

Building Verification Components from Algebraic Principles

Lectures at Universitetet i Oslo 2019

Georg Struth

University of Sheffield

These Lectures

- building components for program correctness in **Isabelle/HOL**
 - program construction (by transformation/refinement)
 - program verification
- simple **principled approach** that separates control/data flow
 - abstract algebra for control
 - concrete semantics for data domains
- in detail: simple construction/verification components for
 - while programs
 - based on Hoare logic
 - Morgan's refinement calculus
 - predicate transformers

all components correct by construction

Principled Approach

algebra	intermediate semantics	concrete semantics
control flow	abstract data flow	concrete data flow
control flow logic	intermediate logic	verification tool

First Instance

KAT

control flow

propositional
Hoare logic

relational KAT

abstract data flow

relational
verification conditions

relational KAT
over store

concrete data flow

verification conditions
based on Hoare logic

Second Instance

KAD	relational KAD	relational KAD over store
control flow	abstract data flow	concrete data flow
transformers (PDL)	relational verification conditions	verification conditions based on transformers

Plan

lectures

1. algebraic foundations (Kleene algebras)
2. mathematical components for these algebras
3. construction/verification components for sequential programs
4. extension to hybrid programs (ongoing research)

exercises

depending on interest we could look at

- algebraic reasoning about programs
- verification examples

Algebraic Foundations

While-Programs

syntax (regular operations)

- + nondeterministic choice
- sequential composition
- * finite iteration
- 0 failure/abort
- 1 skip

abstract semantics

regular expressions $t ::= 0 \mid 1 \mid a \in \Sigma \mid t + t \mid t \cdot t \mid t^*$

Kleene algebra $(K, +, \cdot, 0, 1, *)$

Kleene algebra is algebra of regular expressions

Dioids

definition

a **dioid** (idempotent semiring) is a structure $(S, +, \cdot, 0, 1)$ where

- ▷ $(S, +, 0)$ is a semilattice with least element 0
- ▷ $(S, \cdot, 1)$ is a monoid
- ▷ multiplication distributes over addition
- ▷ zero is left/right annihilator

$$\begin{aligned}x + (y + z) &= (x + y) + z & x + y &= y + x & x + 0 &= x & x + x &= x \\x(yz) &= (xy)z & x1 &= x & 1x &= x \\x(y + z) &= xy + xz & (x + y)z &= xz + yz \\x0 &= 0 & 0x &= 0\end{aligned}$$

Dioids

natural order

- $(S, +)$ is semilattice with partial order $x \leq y \Leftrightarrow x + y = y$
- regular operations preserve order (e.g. $x \leq y \Rightarrow z + x \leq z + y$)
- 0 is least element

opposition

- map $(-)^{\partial} : S \rightarrow S$ swaps order of multiplication

$$0^{\partial} = 0 \quad 1^{\partial} = 1 \quad (x + y)^{\partial} = x^{\partial} + y^{\partial} \quad (x \cdot y)^{\partial} = y^{\partial} \cdot x^{\partial}$$

- $\partial[S]$ is again a dioid—the **opposite dioid**

Kleene Algebras

definition

a **Kleene algebra** is a dioid expanded by star operation that satisfies

$$1 + xx^* \leq x^*$$

$$z + xy \leq y \Rightarrow x^*z \leq y$$

$$1 + x^*x \leq x^*$$

$$z + yx \leq y \Rightarrow zx^* \leq y$$

intuition

- x^*z is least solution of affine linear inequality $z + xy \leq y$
- zx^* is least solution of affine linear inequality $z + yx \leq y$

Models of Kleene Algebra

for programming

- **binary relations** form KAs
- our verification components are based on this model

for proofs

- **(regular) languages** form KAs
- regular expressions are ground terms in KA signature
- KAs are complete for regular expression equivalence
- variety of KA is decidable via automata (PSPACE-complete)

Language Kleene Algebras

let Σ^* denote free monoid with empty word ε over Σ

definition

a **language** is a subset of Σ^*

theorem (soundness)

- $(2^{\Sigma^*}, \cup, \cdot, *, \emptyset, \{\varepsilon\})$ forms the **full language KA** over Σ , where

$$X \cdot Y = \{vw \mid v \in X \wedge w \in Y\}$$

$$X^* = \bigcup_{i \geq 0} X^i$$

and $X^0 = \{\varepsilon\}$, $X^{i+1} = XX^i$

- any subalgebra forms a **language KA**

Regular Language Kleene Algebras

definition

KA morphism $L : T_{KA}(\Sigma) \rightarrow 2^{\Sigma^*}$ generates **regular languages** over Σ :

$$\begin{aligned} L0 &= \emptyset & L1 &= \{\varepsilon\} & La &= \{a\} \text{ for } a \in \Sigma \\ L(s+t) &= Ls \cup Lt & L(s \cdot t) &= Ls \cdot Lt & L(t^*) &= (Lt)^* \end{aligned}$$

theorem (soundness)

- regular languages over Σ form KA
- in particular $KA \vdash s = t \Rightarrow Ls = Lt$ for all $s, t \in T_{KA}(\Sigma)$

Completeness of Kleene Algebra

theorem [Kozen]

$KA \vdash s = t \Leftrightarrow Ls = Lt$ for all $s, t \in T_{KA}(\Sigma)$

consequences

- regular languages over Σ are generated freely by Σ in variety of KA
- KA axiomatises equational theory of regular expressions (as induced by regular language identity)
- equational theory of KA decidable (by automata)

Relation Kleene Algebra

binary relation

subset of $A \times A$

$$R = \{(a, b) \mid a, b \in A\}$$

theorem (soundness)

- $(2^{A \times A}, \cup, \cdot, \emptyset, id, *)$ forms **full relation Kleene algebra** over A , where

$$id = \{(a, a) \mid a \in A\}$$

$$R \cdot S = \{(a, b) \mid \exists c. (a, c) \in R \wedge (c, b) \in S\}$$

$$R^* = \bigcup_{i \geq 0} R^i \quad (\text{reflexive transitive closure of } R)$$

- every subalgebra forms a **relation Kleene algebra**

Relation Kleene Algebra

theorem (completeness)

if $s = t$ holds in class of all relation KAs, then $KA \vdash s = t$

consequence

- equational theory of relation KA is decidable via automata
- this makes KA interesting for program construction/verification

Beyond Equations

quasivariety of KA

undecidable (uniform word problem for semigroups)

quasivariety of regular expressions

KA does not work

- $x^2 = 1 \Rightarrow x = 1$ holds in language KA
- but not for relation $R = \{(0, 1), (1, 0)\}$, which is in KA (with $\{(0, 0), (1, 1)\}, \emptyset$, etc.)

program construction/verification requires
reasoning under assumptions

Kleene Algebras and Sequential Programs

program analysis

- reason about actions and propositions/states
- propositions can be **tests** or **assertions**

relational semantics

- relations model i/o-behaviour of programs on state spaces
- elements $p \leq 1$ represent sets of states/propositions
 - px yields all x -transitions that start from states in p
 - xp , by opposition, yields all x -transitions that end in states in p
- these elements form boolean subalgebras
(join is $+$, meet is \cdot , 0 is least and 1 greatest element)
- they can be used as **tests** or **assertions** in relational semantics

Kleene Algebras with Tests

abstraction

use KA for actions and BA (test algebra) for propositions

definition [Manes/Kozen]

two-sorted structure $(K, B, +, \cdot, \neg, 0, 1, *)$

- BA $(B, +, \cdot, \neg, 0, 1)$ embedded into K
- K models actions, B tests/assertions
- partial operation \neg defined on subalgebra B

Models of KAT

relation KAT

- binary relations form KATs
 - test algebra formed by subsets of id
 - these subidentities are isomorphic to sets of states
- every relation KAT is isomorphic to relation KA
- hence equational theory of relation KAT is still PSPACE-complete

guarded string KAT

- essentially trace KAT in which propositions and actions alternate
- P formed by atoms of free BA generated by finite set G
- guarded strings (and traces) form words over enlarged alphabet
- this implies completeness of KAT for guarded regular languages

KAT and Imperative Programs

algebraic program semantics

while programs (**no assignment**):

abort = 0

skip = 1

$x; y = xy$

if p **then** x **else** y **fi** = $px + \neg py$

while p **do** x **od** = $(px)^* \neg p$

Kleene Algebra with Isabelle

—demo—

Verification Component based on KAT

Outline

KAT

control flow

propositional
Hoare logic

relational KAT

abstract data flow

relational
verification conditions

relational KAT
with store

concrete data flow

verification conditions
from Hoare logic

Verification Component Outline

approach

1. use KAT as abstract algebraic semantics for while-programs
2. define validity of Hoare triples in KAT
3. derive rules of Hoare logic without assignment in KAT
4. derive assignment rule in relation KAT over program store
5. use Isabelle polymorphism to integrate arbitrary data domains
6. use KAT/Hoare logic for verification condition generation
7. use domain-specific Isabelle components to verify programs

tool correct by construction

Hoare Triples in KAT

validity of Hoare triple

$$\vdash \{p\} x \{q\} \Leftrightarrow px\neg q = 0$$

intuition (partial correctness)

if program x is executed from state where p holds and if x terminates, then q must hold in state where x terminates

in relation KAT

$$\begin{aligned} & \forall s, s'. (s, s') \notin px\neg q \\ & \Leftrightarrow \forall s, s'. \neg((s, s) \in p \wedge (s, s') \in x \wedge (s', s') \in \neg q) \\ & \Leftrightarrow \forall s, s'. ((s, s) \in p \wedge (s, s') \in x) \Rightarrow (s', s') \in q \end{aligned}$$

Propositional Hoare Logic

propositional Hoare logic means Hoare logic without assignment rule

theorem [Kozen]

inference rules of PHL derivable in KAT

$$\begin{aligned} & \vdash \{p\} \text{ skip } \{p\} \\ p \leq p' \wedge q' \leq q \wedge \vdash \{p'\} x \{q'\} & \Rightarrow \vdash \{p\} x \{q\} \\ \vdash \{p\} x \{r\} \wedge \vdash \{r\} y \{q\} & \Rightarrow \vdash \{p\} x; y \{q\} \\ \vdash \{pb\} x \{q\} \wedge \vdash \{p\neg b\} y \{q\} & \Rightarrow \vdash \{p\} \text{ if } b \text{ then } x \text{ else } y \text{ fi } \{q\} \\ \vdash \{pb\} x \{p\} & \Rightarrow \vdash \{p\} \text{ while } b \text{ do } x \text{ od } \{\neg bp\} \end{aligned}$$

Store and Assignments

simple store in Isabelle

- stores formalised as functions from variable to values
- generic for any type of data (KAT/relation KAT polymorphic)
- variables formalised as strings
- values can have any type

assignment

$$(v := e) = \{(s, \text{fun_upd } s \ v \ (e \ s)) \mid s \in S\}$$

theorem

all inference rules of HL are derivable in relation KAT with store

$$\vdash \{Q[e/v]\} (v := e) \{Q\}$$

Verification Condition Generation

Hoare logic

- one structural rule per program construct
- can be programmed as **hoare** tactic in Isabelle
- blasts away entire control structure

derivable rules

$$p \leq p' \wedge \vdash \{p'\} x \{q\} \Rightarrow \vdash \{p\} x \{q\}$$

$$p \leq i \wedge \neg pi \leq q \wedge \vdash \{ib\} x \{i\} \Rightarrow \vdash \{p\} \mathbf{while} \ b \ \mathbf{inv} \ i \ \mathbf{do} \ x \ \mathbf{od} \ \{q\}$$

Verification Component with Isabelle

control flow

- Isabelle libraries for KAT include PHL
- **hoare** tactic generates verification conditions automatically from HL

data flow

- modelled generically in relation KAT (with store)
- shallow embedding of simple while-language
- analysed with Isabelle's provers
- functional data types often impersonate imperative data structures
- could use data refinement as justification. . .

— demo —

Refinement Component based on KAT

Refinement KAT

definition

refinement KAT is KAT expanded by **specification statement** $[_ , _]$ and axiom

$$\vdash \{p\} x \{q\} \Leftrightarrow x \leq [p, q]$$

theorem

$(2^{A \times A}, B, \cup, \circ, [_, _], *, \neg, \emptyset, id)$ forms rKAT with

$$[P, Q] = \bigcup \{R \subseteq A \times A \mid \vdash \{P\} R \{Q\}\}$$

Propositional Refinement Calculus

theorem

Morgan's propositional refinement laws are derivable in rKAT ($\sqsubseteq = \geq$)

$$\begin{aligned} p \leq q &\Rightarrow [p, q] \sqsubseteq \mathbf{skip} \\ p \leq p' \wedge q' \leq q &\Rightarrow [p, q] \sqsubseteq [p', q'] \\ [0, 1] &\sqsubseteq x \\ x &\sqsubseteq [1, 0] \\ [p, q] &\sqsubseteq [p, r]; [r, q] \\ [p, q] &\sqsubseteq \mathbf{if } b \mathbf{ then } [bp, q] \mathbf{ else } [\neg bp, q] \mathbf{ fi} \\ [p, \neg bp] &\sqsubseteq \mathbf{while } b \mathbf{ do } [bp, p] \mathbf{ od} \end{aligned}$$

no frame laws for local variables

Refinement Calculus

theorem

assignment laws derivable in relation rKAT

$$P \subseteq Q[e'/x] \Rightarrow [P, Q] \sqsubseteq (v := e)$$

$$Q' \subseteq Q[e'/x] \Rightarrow [P, Q] \sqsubseteq [P, Q']; (v := e)$$

$$P' \subseteq P[e'/x] \Rightarrow [P, Q] \sqsubseteq (v := e); [P'; Q]$$

— demo —

Verification Component based on Predicate Transformers

Outline

KAD	relational KAD	relational KAD over store
control flow	abstract data flow	concrete data flow
transformers (PDL)	relational verification conditions	verification conditions based on transformers

Verification Component Outline

approach

1. use KAD as abstract algebraic semantics for while-programs
2. define partial correctness specification in KAD
3. derive predicate transformer laws without assignment in KAD
4. derive assignment law in relation KAD over program store
5. use Isabelle polymorphism to integrate arbitrary data domains
6. use KAD predicate transformer laws for vcg
7. use domain-specific Isabelle components to verify programs

tool correct by construction

Adding Modalities

motivation

- many applications require different approach to actions/propositions
- systems dynamics by action on state space $K \rightarrow B \rightarrow B$
- computational logics (e.g. PDL) “use” KAs, but how precisely?

modal approach

- actions/propositions via relational (aka Kripke) frames
- modal operators via preimages/images $|x\rangle p / \langle x|p$
- preimages/images via axioms for **domain/codomain**

State Transitions

in KAT

“terminating program x from store p goes to store q ” expressed as

$$px \leq xq \quad \text{or equivalently} \quad px \neg q = 0$$

alternative

“ q contains x -image of p ”

how can we model relational (pre)images directly in semirings?

Adding Modalities

task

- abstract equational axioms for relational domain

$$d x = \{(p, p) \mid \exists q. (p, q) \in x\}$$

- algebraic definition of relational modalities
 - $\langle x \mid p = \text{ran}(px)$ is **image** of p under x
 - $\mid x \rangle p = \text{dom}(xp)$ by opposition, is **preimage** of p under x

two approaches

1. domain as map $K \rightarrow B$ in KAT
2. domain as endo $S \rightarrow S$ that induces B in dioid

Domain Semirings

domain semiring

semiring S with $d : S \rightarrow S$ that satisfies

$$\begin{aligned}x + d x \cdot x &= d x \cdot x & d(x \cdot y) &= d(x \cdot d y) & d(x + y) &= d x + d y \\d x + 1 &= 1 & d 0 &= 0\end{aligned}$$

lemma

domain semirings are dioids

proposition

$d^2 = d$ (domain is retraction), so $x \in d[S] \Leftrightarrow d x = x$

Domain Algebra

theorem

$(d[S], +, \cdot, 0, 1)$ is bounded DL (d induces state space)

notation

- $(d[S], +, \cdot, 0, 1)$ is called **domain algebra** of S
- $p, q, r \dots$ for domain elements

modalities

$$\langle x \rangle y = d(xy)$$

$$\langle x \rangle y = r(yx)$$

how can we obtain boolean state space?

Antidomain Semirings

antidomain semiring

semiring S with endo $a : S \rightarrow S$ that satisfies

$$ax \cdot x = 0 \quad a(x \cdot y) \leq a(x \cdot a^2 y) \quad a^2 x + ax = 1$$

remarks

- domain definable as $d = a^2$ (boolean complement)
- $a[S](= d[S])$ generated is **maximal** BA in $[0, 1]$
- simple axioms induce rich modal calculus...

diamonds again

$$|x\rangle y = d(xy) \quad \langle x|y = r(yx)$$

Dualities for Modalities

$$\langle x \rangle p = d(xp)$$

$$\langle x \rangle p = a(xa(p))$$

$$\langle x | p = r(px)$$

$$[x | p = ar(ar(p)x)$$

Dualities for Modalities

$$\begin{array}{c} |x\rangle p = d(xp) \\ \updownarrow \text{opposition} \\ \langle x|p = r(px) \end{array}$$

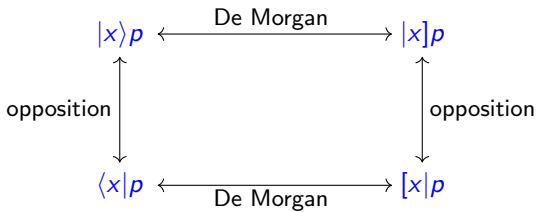
$$\begin{array}{c} [x]p = a(xa(p)) \\ \updownarrow \text{opposition} \\ [x|p = ar(ar(p)x) \end{array}$$

Dualities for Modalities

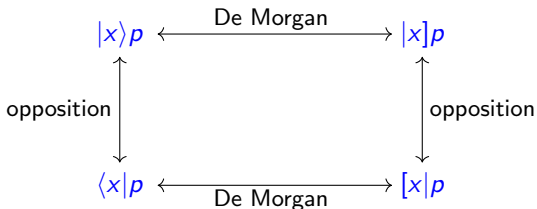
$$\begin{array}{c} \langle x \rangle p = d(xp) \\ \updownarrow \text{opposition} \\ \langle x \rangle p = r(px) \end{array}$$

$$\begin{array}{c} [x]p = \neg \langle x \rangle \neg p \\ \updownarrow \text{opposition} \\ [x]p = \neg \langle x \rangle \neg p \end{array}$$

Dualities for Modalities



Dualities for Modalities



- conjugations

$$(|x]p)q = 0 \Leftrightarrow p(\langle x|q) = 0 \quad (|x]p)q = 0 \Leftrightarrow p([x|q) = 0$$

- adjunctions

$$\langle x | p \leq q \Leftrightarrow p \leq |x] q \quad |x \rangle p \leq q \Leftrightarrow p \leq [x | q$$

Dualities for Modalities

demodalisation

$$\begin{aligned} |x\rangle p \leq q &\Leftrightarrow \neg q x p \leq 0 \Leftrightarrow p \leq [x|q \\ \langle x|p \leq q &\Leftrightarrow p x \neg q \leq 0 \Leftrightarrow p \leq |x]q \end{aligned}$$

properties

- conjugations/adjunctions as **theorem generators**
- dualities as **theorem transformers**

MKA and KAT

theorem

every KA with antidomain is a KAT (... but not conversely)

a Hoare logic is **expressive** if for each command x and postcondition q the weakest liberal precondition is definable

theorem

MKA is expressive for Hoare logic (... KAT isn't)

proof

- $|x]q = wlp(x, q)$ for each command x and postcondition q
- in relational model $|R]Q = \bigcup\{P \mid \{P\} R \{Q\}\}$

Control Elimination

partial correctness specification

$$p \leq [x]q$$

predicate transformer laws

- $[xy]q = [x][y]q$
- $[\text{if } p \text{ then } x \text{ else } y]q = (\neg p + [x]q)(p + [y]q)$
- $p \leq i \wedge i \neg t \leq q \wedge it \leq [x]i \Rightarrow p \leq [\text{while } t \text{ inv } i \text{ do } x]q$

recursive wp/vc computation

Control Elimination

partial correctness specification

$$p \leq [x]q$$

predicate transformer laws

- $[xy]q = [x][y]q$
- $[\text{if } p \text{ then } x \text{ else } y]q = (\neg p + [x]q)(p + [y]q)$
- $p \leq i \wedge i \neg t \leq q \wedge it \leq [x]i \Rightarrow p \leq [\text{while } t \text{ inv } i \text{ do } x]q$

but what about assignment?

Data Integration

1. relational model

- $(\text{Rel}(X), \cup, \cap, \emptyset, Id, *)$ forms **relation MKA** over X with
 - ▷ $aR = \{(a, a) \mid \neg \exists b. (a, b) \in R\}$
 - ▷ $arR = \{(b, b) \mid \neg \exists a. (a, b) \in R\}$
- **subidentities** $\{P \in \text{Rel}(X) \mid P \subseteq Id\}$ form boolean subalgebra

2. relational store model

- store as function $V \rightarrow E$ from variables to values
- assignments defined by $(v := e) = \{(s, s[(es)/v]) \mid s \in E^V\}$
- wp law for assignments derivable: $|v := e|[Q] = [\lambda s. Q s[(es)/v]]$

Modalities vs Predicate Transformers

relation $R \subseteq X \times Y$ gives rise to three transformers:

- **state transformer** $f_R : X \rightarrow 2^Y$ defined by

$$f_R x = \{y \mid (x, y) \in R\}$$

- **conjunctive predicate transformer** $|R| : 2^Y \rightarrow 2^X$ defined by

$$|R|P = \{x \mid f_R x \subseteq P\}$$

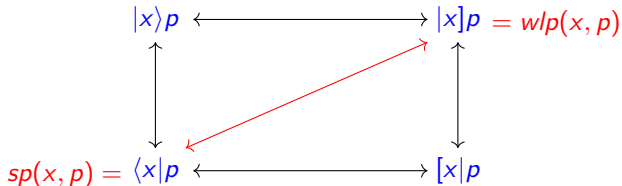
- **disjunctive predicate transformer** $\langle R \rangle : 2^X \rightarrow 2^Y$ defined by

$$\langle R \rangle p = \bigcup \{f_R x \mid x \in p\}$$

Isabelle Verification Component

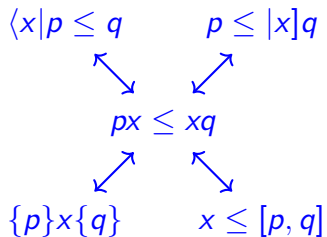
—demo—

SP for Free

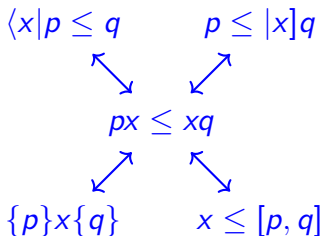


- adjunction $\langle x|p \leq q \Leftrightarrow p \leq |x]q$ dualises wp-laws
- Floyd-style assignment law works well in this setting
- useful for symbolic execution

Hoare Logic and Refinement



Hoare Logic and Refinement



- PHL rules derivable in MKA
- Morgan-style refinement calculus derivable in **refinement MKA**
- but MKA **is** refinement calculus
- assignment rules derivable in relational store model
- we link into KAT/rKAT components instead

—demo—

Literature

- algebra and HL
 - ▶ Kozen, *On Hoare Logic and Kleene Algebra with Tests*
 - ▶ Desharnais, Struth, *Internal Axioms for Domain Semirings*
 - ▶ Möller, Struth, *Algebras of Modal Operators and Partial Correctness*
- refinement calculi
 - ▶ Morgan, *Programming from Specifications*
 - ▶ Back, von Wright, *Refinement Calculus: A Systematic Introduction*
- Isabelle formalisations
 - ▶ Armstrong, Gomes, Struth, *Building Program Construction and Verification Tools from Algebraic Principles*
 - ▶ Struth, *Hoare Semigroups*

Conclusion

- principled approach to program correctness tools in Isabelle
 - use algebra at control flow layer
 - link with relation/predicate transformer semantics and store
 - derived Hoare logics or refinement calculi
- all algebras used have decidable fragments
- sequential program verification works smoothly
- concurrency verification (still) more tedious
- prototyping fast, simple, adaptable
- resulting tools lightweight