

# Can we automate proof of contextual equivalence using Logical Relations?

Guilhem Jaber

PPS, IRIF, Université Paris Diderot



Workshop on Program Equivalence  
April 12th 2016

- Present techniques to prove *contextual equivalence* of programs:
  - ↪ Program as a black-box
  - ↪ the context can call it many time and whenever he wants,

- Present techniques to prove *contextual equivalence* of programs:
  - ↪ Program as a black-box
  - ↪ the context can call it many time and whenever he wants,
- for a functional language:
  - ↪ Open programs and callbacks,
  - ↪ Reentrant calls,

- Present techniques to prove *contextual equivalence* of programs:
  - ↪ Program as a black-box
  - ↪ the context can call it many time and whenever he wants,
- for a functional language:
  - ↪ Open programs and callbacks,
  - ↪ Reentrant calls,
- with references:
  - ↪ Local states,
  - ↪ Disclosure between the program and the context.

- Present techniques to prove *contextual equivalence* of programs:
  - ↪ Program as a black-box
  - ↪ the context can call it many time and whenever he wants,
- for a functional language:
  - ↪ Open programs and callbacks,
  - ↪ Reentrant calls,
- with references:
  - ↪ Local states,
  - ↪ Disclosure between the program and the context.

## Kripke Logical Relations

- Present techniques to prove *contextual equivalence* of programs:
  - ↪ Program as a black-box
  - ↪ the context can call it many time and whenever he wants,
- for a functional language:
  - ↪ Open programs and callbacks,
  - ↪ Reentrant calls,
- with references:
  - ↪ Local states,
  - ↪ Disclosure between the program and the context.

## Kripke Logical Relations

Can we automate them ?

## For what kind of Language: RefML

A typed functional programs : `fun test(n, g) = n + g(1)`

## For what kind of Language: RefML

A typed functional programs : `fun test(n, g) = n + g(1)`

with Integers and Booleans: `if b then 0 else n + 1`

with pairs: `<u, v>`

## For what kind of Language: RefML

A typed functional programs :	<code>fun test(n, g) = n + g(1)</code>
with Integers and Booleans:	<code>if b then 0 else n + 1</code>
with pairs:	<code>&lt;u, v&gt;</code>
with higher-order references:	<code>ref 2, ref (<math>\lambda x.M</math>)</code>
stored in heap via locations:	<code>(ref v, h) <math>\rightarrow</math> (<math>\ell, h \cdot [\ell \mapsto v]</math>)</code> <code>(<math>\ell</math> fresh in <math>h</math>)</code>
mutable:	<code>x :=!x + 1</code>

## For what kind of Language: RefML

A typed functional programs :	<code>fun test(n, g) = n + g(1)</code>
with Integers and Booleans:	<code>if b then 0 else n + 1</code>
with pairs:	<code>&lt;u, v&gt;</code>
with higher-order references:	<code>ref 2, ref (<math>\lambda x.M</math>)</code>
stored in heap via locations:	<code>(ref v, h) <math>\rightarrow</math> (<math>l, h \cdot [l \mapsto v]</math>)</code> <code>(<math>l</math> fresh in <math>h</math>)</code>
mutable:	<code>x :=!x + 1</code>
No pointer arithmetic:	<code>(<math>l + 1</math>)</code> is ill-typed
But equality test:	<code><math>l_1 == l_2</math></code> is well-typed

## For what kind of Language: RefML

A typed functional programs :	<code>fun test(n, g) = n + g(1)</code>
with Integers and Booleans:	<code>if b then 0 else n + 1</code>
with pairs:	<code>&lt;u, v&gt;</code>
with higher-order references:	<code>ref 2, ref (<math>\lambda x.M</math>)</code>
stored in heap via locations:	<code>(ref v, h) <math>\rightarrow</math> (<math>l, h \cdot [l \mapsto v]</math>)</code> <code>(<math>l</math> fresh in <math>h</math>)</code>
mutable:	<code>x := !x + 1</code>
No pointer arithmetic:	<code>(<math>l + 1</math>)</code> is ill-typed
But equality test:	<code><math>l_1 == l_2</math></code> is well-typed
Full recursion (via Landin “knot”).	

# Contextual Equivalence

Contextual equivalence of  $M_1, M_2$ :

$$\forall C. \forall h. (C[M_1] \Downarrow, h) \iff (C[M_2] \Downarrow, h)$$

Observation  $(M \Downarrow, h)$  : Termination.

# Contextual Equivalence

Contextual equivalence of  $M_1, M_2$ :

$$\forall C. \forall h. (C[M_1] \Downarrow, h) \iff (C[M_2] \Downarrow, h)$$

Observation  $(M \Downarrow, h)$  : Termination.

- Robust w.r.t the choice of observation.
- Depend on the language contexts are written in.

# Contextual Equivalence

Contextual equivalence of  $M_1, M_2$ :

$$\forall C. \forall h. (C[M_1] \Downarrow, h) \iff (C[M_2] \Downarrow, h)$$

Observation  $(M \Downarrow, h)$  : Termination.

- Robust w.r.t the choice of observation.
- Depend on the language contexts are written in.
- Undecidable in general
  - ↳ Even in a finitary setting (finite datatypes, no recursion): Murawski & Tzevelekos

# Contextual Equivalence

Contextual equivalence of  $M_1, M_2$ :

$$\forall C. \forall h. (C[M_1] \Downarrow, h) \iff (C[M_2] \Downarrow, h)$$

Observation  $(M \Downarrow, h)$  : Termination.

- Robust w.r.t the choice of observation.
- Depend on the language contexts are written in.
- Undecidable in general
  - ↪ Even in a finitary setting (finite datatypes, no recursion): Murawski & Tzevelekos
- Many techniques to reason on contextual equivalence:
  - ↪ Algorithmic Game Semantics,
  - ↪ Environmental Bisimulations, Open Bisimulations,
  - ↪ Kripke Logical Relations.

## Synchronization of Callbacks (I/II)

`fun test1(f1) = f1()` is **not** equivalent to  
`fun test2(f2) = f2(); f2()`

# Synchronization of Callbacks (I/II)

`fun test1(f1) = f1()` is **not** equivalent to  
`fun test2(f2) = f2(); f2()`

- ↪ Contexts can check how many time  $f_1, f_2$  are called.
- ↪ Callbacks are fully observable!

$C[\bullet] \stackrel{def}{=} \text{let } x = \text{ref } 0 \text{ in } \bullet (\lambda_.x := !x + 1); \text{if } !x > 1 \text{ then } \Omega \text{ else}()$   
can discriminate them.

## Synchronization of Callbacks (II/II)

`fun add1(f1) = (f1 1) + (f1 2)` is **not** equivalent to  
`fun add2(f2) = (f2 2) + (f2 1)`

## Synchronization of Callbacks (II/II)

`fun add1(f1) = (f1 1) + (f1 2)` is **not** equivalent to  
`fun add2(f2) = (f2 2) + (f2 1)`

↪ Arguments given to callbacks must be related.

$C[\bullet] \stackrel{def}{=} \text{let } x = \text{ref } 0 \text{ in } \bullet (\lambda y.x := y); \text{if } !x == 1 \text{ then } \Omega \text{ else}()$   
can discriminate them.

## Disclosure of Locations (I/II)

`fun const1() = let x = ref 0 in 1` is equivalent to  
`fun const2() = 1`

- ↪ The creation of the reference bounded to `x` is not observable by the context.
- ↪ It is private to the term!

## Disclosure of Locations (II/II)

`fun discl1(f) = let x = ref0 in fx; x := 1`

is **not** equivalent to

`fun discl2(f) = let x = ref0 in fx; x := 2`

## Disclosure of Locations (II/II)

```
fun discl1(f) = let x = ref 0 in fx; x := 1
```

is **not** equivalent to

```
fun discl2(f) = let x = ref 0 in fx; x := 2
```

- ↪ The reference bound to  $x$  is disclosed to the context.
- ↪ It can look inside afterwards to see what is stored.

$C[\bullet] \stackrel{\text{def}}{=} \text{let } z = \text{ref}(\text{ref } 0) \text{ in } \bullet (\lambda y. z := y); \text{if } !!z == 1 \text{ then } \Omega \text{ else}()$   
can discriminate them.

# Representation Independence

The two following programs are equivalent:

```
let c1 = ref0
    fun inc1() . c1 := !c1 + 1
    fun get1() . !c1
in ⟨inc1, get1⟩
```

---

```
let c2 = ref0
    fun inc2() . c2 := !c2 - 1
    fun get2() . -!c2
in ⟨inc2, get2⟩
```

Need a relational invariant between  $c_1$  and  $c_2$ .

# Quiz

Are the two following programs equivalent?

```
let c1 = ref0
    fun inc1(f) = f(); c1 := !c1 + 1
    fun get1() = !c1
in ⟨inc1, get1⟩
```

---

```
let c2 = ref0
    fun inc2(f) = let n = !c2 in f(); c2 := n + 1
    fun get2() = !c2
in ⟨inc2, get2⟩
```

# Quiz

Are the two following programs equivalent?

```
let c1 = ref0
    fun inc1(f) = f(); c1 := !c1 + 1
    fun get1() = !c1
in ⟨inc1, get1⟩
```

---

```
let c2 = ref0
    fun inc2(f) = let n = !c2 in f(); c2 := n + 1
    fun get2() = !c2
in ⟨inc2, get2⟩
```

No, because of reentrant calls !

$$C[\bullet] \stackrel{def}{=} \text{let } \langle \text{inc}, \text{get} \rangle = \bullet \text{ in let } d = \text{get}() \text{ in} \\ \text{inc}(\lambda \_ . \text{inc} (\lambda x . x)); \text{ if } \text{get}() \neq d + 2 \text{ then } \Omega \text{ else } ()$$

can discriminate them.

# Logical Relations

Binary relations  $\mathcal{E} \llbracket \tau \rrbracket, \mathcal{V} \llbracket \tau \rrbracket$  on closed terms and values

$\rightsquigarrow$  inductively defined on types.

$$\mathcal{V} \llbracket \text{Int} \rrbracket \stackrel{\text{def}}{=} \{(n, n) \mid n \in \mathbb{Z}\}$$

$$\mathcal{V} \llbracket \tau \rightarrow \sigma \rrbracket \stackrel{\text{def}}{=} \{(\lambda x_1. M_1, \lambda x_2. M_2) \mid \forall (v_1, v_2) \in \mathcal{V} \llbracket \tau \rrbracket. \\ ((\lambda x_1. M_1)v_1, (\lambda x_2. M_2)v_2) \in \mathcal{E} \llbracket \sigma \rrbracket\}$$

$$\mathcal{E} \llbracket \tau \rrbracket \stackrel{\text{def}}{=} \{(M_1, M_2) \mid (M_1 \uparrow \wedge M_2 \uparrow) \\ \vee ((M_1 \mapsto^* v_1) \wedge (M_2 \mapsto^* v_2) \wedge (v_1, v_2) \in \mathcal{V} \llbracket \tau \rrbracket)\}$$

## Operational reasoning for functions with local state

*Andrew Pitts and Ian Stark*

### Abstract

Languages such as ML or Lisp permit the use of recursively defined function expressions with locally declared storage locations. Although this can be very convenient from a programming point of view it severely complicates the properties of program equivalence even for relatively simple fragments of such languages—such as the simply typed fragment of Standard ML with integer-valued references considered here. This paper presents a method for reasoning about *contextual equivalence* of programs involving this combination of functional and procedural features. The method is based upon the use of a certain kind of *logical relation* parameterised by relations between program states. The form of this logical relation is novel, in as much as it involves relations not only between program expressions, but also between program continuations (also known as *evaluation contexts*). The authors found this approach necessary in order to establish the ‘Fundamental Property of logical relations’ in the presence of both dynamically allocated local state and recursion. The logical relation characterises contextual equivalence and yields a proof of the best known context lemma for this kind of language—the Mason-Talcott ‘ciu’ theorem. Moreover, it is shown that the method can prove examples where such a context lemma is not much help and which involve representation independence, higher order memoising functions, and profiling functions.

# Kripke Logical Relations

Extension to languages with references.

- ↪ Need *worlds*  $w$ , i.e. invariants on heaps,
- ↪ Extensible on disjoint parts :  $w' \sqsupseteq w$
- ↪ Parametrize the definition of logical relations with such worlds.

# Kripke Logical Relations

Extension to languages with references.

↪ Need *worlds*  $w$ , i.e. invariants on heaps,

↪ Extensible on disjoint parts :  $w' \sqsupseteq w$

↪ Parametrize the definition of logical relations with such worlds.

$$\mathcal{E} \llbracket \tau \rrbracket w \stackrel{\text{def}}{=} \left\{ (M_1, M_2) \mid \forall (h_1, h_2) : w. ((M_1, h_1) \uparrow \wedge (M_2, h_2) \uparrow) \right. \\ \left. \vee (((M_1, h_1) \mapsto^* (v_1, h'_1)) \wedge ((M_2, h_2) \mapsto^* (v_2, h'_2))) \right. \\ \left. \exists w' \sqsupseteq w. (h'_1, h'_2) \in w' \wedge (v_1, v_2) \in \mathcal{V} \llbracket \tau \rrbracket w' \right\}$$

# Kripke Logical Relations

Extension to languages with references.

- ↪ Need *worlds*  $w$ , i.e. invariants on heaps,
- ↪ Extensible on disjoint parts :  $w' \sqsupseteq w$
- ↪ Parametrize the definition of logical relations with such worlds.

$$\mathcal{E} \llbracket \tau \rrbracket w \stackrel{\text{def}}{=} \left\{ (M_1, M_2) \mid \forall (h_1, h_2) : w. ((M_1, h_1) \uparrow \wedge (M_2, h_2) \uparrow) \right. \\ \left. \vee (((M_1, h_1) \mapsto^* (v_1, h'_1)) \wedge ((M_2, h_2) \mapsto^* (v_2, h'_2))) \right. \\ \left. \exists w' \sqsupseteq w. (h'_1, h'_2) \in w' \wedge (v_1, v_2) \in \mathcal{V} \llbracket \tau \rrbracket w' \right\}$$

$$\mathcal{V} \llbracket \tau \rightarrow \sigma \rrbracket w \stackrel{\text{def}}{=} \left\{ (\lambda x_1. M_1, \lambda x_2. M_2) \mid \forall w' \sqsupseteq w. \forall (v_1, v_2) \in \mathcal{V} \llbracket \tau \rrbracket w'. \right. \\ \left. ((\lambda x_1. M_1)v_1, (\lambda x_2. M_2)v_2) \in \mathcal{E} \llbracket \sigma \rrbracket w' \right\}$$

# Equivalence of the Representation Independence example (I/II)

```
let c1 = ref0
    fun inc1() = c1 := !c1 + 1
    fun get1() = !c1
in ⟨inc1, get1⟩
```

---

```
let c2 = ref0
    fun inc2() = c2 := !c2 - 1
    fun get2() = -!c2
in ⟨inc2, get2⟩
```

# Equivalence of the Representation Independence example (I/II)

- Take  $w = \{(h_1, h_2) \mid h_1(c_1) = -h_2(c_2)\}$

# Equivalence of the Representation Independence example (I/II)

- Take  $w = \{(h_1, h_2) \mid h_1(c_1) = -h_2(c_2)\}$
- Prove  $(\text{inc}_1, \text{inc}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Unit} \rrbracket w$

# Equivalence of the Representation Independence example (I/II)

- Take  $w = \{(h_1, h_2) \mid h_1(c_1) = -h_2(c_2)\}$
- Prove  $(\text{inc}_1, \text{inc}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Unit} \rrbracket w$ 
  - $\rightsquigarrow$  Take  $(h_1, h_2) \in w$ ,

# Equivalence of the Representation Independence example (I/II)

- Take  $w = \{(h_1, h_2) \mid h_1(c_1) = -h_2(c_2)\}$
- Prove  $(\text{inc}_1, \text{inc}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Unit} \rrbracket w$ 
  - $\rightsquigarrow$  Take  $(h_1, h_2) \in w$ ,
  - $\rightsquigarrow (\text{inc}_1(), h_j) \mapsto^* ((), h'_j)$ ,

# Equivalence of the Representation Independence example (I/II)

- Take  $w = \{(h_1, h_2) \mid h_1(c_1) = -h_2(c_2)\}$
- Prove  $(\text{inc}_1, \text{inc}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Unit} \rrbracket w$ 
  - $\rightsquigarrow$  Take  $(h_1, h_2) \in w$ ,
  - $\rightsquigarrow (\text{inc}_i(), h_j) \mapsto^* ((), h'_j)$ ,
  - $\rightsquigarrow h'_1 = h_1[c \mapsto (h_1(c) + 1)]$  and  $h'_2 = h_2[c \mapsto (h_2(c) - 1)]$ ,

# Equivalence of the Representation Independence example (I/II)

- Take  $w = \{(h_1, h_2) \mid h_1(c_1) = -h_2(c_2)\}$
- Prove  $(\text{inc}_1, \text{inc}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Unit} \rrbracket w$ 
  - $\rightsquigarrow$  Take  $(h_1, h_2) \in w$ ,
  - $\rightsquigarrow (\text{inc}_i(), h_i) \mapsto^* ((), h'_i)$ ,
  - $\rightsquigarrow h'_1 = h_1[c \mapsto (h_1(c) + 1)]$  and  $h'_2 = h_2[c \mapsto (h_2(c) - 1)]$ ,
  - $\rightsquigarrow$  So that  $(h'_1, h'_2) \in w$  !

# Equivalence of the Representation Independence example (I/II)

- Take  $w = \{(h_1, h_2) \mid h_1(c_1) = -h_2(c_2)\}$
- Prove  $(\text{inc}_1, \text{inc}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Unit} \rrbracket w$ 
  - $\rightsquigarrow$  Take  $(h_1, h_2) \in w$ ,
  - $\rightsquigarrow (\text{inc}_i(), h_i) \mapsto^* ((), h'_i)$ ,
  - $\rightsquigarrow h'_1 = h_1[c \mapsto (h_1(c) + 1)]$  and  $h'_2 = h_2[c \mapsto (h_2(c) - 1)]$ ,
  - $\rightsquigarrow$  So that  $(h'_1, h'_2) \in w$  !
- Prove  $(\text{get}_1, \text{get}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Int} \rrbracket w$

# Equivalence of the Representation Independence example (I/II)

- Take  $w = \{(h_1, h_2) \mid h_1(c_1) = -h_2(c_2)\}$
- Prove  $(\text{inc}_1, \text{inc}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Unit} \rrbracket w$ 
  - $\rightsquigarrow$  Take  $(h_1, h_2) \in w$ ,
  - $\rightsquigarrow (\text{inc}_i(), h_i) \mapsto^* ((), h'_i)$ ,
  - $\rightsquigarrow h'_1 = h_1[c \mapsto (h_1(c) + 1)]$  and  $h'_2 = h_2[c \mapsto (h_2(c) - 1)]$ ,
  - $\rightsquigarrow$  So that  $(h'_1, h'_2) \in w$  !
- Prove  $(\text{get}_1, \text{get}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Int} \rrbracket w$ 
  - $\rightsquigarrow$  Take  $(h_1, h_2) \in w$ ,

# Equivalence of the Representation Independence example (I/II)

- Take  $w = \{(h_1, h_2) \mid h_1(c_1) = -h_2(c_2)\}$
- Prove  $(\text{inc}_1, \text{inc}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Unit} \rrbracket w$ 
  - $\rightsquigarrow$  Take  $(h_1, h_2) \in w$ ,
  - $\rightsquigarrow (\text{inc}_i(), h_i) \mapsto^* ((), h'_i)$ ,
  - $\rightsquigarrow h'_1 = h_1[c \mapsto (h_1(c) + 1)]$  and  $h'_2 = h_2[c \mapsto (h_2(c) - 1)]$ ,
  - $\rightsquigarrow$  So that  $(h'_1, h'_2) \in w$  !
- Prove  $(\text{get}_1, \text{get}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Int} \rrbracket w$ 
  - $\rightsquigarrow$  Take  $(h_1, h_2) \in w$ ,
  - $\rightsquigarrow (\text{get}_1(), h_1) \mapsto^* (h_1(c_1), h_1)$  and  $(\text{get}_2, h_2) \mapsto^* (-h_2(c_2), h_2)$ ,

# Equivalence of the Representation Independence example (I/II)

- Take  $w = \{(h_1, h_2) \mid h_1(c_1) = -h_2(c_2)\}$
- Prove  $(\text{inc}_1, \text{inc}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Unit} \rrbracket w$ 
  - ↪ Take  $(h_1, h_2) \in w$ ,
  - ↪  $(\text{inc}_1(), h_i) \mapsto^* ((), h'_i)$ ,
  - ↪  $h'_1 = h_1[c \mapsto (h_1(c) + 1)]$  and  $h'_2 = h_2[c \mapsto (h_2(c) - 1)]$ ,
  - ↪ So that  $(h'_1, h'_2) \in w$  !
- Prove  $(\text{get}_1, \text{get}_2) \in \mathcal{V} \llbracket \text{Unit} \rightarrow \text{Int} \rrbracket w$ 
  - ↪ Take  $(h_1, h_2) \in w$ ,
  - ↪  $(\text{get}_1(), h_1) \mapsto^* (h_1(c_1), h_1)$  and  $(\text{get}_2, h_2) \mapsto^* (-h_2(c_2), h_2)$ ,
  - ↪ Need to prove  $h_1(c_1) = -h_2(c_2)$ : Straightforward from  $(h_1, h_2) \in w$  !

## Invariants are not Enough

`let x = ref0 in fun awk1(f) = x := 1; f(); !x`  $\simeq$  `fun awk2(f) = f(); 1`

## Invariants are not Enough

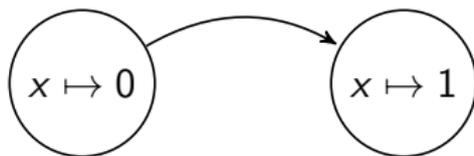
`let x = ref0 in fun awk1(f) = x := 1; f(); !x`  $\simeq$  `fun awk2(f) = f(); 1`

Need transition system of Invariants !

# Invariants are not Enough

`let x = ref0 in fun awk1(f) = x := 1; f(); !x`  $\simeq$  `fun awk2(f) = f(); 1`

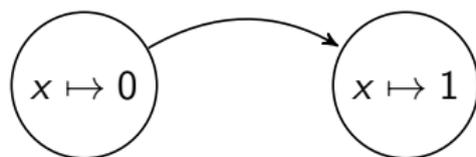
Need transition system of Invariants !



# Invariants are not Enough

`let x = ref0 in fun awk1(f) = x := 1; f(); !x`  $\simeq$  `fun awk2(f) = f(); 1`

Need transition system of Invariants !



- Future world  $w' \sqsupseteq w$  can either:
  - Add a new transition system of invariants on a disjoint part of the heaps,
  - Take a transition to get the following invariant.

## The Impact of Higher-Order State and Control Effects on Local Relational Reasoning

Derek Dreyer  
MPI-SWS  
dreyer@mpi-sws.org

Georg Neis  
MPI-SWS  
neis@mpi-sws.org

Lars Birkedal  
IT University of Copenhagen  
birkedal@itu.dk

### Abstract

Reasoning about program equivalence is one of the oldest problems in semantics. In recent years, useful techniques have been developed, based on bisimulations and logical relations, for reasoning about equivalence in the setting of increasingly realistic languages—languages nearly as complex as ML or Haskell. Much of the recent work in this direction has considered the interesting representation independence principles *enabled* by the use of local state, but it is also important to understand the principles that *powerful* features like higher-order state and control effects *disable*. This latter topic has been broached extensively within the framework of game semantics, resulting in what Abramsky dubbed the “semantic cube”: fully abstract game-semantic characterizations of various axes in the design space of ML-like languages. But when it comes to reasoning about many actual examples, game semantics does not yet supply a useful technique for proving equivalences.

In this paper, we marry the aspirations of the semantic cube to the powerful proof method of *step-indexed Kripke logical relations*. Building on recent work of Ahmed, Dreyer, and Rossberg, we define the first fully abstract logical relation for an ML-like language with recursive types, abstract types, general references and call/cc. We then show how, under orthogonal restrictions to the expressive power of our language—namely, the restriction to first-order state and/or the removal of call/cc—we can enhance the proving power of our possible-worlds model in correspondingly orthogonal ways, and we demonstrate this proving power on a range of interesting examples. Central to our story is the use of *state transition systems* to model the way in which properties of local state evolve over time.

### 1. Introduction

Reasoning about program equivalence is one of the oldest problems in semantics, with applications to program verification (“Is an optimized program equivalent to some reference implementation?”), compiler correctness (“Does a program transformation preserve the semantics of the source program?”), representation independence (“Can we modify the internal representation of an abstract data type without affecting the behavior of clients?”), and more besides.

The canonical notion of program equivalence for many applications is *observational* (or *contextual*) equivalence. Two programs are observationally equivalent if no program context can distinguish them by getting them to exhibit observably different input/output behavior. Reasoning about observational equivalence directly is difficult, due to the universal quantification over program contexts. Consequently, there has been a huge amount of work on developing useful models and logics for observational equivalence, and in recent years this line of work has scaled to handle increasingly realistic languages—languages nearly as complex as ML or Haskell, with features like general recursive types, general (higher-order) mutable references, and first-class continuations.

The focus of much of this recent work—*e.g.*, environmental bisimulations [36, 17, 32, 35], normal form bisimulations [34, 16], step-indexed Kripke logical relations [4, 2, 3]—has been on establishing some effective techniques for reasoning about programs that actually *use* the interesting, semantically complex features (state, continuations, etc.) of the languages being modeled. For instance, most of the work on languages with state concerns the various kinds of representation independence principles that arise due to the use of *local state* as an abstraction mechanism

# How to automate reasoning on Kripke Logical Relations

Obstacles to automate Kripke Logical Relations:

- Definition of  $\mathcal{V} \llbracket \tau \rightarrow \sigma \rrbracket$  quantify over pairs of values in  $\mathcal{V} \llbracket \tau \rrbracket$ ,
- Future worlds  $w' \sqsupseteq w$  can add arbitrary new systems of invariants.

Solutions:

# How to automate reasoning on Kripke Logical Relations

Obstacles to automate Kripke Logical Relations:

- Definition of  $\mathcal{V} \llbracket \tau \rightarrow \sigma \rrbracket$  quantify over pairs of values in  $\mathcal{V} \llbracket \tau \rrbracket$ ,
- Future worlds  $w' \sqsupseteq w$  can add arbitrary new systems of invariants.

Solutions:

- Reason on **open** terms with **functional names** (i.e. uninterpreted functions)
  - ↪ Make the full control flow apparent in the operational reduction,
  - ↪ Keep track of related functional names in an environment  $e$ ,

# How to automate reasoning on Kripke Logical Relations

Obstacles to automate Kripke Logical Relations:

- Definition of  $\mathcal{V} \llbracket \tau \rightarrow \sigma \rrbracket$  quantify over pairs of values in  $\mathcal{V} \llbracket \tau \rrbracket$ ,
- Future worlds  $w' \sqsupseteq w$  can add arbitrary new systems of invariants.

Solutions:

- Reason on **open** terms with **functional names** (i.e. uninterpreted functions)
  - ↪ Make the full control flow apparent in the operational reduction,
  - ↪ Keep track of related functional names in an environment  $e$ ,
- Fix a “World Transition System”  $\mathcal{A}$ :
  - ↪ Transitions specify only the evolution of worlds
  - ↪ Transition system represents the control flow between the term and its environment.

# How to automate reasoning on Kripke Logical Relations

Obstacles to automate Kripke Logical Relations:

- Definition of  $\mathcal{V} \llbracket \tau \rightarrow \sigma \rrbracket$  quantify over pairs of values in  $\mathcal{V} \llbracket \tau \rrbracket$ ,
- Future worlds  $w' \sqsupseteq w$  can add arbitrary new systems of invariants.

Solutions:

- Reason on **open** terms with **functional names** (i.e. uninterpreted functions)
  - ↪ Make the full control flow apparent in the operational reduction,
  - ↪ Keep track of related functional names in an environment  $e$ ,
- Fix a “World Transition System”  $\mathcal{A}$ :
  - ↪ Transitions specify only the evolution of worlds
  - ↪ Transition system represents the control flow between the term and its environment.

# How to automate reasoning on Kripke Logical Relations

Obstacles to automate Kripke Logical Relations:

- Definition of  $\mathcal{V} \llbracket \tau \rightarrow \sigma \rrbracket$  quantify over pairs of values in  $\mathcal{V} \llbracket \tau \rrbracket$ ,
- Future worlds  $w' \sqsupseteq w$  can add arbitrary new systems of invariants.

Solutions:

- Reason on **open** terms with **functional names** (i.e. uninterpreted functions)
  - ↪ Make the full control flow apparent in the operational reduction,
  - ↪ Keep track of related functional names in an environment  $e$ ,
- Fix a “World Transition System”  $\mathcal{A}$ :
  - ↪ Transitions specify only the evolution of worlds
  - ↪ Transition system represents the control flow between the term and its environment.

# How to automate reasoning on Kripke Logical Relations

Obstacles to automate Kripke Logical Relations:

- Definition of  $\mathcal{V} \llbracket \tau \rightarrow \sigma \rrbracket$  quantify over pairs of values in  $\mathcal{V} \llbracket \tau \rrbracket$ ,
- Future worlds  $w' \sqsupseteq w$  can add arbitrary new systems of invariants.

Solutions:

- Reason on **open** terms with **functional names** (i.e. uninterpreted functions)
  - ↪ Make the full control flow apparent in the operational reduction,
  - ↪ Keep track of related functional names in an environment  $e$ ,
- Fix a “World Transition System”  $\mathcal{A}$ :
  - ↪ Transitions specify only the evolution of worlds
  - ↪ Transition system represents the control flow between the term and its environment.

Give rise to our definition of “**Kripke Open Bisimulations**”  
(Joint Work with **N. Tabareau**)  
(Published at APLAS'15)

$(M_1, M_2) \in \mathcal{E}_{\mathcal{A}} \llbracket \tau \rrbracket_e w$  when, for all  $(h_1, h_2) \in w$ ,

$(M_1, M_2) \in \mathcal{E}_{\mathcal{A}} \llbracket \tau \rrbracket_e w$  when, for all  $(h_1, h_2) \in w$ ,

- Either both  $(M_1, h_1) \uparrow$  and  $(M_2, h_2) \uparrow$ ,

$(M_1, M_2) \in \mathcal{E}_{\mathcal{A}} \llbracket \tau \rrbracket_e w$  when, for all  $(h_1, h_2) \in w$ ,

- Either both  $(M_1, h_1) \uparrow$  and  $(M_2, h_2) \uparrow$ ,
- Or both  $(M_1, h_1) \mapsto^* (v_1, h'_1)$  and  $(M_2, h_2) \mapsto^* (v_2, h'_2)$  and there exists  $w' \sqsupseteq_{\mathcal{A}} w$  with  $(h'_1, h'_2) \in w'$  and  $(v_1, v_2) \in \mathcal{V}_{\mathcal{A}} \llbracket \tau \rrbracket_e w'$ ,

# Relations on Terms

$(M_1, M_2) \in \mathcal{E}_{\mathcal{A}} \llbracket \tau \rrbracket_e w$  when, for all  $(h_1, h_2) \in w$ ,

- Either both  $(M_1, h_1) \uparrow$  and  $(M_2, h_2) \uparrow$ ,
- Or both  $(M_1, h_1) \mapsto^* (v_1, h'_1)$  and  $(M_2, h_2) \mapsto^* (v_2, h'_2)$  and there exists  $w' \sqsupseteq_{\mathcal{A}} w$  with  $(h'_1, h'_2) \in w'$  and  $(v_1, v_2) \in \mathcal{V}_{\mathcal{A}} \llbracket \tau \rrbracket_e w'$ ,
- Or both  $(M_1, h_1) \mapsto^* (K_1[f_1 \ v_1], h'_1)$  and  $(M_2, h_2) \mapsto^* (K_2[f_2 \ v_2], h'_2)$  with  $(f_1, f_2, \sigma \rightarrow \sigma') \in e$  and there exists  $w' \sqsupseteq_{\mathcal{A}} w$  with  $(h'_1, h'_2) \in w'$  and  $(K_1, K_2) \in \mathcal{K}_{\mathcal{A}} \llbracket \sigma', \tau \rrbracket_e w'$ ,  $(v_1, v_2) \in \mathcal{V}_{\mathcal{A}} \llbracket \sigma \rrbracket_e w'$ ,

$(M_1, M_2) \in \mathcal{E}_{\mathcal{A}} \llbracket \tau \rrbracket_e w$  when, for all  $(h_1, h_2) \in w$ ,

- Either both  $(M_1, h_1) \uparrow$  and  $(M_2, h_2) \uparrow$ ,
- Or both  $(M_1, h_1) \mapsto^* (v_1, h'_1)$  and  $(M_2, h_2) \mapsto^* (v_2, h'_2)$  and there exists  $w' \sqsupseteq_{\mathcal{A}} w$  with  $(h'_1, h'_2) \in w'$  and  $(v_1, v_2) \in \mathcal{V}_{\mathcal{A}} \llbracket \tau \rrbracket_e w'$ ,
- Or both  $(M_1, h_1) \mapsto^* (K_1[f_1 \ v_1], h'_1)$  and  $(M_2, h_2) \mapsto^* (K_2[f_2 \ v_2], h'_2)$  with  $(f_1, f_2, \sigma \rightarrow \sigma') \in e$  and there exists  $w' \sqsupseteq_{\mathcal{A}} w$  with  $(h'_1, h'_2) \in w'$  and  $(K_1, K_2) \in \mathcal{K}_{\mathcal{A}} \llbracket \sigma', \tau \rrbracket_e w'$ ,  $(v_1, v_2) \in \mathcal{V}_{\mathcal{A}} \llbracket \sigma \rrbracket_e w'$ ,
- (+ deferred divergence).

## Theorem

Two terms  $M_1, M_2$  of type  $\tau$  are contextually equivalent iff there exists a WTS  $\mathcal{A}$  s.t.  $(M_1, M_2) \in \mathcal{E}_{\mathcal{A}} \llbracket \tau \rrbracket w_0$ .

- Proof: Correspondence with “simple” bisimulations on the LTS generating traces.
- Soundness:
  - ↪ WTS  $\mathcal{A}$  as an abstraction of the LTS generating traces.
- Completeness:
  - ↪ Not for free: no more biorthogonality,
  - ↪ Need *exhaustive* WTS,
  - ↪ First complete method which does not rely on closure in its definition.

- $\mathcal{E}_{\mathcal{A}} \llbracket \tau \rrbracket_e w$  uses operational reduction
  - ↪ Replace it with symbolic execution,
  - ↪ Generate arithmetic constraints,
  - ↪ Block on callbacks and recursive calls.
- Quantification over future worlds
  - ↪ Abstract over it using temporal logic,
  - ↪ Temporal modalities  $\Box\phi, X(\phi)$ .
- Give rise to Temporal Symbolic Kripke Open Bisimulations  $\mathbb{E} \llbracket \tau \rrbracket$ .
  - ↪ Model check  $\mathcal{A} \models \mathbb{E} \llbracket \tau \rrbracket$ .

## An Example: Callback with lock

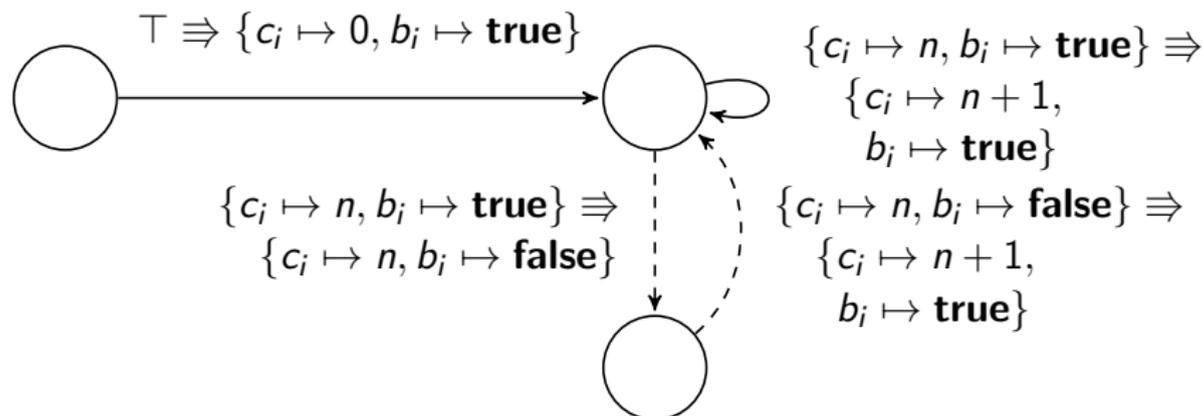
```
let b1 = ref true, c1 = ref 0,  
    fun inc1(f) = if !b1 then {  
        b1 := false;  
        f(); c1 := !c1 + 1;  
        b1 := true } else (),  
    fun get1() = !c1  
in ⟨inc1, get1⟩
```

```
let b2 = ref true, c2 = ref 0,  
    fun inc2(f) = if !b2 then {  
        b2 := false;  
        let n = !c2 in f(); c2 := n + 1;  
        b2 := true } else (),  
    fun get2() = !c2  
in ⟨inc2, get2⟩
```

## An Example: Callback with lock

```
let  $b_1 = \text{ref } \mathbf{true}$ ,  $c_1 = \text{ref } 0$ ,  
  fun  $\text{inc}_1(f) = \text{if } !b_1 \text{ then } \left\{ \begin{array}{l} b_1 := \mathbf{false}; \\ f(); c_1 := !c_1 + 1; \\ b_1 := \mathbf{true} \end{array} \right\} \text{ else } ()$ ,  
  fun  $\text{get}_1() = !c_1$   
in  $\langle \text{inc}_1, \text{get}_1 \rangle$ 
```

```
let  $b_2 = \text{ref } \mathbf{true}$ ,  $c_2 = \text{ref } 0$ ,  
  fun  $\text{inc}_2(f) = \text{if } !b_2 \text{ then } \left\{ \begin{array}{l} b_2 := \mathbf{false}; \\ \text{let } n = !c_2 \text{ in } f(); c_2 := n + 1; \\ b_2 := \mathbf{true} \end{array} \right\} \text{ else } ()$ ,  
  fun  $\text{get}_2() = !c_2$   
in  $\langle \text{inc}_2, \text{get}_2 \rangle$ 
```



# Automatically generated temporal formula

$\mathbb{E}[\llbracket \tau \rrbracket](M_1^{clb}, M_2^{cbl})$  is equal to

$$\begin{aligned} & \mathcal{H}N_0.(\mathcal{N}_1\ell_2.(\mathcal{N}_1\ell_1.(\mathcal{N}_2\ell_4.(\mathcal{N}_2\ell_3.(\mathbf{X}((\ell_2 \mapsto_1 \mathbf{0}) \wedge (\ell_1 \mapsto_1 \mathbf{true}) \wedge (\ell_4 \mapsto_2 \mathbf{0}) \wedge (\ell_3 \mapsto_2 \mathbf{true}) \wedge \\ & ((\square(\mathcal{H}N_5.(\forall x_6, x_7, x_8, x_9.(((\ell_2 \mapsto_1 x_6) \wedge (\ell_1 \mapsto_1 x_7) \wedge (\ell_4 \mapsto_2 x_8) \wedge (\ell_3 \mapsto_2 x_9)) \Rightarrow \\ & ((\mathbf{X}(((x_7 = \mathbf{true}) \wedge (x_9 = \mathbf{true})) \Rightarrow ((\ell_2 \mapsto_1 x_6) \wedge (\ell_1 \mapsto_1 \mathbf{false}) \wedge (\ell_4 \mapsto_2 x_8) \wedge (\ell_3 \mapsto_2 \mathbf{false})) \wedge \\ & (\square_{\text{pub}}(\forall x_{10}, x_{11}, x_{13}, x_{14}.(((\ell_2 \mapsto_1 x_{10}) \wedge (\ell_1 \mapsto_1 x_{11}) \wedge (\ell_4 \mapsto_2 x_{13}) \wedge (\ell_3 \mapsto_2 x_{14})) \Rightarrow \\ & (\mathbf{X}(\forall x_{12}, x_{15}.(((x_{12} = x_{10} + 1) \wedge (x_{15} = x_8 + 1)) \Rightarrow ((\ell_2 \mapsto_1 x_{12}) \wedge (\ell_1 \mapsto_1 \mathbf{true}) \wedge \\ & (\ell_4 \mapsto_2 x_{15}) \wedge (\ell_3 \mapsto_2 \mathbf{true}) \wedge (\mathbf{P}_{\text{pub}}(N_5)))))))))) \wedge (\mathbf{not}((x_7 = \mathbf{true}) \wedge (x_9 = \mathbf{false}))) \wedge \\ & (\mathbf{not}((x_7 = \mathbf{false}) \wedge (x_9 = \mathbf{true}))) \wedge (\mathbf{X}(((x_7 = \mathbf{false}) \wedge (x_9 = \mathbf{false})) \Rightarrow \\ & ((\ell_2 \mapsto_1 x_6) \wedge (\ell_1 \mapsto_1 x_7) \wedge (\ell_4 \mapsto_2 x_8) \wedge (\ell_3 \mapsto_2 x_9) \wedge (\mathbf{P}_{\text{pub}}(N_5)))))))))) \\ & \wedge (\square(\mathcal{H}N_{16}.(\forall [x_{17}, x_{18}, x_{19}, x_{20}].(((\ell_2 \mapsto_1 x_{17}) \wedge (\ell_1 \mapsto_1 x_{18}) \wedge (\ell_4 \mapsto_2 x_{19}) \wedge (\ell_3 \mapsto_2 x_{20})) \Rightarrow \\ & (\mathbf{X}((\ell_2 \mapsto_1 x_{17}) \wedge (\ell_1 \mapsto_1 x_{18}) \wedge (\ell_4 \mapsto_2 x_{19}) \wedge (\ell_3 \mapsto_2 x_{20}) \wedge (x_{17} = x_{19}) \wedge (\mathbf{P}_{\text{pub}}(N_{16})))))))))) \\ & \wedge (\mathbf{P}_{\text{pub}}(N_0)))))) \end{aligned}$$

# Translation to SMT-LIB

```
(assert (exists ((s21 Int)(h22 Heap)(h23 Heap)(l2 Int)(l1 Int)) (and (not (= l1 l2))
(exists ((l4 Int)(l3 Int)) (and (not (= l3 l4)) (exists ((s25 Int)(h26 Heap)(h27 Heap)(S28 LocSpan))
(and (TransPriv s21 s25 h22 h23 h26 h27 S28) (and (= (select h26 l2) 0) (= (select h26 l1) 0)
(= (select h27 l4) 0) (= (select h27 l3) 0) (and (forall ((s29 Int)(h30 Heap)(h31 Heap)(S32 LocSpan))
(=> (TransPrivT s25 s29 h26 h27 h30 h31 S32) (forall ((x6 Int)(x7 Int)(x8 Int)(x9 Int))
(=> (and (= (select h30 l2) x6) (= (select h30 l1) x7) (= (select h31 l4) x8) (= (select h31 l3) x9) )
(and (exists ((s33 Int)(h34 Heap)(h35 Heap)(S36 LocSpan)) (and (TransPriv s29 s33 h30 h31 h34 h35 S36)
(=> (and (= x7 0) (= x9 0) ) (and (= (select h34 l2) x6) (= (select h34 l1) 1) (= (select h35 l4) x8)
(= (select h35 l3) 1) (forall ((s37 Int)(h38 Heap)(h39 Heap)(S40 LocSpan))
(=> (TransPubT s33 s37 h34 h35 h38 h39 S40) (forall ((x10 Int)(x11 Int)(x13 Int)(x14 Int))
(=> (and (= (select h38 l2) x10) (= (select h38 l1) x11) (= (select h39 l4) x13) (= (select h39 l3) x14))
(exists ((s41 Int)(h42 Heap)(h43 Heap)(S44 LocSpan)) (and (TransPriv s37 s41 h38 h39 h42 h43 S44)
(forall ((x12 Int)(x15 Int)) (=> (and (= x12 (+ x10 1)) (= x15 (+ x8 1)) ) (and (= (select h42 l2) x12)
(= (select h42 l1) 0) (= (select h43 l4) x15) (= (select h43 l3) 0)
(TransPub s29 s41 h30 h31 h42 h43 S44))))))))))))))
(not (and (= x7 0) (= x9 1) )) (not (and (= x7 1) (= x9 0) ))
(exists ((s45 Int)(h46 Heap)(h47 Heap)(S48 LocSpan)) (and (TransPriv s29 s45 h30 h31 h46 h47 S48)
(=> (and (= x7 1) (= x9 1) ) (and (= (select h46 l2) x6) (= (select h46 l1) x7) (= (select h47 l4) x8)
(= (select h47 l3) x9) (TransPub s29 s45 h30 h31 h46 h47 S48))))))))))
(forall ((s49 Int)(h50 Heap)(h51 Heap)(S52 LocSpan))(=> (TransPrivT s25 s49 h26 h27 h50 h51 S52)
(forall ((x17 Int)(x18 Int)(x19 Int)(x20 Int)) (=> (and (= (select h50 l2) x17) (= (select h50 l1) x18)
(= (select h51 l4) x19) (= (select h51 l3) x20) ) (exists ((s53 Int)(h54 Heap)(h55 Heap)(S56 LocSpan))
(and (TransPriv s49 s53 h50 h51 h54 h55 S56) (and (= (select h54 l2) x17) (= (select h54 l1) x18)
(= (select h55 l4) x19) (= (select h55 l3) x20) (= x17 x19) (TransPub s49 s53 h50 h51 h54 h55 S56))))))))))
(TransPub s21 s25 h22 h23 h26 h27 S28)))))))))

(check-sat)
```

- Combine Symbolic, Temporal with Circular reasoning
  - ↪ Deal with recursive functions,
  - ↪ Synchronization of recursive calls,
  - ↪ Simple heuristics to unfold recursive functions,
  - ↪ Circular proof system.
  
- Computation of the transitive closure of the WTS  $\mathcal{A}$ 
  - ↪ Undecidable in general,
  - ↪ Over-approximations via abstractions.
  
- Automatic Computation of  $\mathcal{A}$ .