On Algebra of Program Correctness and Incorrectness

Bernhard Möller, Peter O'Hearn and Tony Hoare

RAMiCS Nov 2021

Context

imperative programs may be reasoned about in two ways:

- show *absence* of bugs
  - safety aspect: nothing bad can ever happen
  - logics of Floyd/Hoare/Dijkstra (1967–1975), briefly (HL)
- show *presence* of bugs to correct them
  - liveness aspect: something expected (but not necessarily good) can always happen
  - Incorrectness Logic (IL) of O'Hearn (POPL 2020)

Assertion Triples

- classical: HL triples

$$\{pre\}\ code\ \{post\}$$

  $code$ guarantees to lead from pre-states to $post$-states

- IL triples

$$[pre]\ code\ [post]$$

  all $post$-states can be reached under $code$ from some pre-state

- when $post$ specifies erroneous situations: all pre-states might trigger an error

Relational Formalisation

for a simple semantics consider programs as relations between states

- $\mathsf{img}(P, C)$ means the image of state set P under relation C
- with this, the triples can be expressed as

$$\{P\}\ C\ \{Q\} \quad \Leftrightarrow_{df} \quad \mathsf{img}(P, C) \subseteq Q \qquad \text{(over-approximation)}$$

$$[P]\ C\ [Q] \quad \Leftrightarrow_{df} \quad Q \subseteq \mathsf{img}(P, C) \qquad \text{(under-approximation)}$$

- IL-like triples were already introduced by de Vries/Koutavas 2011, but for a quite different purpose

Programming Constructs

- essential *atomic commands*:
  - `skip`
  - `abort`
  - `assume(`p`)`
  - expression assignment `x := e`
  - nondeterministic assignment `x := nondet()`
    (also used to "forget" previous value of `x`)
- general *commands* constructed from atomic ones using
  - choice $\cup$
  - sequential composition ;
  - Kleene star * (reflexive transitive closure)

further constructs defined in terms of these

- $\texttt{if}\ P\ \texttt{then}\ C\ \texttt{else}\ V' =_{df} (\texttt{assume}(P); V) \cup (\texttt{assume}(\neg P); C')$
- $\texttt{while}\ P\ \texttt{do}\ C =_{df} (\texttt{assume}(P)\ ;\ C)^*\ ;\ \texttt{assume}(\neg P)$

A Toy Example (More Interesting Ones in the Paper)

here is a triple that is true in IL but not in HL:

$$[\mathtt{x} = 0]\ (\mathtt{x} := \mathtt{x} + 1)^*\ [\mathtt{x} > 0]$$

since we can iterate the loop body an arbitrary number of times, every value $x > 0$ is reachable from any pre-state

contrarily, the triple

$$\{\mathtt{x} = 0\}\ (\mathtt{x} := \mathtt{x} + 1)^*\ \{\mathtt{x} > 0\}$$

fails, since the loop body might be iterated zero times □

Algebraic Abstraction: Modal Semirings

abstraction:

relations R	$\rightarrow$	elements a of a modal semiring
predicates P	$\rightarrow$	tests p
$\mathsf{img}(P, C)$	$\rightarrow$	$\langle a \vert p$ (backward diamond)
`abort`	$\rightarrow$	0
`skip`	$\rightarrow$	1
$\cup$	$\rightarrow$	$+$
;	$\rightarrow$	$\cdot$
$\subseteq$	$\rightarrow$	$\leq$ (semiring order)

- $\langle a|p$ equals the strongest postcondition $\mathsf{sp}(a,p)$
- which is symmetric to the weakest liberal precondition $\mathsf{wlp}(a,p)$, represented by the forward box $|a]p$
- Galois connection

$$\langle a|p \leq q \quad \Leftrightarrow \quad p \leq |a]q$$

- hence $\langle a|$ preserves all existing suprema, $|a]$ all existing infima
- still this is not immediately useful, since the rhs of the IL triple $q \leq \langle a|p$ is the reverse of the one in the Galois connection
- nevertheless there is a lot of (albeit sometimes treacherous) symmetry between classical and IL triples

Inference Rules

- some rules for the non-looping constructs
  (side by side with corresponding HL rules)

$$\frac{[p]\ b\ [q] \qquad b \leq a}{[p]\ a\ [q]} \qquad \frac{\{p\}\ b\ \{q\} \qquad b \geq a}{\{p\}\ a\ \{q\}}$$

$$\frac{[p]\ a\ [q]}{[p]\ a + b\ [q]} \qquad \frac{\{p\}\ a + b\ \{q\}}{\{p\}\ a\ \{q\}}$$

Inference Rules

- rules of sequence and consequence

$$\frac{[p]\ a\ [r] \quad [r]\ b\ [q]}{[p]\ a \cdot b\ [q]} \qquad \frac{\{p\}\ a\ \{r\} \quad \{r\}\ b\ \{q\}}{\{p\}\ a \cdot b\ \{q\}}$$

$$\frac{p' \geq p \quad [p]\ a\ [q] \quad q \geq q'}{[p']\ a\ [q']} \qquad \frac{p' \leq p \quad \{p\}\ a\ \{q\} \quad q \leq q'}{\{p'\}\ a\ \{q'\}}$$

Rules for Loops

For finite iteration things work out well:

$$[p]\ a^n\ [q] \ \Rightarrow\ [p]\ a^{\leq n}\ [q] \ \Rightarrow\ [p]\ a^*\ [q]$$

$$\frac{\forall\, i \leq n : ([p]\ a^i\ [q_i])}{[p]\ a^*\ [\bigvee_{i \leq n} q_i]}$$

also the expected fold rules hold:

$$\frac{[p]\ a\ [q]}{[p]\ a^*\ [q]} \qquad \frac{[p]\ a \cdot a^*\ [q]}{[p]\ a^*\ [q]}$$

these rules follow directly from isotony

Rules for Loops

- apart from that, however, things get quite different from HL
- we have $\langle a^* | p \;=\; \mu f$ where $f(q) = p + \langle a | q$
- hence least fixed point induction applies (*diamond star induction*)
- this yields the HL inference rule

$$\frac{p \leq q \qquad \{q\}\ a\ \{q\}}{\{p\}\ a^*\ \{q\}}$$

  where q is an invariant of a
- but $[p]\ a\ [q]$ is equivalent to $q \leq \mu f$ and no general proof principle for such inequations is available
- in particular, invariants are not useful any more

Rules for Loops

- one possible loop rule is (similar to de Vries/Koutavas 2011)

$$\frac{\forall\, n \in \mathbb{N} :\ [p_n]\ a\ [p_{n+1}]}{[p_0]\ a^*\ [\bigvee_{n\in\mathbb{N}} p_n]} \quad \text{(backwards variant)}$$

- the p_i are in a sense counterparts of variants as used in termination proofs for while programs
- the rule says that the iteration covers at least part of the full transitive a-image of p_0
- however, the infinite disjunction $\bigvee_{n\in\mathbb{N}} p_n$ is problematic, since its existence is not guaranteed in general semirings

Rules for Loops

- call a modal semiring *countably test complete (CTC)* if every countable set $\{p_n \mid n \in \mathbb{N}\}$ has a supremum, denoted by $\bigvee_{n\in\mathbb{N}} p_n$
- by elementary order theory this is equivalent to saying that every countable ascending chain $p_0 \leq p_2 \leq \cdots$ has a supremum
- henceforth we assume a CTC underlying semiring
- in practical applications $\bigvee_{n\in\mathbb{N}} p_n$ is written as $\exists\, n : p_n$
- also we may replace $\geq$ between tests by $\Leftarrow$

Rules for Loops

- recall the loop rule

$$\frac{\forall\, n \in \mathbb{N} :\ [p_n]\ a\ [p_{n+1}]}{[p_0]\ a^*\ [\bigvee_{n\in\mathbb{N}} p_n]} \quad \text{(backwards variant)}$$

- by choosing some predicate r above p_0 and using the rule of consequence we obtain a more general version:

$$\frac{r \geq p_0 \qquad \forall\, n \in \mathbb{N} : [p_n]\ a\ [p_{n+1}] \qquad (\bigvee_{n\in\mathbb{N}} p_n) \ \geq\ q}{[r]\ a^*\ [q]}$$

- expresses that one may need to iterate indefinitely to cover q

Toy Example Cont'd

- we first prove $[\mathtt{x} = 0]\ (\mathtt{x} := \mathtt{x} + 1)^*\ [\mathtt{x} \geq 0]$
- for this, find P_n with

$$(\mathtt{x} = 0) \Leftarrow P_0 \qquad [P_n]\ \mathtt{x := x+1}\ [P_{n+1}] \qquad (\exists\, n : P_n) \Leftarrow (\mathtt{x} \geq 0)$$

- these conditions can be satisfied by choosing

$$P_n =_{df} (\mathtt{x} = n \wedge n \geq 0)$$

- now by the rule of consequence we can shrink the post-condition to $\mathtt{x} > 0$ □

More on Algebra

the semiring of relations has an extremely rich structure:

- its carrier is the power set $\mathcal{P}(M \times M)$ for set M of states
- hence a complete lattice and even a Boolean quantale
- in particular it is CTC
- the completeness is also deployed in the standard closed star representation:

$$C^* = \bigcup_{n \in \mathbb{N}} C^i$$

- can we make do with weaker algebraic concepts for expressing IL?

More on Algebra

- in connection with the Kleene star quantales have already been weakened to $*$-continuous semirings [Kozen 1980]
- these need suprema only for sets $\{a^n \mid n \in \mathbb{N}\}$
- composition $\cdot$ needs to distribute only through such suprema
- but maybe we can get away with still weaker assumptions?

- yes, we can!
- as a kind of surprise, CTC is already enough!

Theorem the presented IL calculus is relatively complete (i.e., allows proving all valid IL triples)

- as a quantale, relation algebra is CTC
- hence this result subsumes the concrete completeness proof in Peter O'Hearn's original POPL 2020 paper
- approach analogous to an earlier one for HL (Möller/Struth 2005)

let's be a bit more precise in a sketch of the proof

- by isotony and diamond star induction, every CTC modal Kleene is "$*$-continuous under the diamond", i.e., all elements a and tests p satisfy $\langle a^* | p = \bigvee_{n \in \mathbb{N}} \langle a^n | p$
- a *command* is a Kleene algebra element generated from 0, 1 and a set of atomic commands and arbitrary tests using $+$, $\cdot$ and *

- we assume for every atomic command a the axiom

$$\frac{}{[p]\ a\ [\langle a|p]}$$

  (in the concrete relational version of IL this holds)
- induction on the generation structure of command a in a modal CTC Kleene algebra shows that all triples $[p]\ a\ [\langle a|p]$ are provable
- completeness now follows since for any valid IL triple $[p]\ a\ [q]$, i.e., $q \leq \langle a|p$, we infer from $[p]\ a\ [\langle a|p]$ using the rule of consequence that $[p]\ a\ [q]$ is provable

Conclusion

- Kleene algebra abstracts basic principles of imperative programs
- naturally forms a foundation for the (partial) correctness logic HL
- we have shown that it also allows a foundation for the incorrectness logic IL
- the corresponding triples can be directly translated into inequational formulas of Modal Kleene algebra,
- dual to the ones for HL
- hence Modal Kleene algebra can be said to unify correctness and incorrectness logic
- the same applies to a quite recent equivalent, independently developed, approach using KATs with top elements by Zhang et al. (to appear at POPL 2022)

Conclusion

What else is in the paper?

- application of the logic for disproving conjectured HL triples
- refined semantics with error handling
- a variant of incorrectness logic
- pinpointing the properties really relevant to relative completeness of the logic

Conclusion

some open questions

- how to extend correctness and incorrectness logics from mere safety to liveness or hyperproperties
- how to extend our results to further programming features
- examples: Concurrent Kleene Algebra and Concurrent Separation Logic

Some References

- D. Kozen: A representation theorem for models of $*$-free PDL. In J. de Bakker, J. van Leeuwen (eds): Automata, Languages and Programming. LNCS 85. Springer 1980, 351–362
- E. de Vries, V. Koutavas: Reverse Hoare logic. 9th SEFM. LNCS 7041. pp 155–171, 2011
- B. Möller, G. Struth: Algebras of modal operators and partial correctness. Theoretical Computer Science 351, 221–239 (2006)
- P. O'Hearn: Incorrectness logic. POPL 2020, 10:1–10:32
- Cheng Zhang, Arthur Azevedo de Amorim, Marco Gaboardi: On incorrectness logic and Kleene algebra with top and tests. Accepted for POPL 2022