u
$$

u; $p^* = q^*; u$

raise *e* = raise *e*; 1 = 1;raise *e* = raise *e*

$$
raise e] = raise e; 1^* = | 1^*; raise e
$$

raise	e = raise	e ; 1 = 1; raise	e = raise	e
raise	e	= raise	e ; 1* = $\boxed{1^*$; raise	e

 \bigoplus Need not hold in may-diverge Kleene algebras

Restricting Uniformity

$$
u; p = q; u
$$

u; $p^* = q^*; u$

 \bigoplus Need not hold in may-diverge Kleene algebras \Rightarrow Restrict to linear *u*:

$$
u; o = 0
$$
 $u; (p+q) = u; p + u; q$

KiCT

Kleene-iteration category with tests (KiCT)

- Category with coproducts and nondeterminism
- Selected class of tests
- Selected class of linear tame morphisms
- Kleene iteration
- **Laws:**

$$
o; p = o \t (p + q); r = p; r + q; r
$$

$$
p^* = 1 + p; p^* \t (p + q)^* = p^*; (q; p^*)^*
$$

$$
\frac{u; p^* = q^*; u}{u; p = q; u}
$$

with tame *u*

- $KiCT + (1^* = 1)$ with all morphisms tame = Kleene-Kozen with tests and coproducts
- \blacksquare KiCT with expressive tests = tame-uniform Conway iteration + non-determinism
- Free KiCT = non-deterministic rational trees w.r.t. may-diverge nondeterminism

What is generic core of Kleene iteration?

KiCT:

- \odot Core reasoning principles
- |⊙ Robustness under adding features
- Generic completeness argument \odot
- Compatibility with classical program semantics ∞

 30

What is generic core of Kleene iteration?

KiCT:

- \odot Core reasoning principles
- の Robustness under adding features
- Generic completeness argument \odot
- Compatibility with classical program semantics \odot

But what is KiCT without coproducts?

 30

Hypothetical Route

- If everything is tame (Kleene algebra), this is essentially what happens
- What if nothing is tame (Process algebra)?

Milner's Conundrum

- Milner^{*} realized that "regular behaviours" are properly more general than "*-behaviours"
- Simplest example

$$
\begin{cases}\nX = 1 + a; Y \\
Y = 1 + b; X\n\end{cases}
$$

We can pass to $X = 1 + a$; $(1 + b; X)$, but not to *X* = $(ab)^*(1 + a)$

- \blacksquare This descrepancy \approx failure of matrix construction/Kleene theorem
- **Milner's solution is equivalent to using coproducts in the language**
- He also proposed a modification of Salomaa's system for *-behaviours – proven complete only recently (Grabmayer)

[˚]R. Milner, A complete inference system for a class of regular behaviours, 1984

- KiCTs reframe Kleene algebra principles in categorical setting and succeed with various yardsticks
- KiCTs without coproducts would be a hypothetical most basic notions of Kleene iteration
- \blacksquare **Open Problem:** Can it ever be found?