Pushing the Limits of Kleene Algebra

Sergey Goncharov

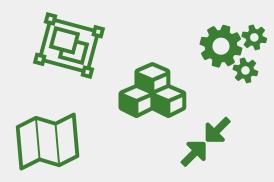
FAU Erlangen-Nürnberg

RAMICS-AiML 2024, 19-22 August 2024

Overview

What I will talk about:

- Compositionality
- Modularity
- Genericity
- Design
- Semantics



What I wont talk about:

- Efficiency
- Optimization
- Computation Complexity



|≣]

Goncharov, "Shades of Iteration: From Elgot to Kleene", 2023, WADT 2022

Goncharov and Uustalu, "A Unifying Categorical View of |≣] Nondeterministic Iteration and Tests", 2024, CONCUR 2024

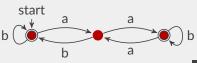
Historical Overture

Regular Events





- Kleene star $e \mapsto e^*$
- Kleene theorem
 - Syntax for finite state machines
 - Algebraic equational reasoning



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

Regular expressions over Σ :

$$e, e_1, e_2 \coloneqq (a \in \Sigma) \mid O \mid 1 \mid e_1 + e_2 \mid e_1e_2 \mid e^*$$

Language interpretation:

$$\begin{bmatrix} \mathbf{0} \end{bmatrix} = \{ \} \qquad \begin{bmatrix} \mathbf{e}_1 \mathbf{e}_2 \end{bmatrix} = \{ xy \mid x \in \llbracket \mathbf{e}_1 \rrbracket, y \in \llbracket \mathbf{e}_2 \rrbracket \}$$
$$\begin{bmatrix} \mathbf{1} \end{bmatrix} = \{ \mathbf{\varepsilon} \} \qquad \begin{bmatrix} \mathbf{e}_1 + \mathbf{e}_2 \rrbracket = \llbracket \mathbf{e}_1 \rrbracket \cup \llbracket \mathbf{e}_2 \rrbracket$$
$$\begin{bmatrix} \mathbf{e}^* \rrbracket = \{ \mathbf{\varepsilon} \} \cup \llbracket \mathbf{e} \rrbracket \cup \llbracket \mathbf{e} \rrbracket \cup \ldots$$

Language L ⊆ Σ* is regular iff L = [[e]] for some regular expression e with [[a]] = a for a ∈ Σ

Other interpretations

► Yes, e.g. relational one!

Complete reasoning system for regular expressions

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

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

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

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

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

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

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

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

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

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

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

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

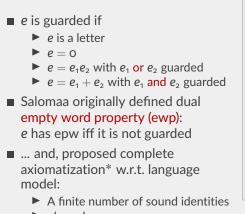
$$1 + a(ba)*b = 1 + a(1 + (ba)(ba)*)b$$

= 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

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

Salomaa's Complete Axiomatization



standalone/salomaa-photo.

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

^{*}A. Salomaa, Two Complete Axiom Systems for the Algebra of Regular Events, 1966

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
- Hence, no finite axiomatizability (even on one-letter alphabet)



So,

- 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:

 Power identities do not suffice, e.g. they do not imply

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

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



 \mathcal{G} \Rightarrow Bridge between algebra and automata (represented by matrices)

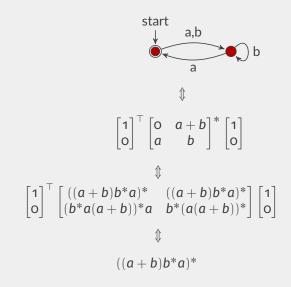
*J. H. Conway, Regular Algebra and Finite Machines, 1971

Matrices of Regular Expressions

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

"1" is
$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$
 $\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} a+a' & b+b' \\ c+c' & d+d' \end{bmatrix}$
"o" is $O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} aa'+bc' & ab'+bd' \\ ca'+dc' & cb'+dd' \end{bmatrix}$

■ 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

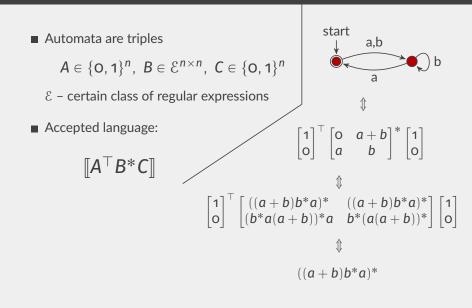


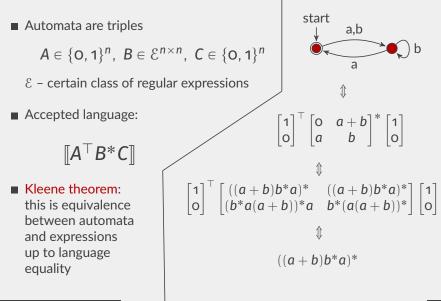
Automata are triples

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

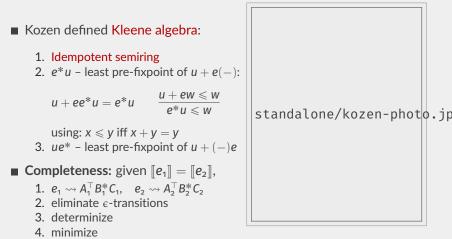
1 0

 \mathcal{E} – certain class of regular expressions





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

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

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

■ All fixpoints are least (pre-)fixpoints

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

Induction rules

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

encompass infinitely many identities, critical for completeness

Tests for Control

- Another reading: Algebra elements = programs
 - ▶ 0 divergence and/or deadlock, 1 neutral program, etc.
- Kleene algebra with tests (KAT) adds control via tests:
 - Kleene sub-algebra B
 - ► B is Boolean algebra under (0, 1, ; , +)
- This enables encodings:
 - Branching (if b then p else q) as $b; p + \overline{b}; q$ 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:
 - ▶ modal ⇒ modal Kleene algebra (Struth et al.)
 - ► stateful ⇒ KAT + B! (Grathwohl, Kozen, Mamouras)
 - ► concurrent ⇒ concurrent Kleene algebra (Hoare et al.)
 - ▶ nominal ⇒ nominal Kleene algebra (Kozen et al.)
 - ► differential equations ⇒ differential dynamic logic (Platzer et al.)
 - network primitives > NetKAT (Foster et al.)
 - etc., etc., etc.

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

Pushing Limits

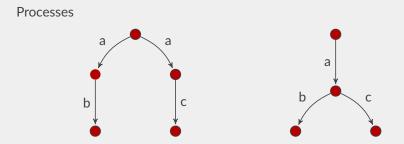
Assumming programs raise exceptions: raise $e_i =$ "raise exception e_i ",

raise e_1 = raise e_1 ; o = o = raise e_2 ; o = raise e_2

- So, we cannot have more than one exception
 - ... 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

Identity

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

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

Alternatively:

$$1^* = 1$$

Hence deadlock = divergence



What is generic core of Kleene iteration?

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

Categorifying Iteration

From Algebras to Categories

■ Categories ≈ many-sorted monoids:

$$n_A: A \to A$$
 (unit) $\frac{p: A \to B \quad q: B \to C}{p; q: A \to C}$ (multiplication)

Objects A, B, ... - sorts, Morphisms p: A → B - programs
 Fact: monoid = single-object category

Kleene-Kozen categories – additionaly

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

subject to Kleene algebra laws

- Fact: Kleene algebra = single-object Kleene-Kozen category
- Example: Category of relations = relational interpretation

• Tests = particular morphisms $b: A \rightarrow A$

Coproducts and Elgot Iteration

- Coproducts A ⊕ B can be thought of as disjoint unions A ⊕ B
- Elgot iteration:

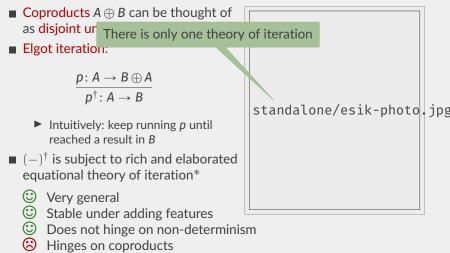
$$\frac{p\colon A\to B\oplus A}{p^{\dagger}\colon A\to B}$$

- Intuitively: keep running p until reached a result in B
- (-)[†] is subject to rich and elaborated equational theory of iteration*
 - Very general
 - Stable under adding features
 - Does not hinge on non-determinism
 - 😕 Hinges on coproducts
 - Quasi-equational axiomatizations little explored

standalone/esik-photo	.jpg
-----------------------	------

^{*}S. Bloom, Z. Ésik, Iteration Theories, 1993

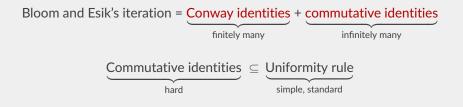
Coproducts and Elgot Iteration



😕 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
 - 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



Can we formulate uniform Conway iteration via familiar while-loops

Control in Category

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

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

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

In particular: ~ d = <u>if</u> d <u>then</u> ff <u>else</u> tt, (d || e) = <u>if</u> d <u>then</u> tt <u>else</u> e
 Various expected laws are entailed, but some are not, e.g.

$d \parallel tt \neq tt$

Uniform Conway While-Operator

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

Axioms:

while
$$d \text{ do } p = \text{if } d \text{ then } p$$
; (while $d \text{ do } p$) else 1
while $(d \parallel e) \text{ do } p = (\text{while } d \text{ do } p)$; while $e \text{ do } (p; \text{while } d \text{ do } p)$
while $(d \&\& (e \parallel tt)) \text{ do } p = \text{while } d \text{ do } (\text{if } e \text{ then } p \text{ else } p)$

Uniformity Rule:

 $\frac{u; \underline{if} \ d \underline{then} \ p; tt \underline{else} \ ff = \underline{if} \ e \underline{then} \ q; u; tt \underline{else} \ v; ff}{u; \underline{while} \ d \ \underline{do} \ p = (\underline{while} \ e \ \underline{do} \ q); v}$

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:

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

Test-based 'if' and 'while':

Axioms:

while $b \operatorname{do} p = \operatorname{if} b \operatorname{then} p$; (while $b \operatorname{do} p$) else 1

while $(b \lor c)$ do p = (while b do p); while c do (p; while b do p)

Uniformity:

 $\frac{u; b; p = c; q; u}{u; \text{ while } b \text{ do } p = (\text{while } c \text{ do } q); v}$

Alternative axiomatization: idempotent semiring, and

$$p^* = 1 + p; p^*$$
 $(p + q)^* = p^*; (q; p^*)^*$
 $1^* = 1$ $\frac{u; p = q; u}{u; p^* = q^*; u}$

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

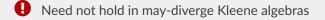
Restricting Uniformity

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

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



Restricting Uniformity

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

● Need not hold in may-diverge Kleene algebras ⇒ Restrict to linear u:

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

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;
$$p = O$$
 $(p+q)$; $r = p$; $r + q$; r
 $p^* = 1 + p$; p^* $(p+q)^* = p^*$; $(q; p^*)^*$
 $\frac{u; p^* = q^*; u}{u; p = q; u}$

with tame u

- KiCT + (1* = 1) with all morphisms tame = Kleene-Kozen with tests and coproducts
- 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:

- Core reasoning principles
- Robustness under adding features
- Generic completeness argument
- Ompatibility with classical program semantics

What is generic core of Kleene iteration?

KiCT:

- Ore reasoning principles
- Robustness under adding features
- Ompatibility with classical program semantics

But what is KiCT without coproducts?

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

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

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

- This descrepancy ≈ 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
- Open Problem: Can it ever be found?