

The Marriage of Bisimulations and Kripke Logical Relations

Chung-Kil Hur Derek Dreyer
Georg Neis Viktor Vafeiadis

Max Planck Institute for Software Systems (MPI-SWS)
Kaiserslautern and Saarbrücken, Germany

POPL 2012
Philadelphia, USA

Canonical definition: Contextual equivalence

- Observable equivalence under an arbitrary context
- **Hard to reason about**, due to the quantification over arbitrary contexts

Various methods developed for local reasoning

- **Bisimulations** and **Kripke Logical Relations (KLRs)**
- Handle higher-order functions, abstract types, recursive types, general references, exceptions, continuations, etc.

Motivation #1: Marrying complementary approaches

KLRs' treatment of **local state** is **more powerful**.

- **Transition systems** for controlling evolution of state.
- Subsumes the power of **environmental bisimulations**.

Bisimulations' treatment of **recursion** is **cleaner**.

- **Coinduction** simpler and more direct than **step-indexing**.

Can we join them together in a single method?

Goal: compositional equivalences between programs
in different languages

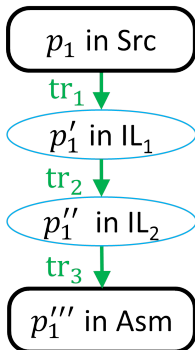
- e.g., compositional certified compilation

Motivation #2: Inter-language reasoning

Goal: compositional equivalences between programs in different languages

- e.g., compositional certified compilation

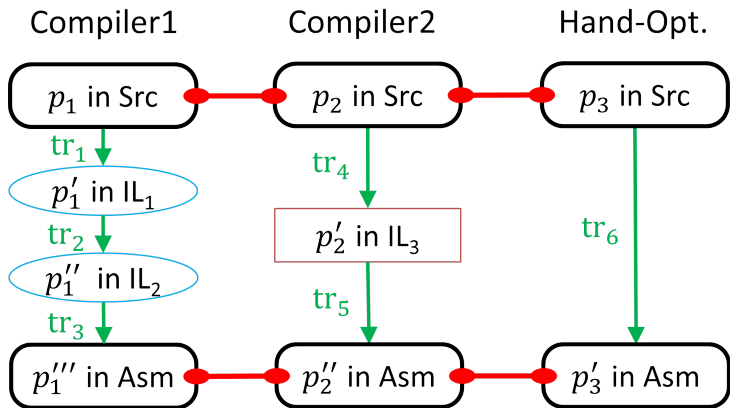
Compiler



Motivation #2: Inter-language reasoning

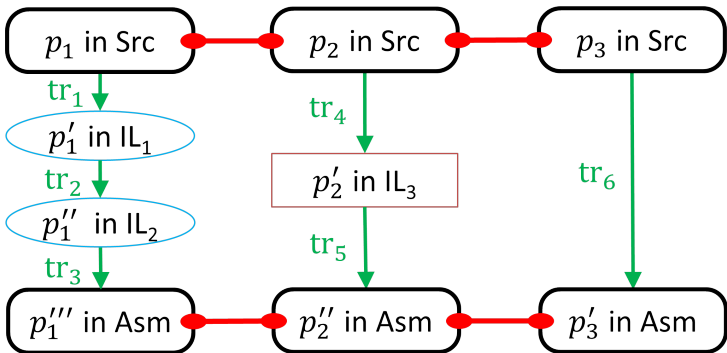
Goal: compositional equivalences between programs in different languages

- e.g., compositional certified compilation



Motivation #2: Inter-language reasoning

- **Horizontal compositionality** is preservation of equivalence under **linking of modules**.
- **Vertical compositionality** is **transitive composition** of equivalence proofs.



Motivation #2: Inter-language reasoning

KLRs are **not transitively composable**

- Due to their use of “**step-indexing**” for recursive features
- Hur et al. [ICFP09, POPL11] only studied one-pass compilers

Bisim's do **not scale** (in an obvious way) to **inter-language** reasoning

- Due to their use of “**syntactic**” **devices** for H-O functions

Can we remove these limitations?

A new method for local relational reasoning:

Relation Transition Systems (RTSs)

- Combines the “most appealing” features of KLRs and bisimulations
- Potential to scale to inter-language reasoning
 - Does not rely on syntactic devices for H-O functions
 - Supports transitive composition of equivalence proofs

A new method for local relational reasoning:

Key idea

Don't just **support** local reasoning. **Demand** it!

- **Supports transitive composition** of equivalence proofs

Existing methods support local reasoning
but don't demand it

- There's nothing preventing one from sneaking a “brute-force” proof in through the back door

Our method will demand **strictly local reasoning**

- Brute-force proofs will not be permitted!

Benefit of our approach: **More compositionality**

$$\tau \in \mathbf{Type} ::= \alpha \mid \tau_{\text{base}} \mid \tau_1 \rightarrow \tau_2 \mid \tau_1 \times \tau_2 \mid \mu\alpha. \tau$$

$$v_1 \approx v_2 : \mathcal{T} \stackrel{\text{def}}{=} \text{---}$$

If you want to prove v_1 equivalent to v_2 ,

Coinductive approach (similar to bisimulations)

$$v_1 \approx v_2 : \tau \stackrel{\text{def}}{=} \exists \sim_L. v_1 \sim_L v_2 : \tau$$

If you want to prove v_1 equivalent to v_2 ,

- 1 Find a “local knowledge” \sim_L relating v_1 and v_2

$$v_1 \approx v_2 : \tau \stackrel{\text{def}}{=} \exists \sim_L. v_1 \sim_L v_2 : \tau \\ \wedge \textit{consistent}(\sim_L)$$

If you want to prove v_1 equivalent to v_2 ,

- 1 Find a “local knowledge” \sim_L relating v_1 and v_2
- 2 Show that \sim_L is consistent

$$v_1 \approx v_2 : \tau \stackrel{\text{def}}{=} \exists \boxed{\sim_L}. v_1 \sim_L v_2 : \tau \wedge \textit{consistent}(\sim_L)$$

If you want to prove v_1 equivalent to v_2 ,

- 1 Find a “local knowledge” \sim_L relating v_1 and v_2
- 2 Show that \sim_L is consistent

- 1 Restrict \sim_L to only function types
- 2 Derive $\overline{\sim_L}$ from \sim_L by induction

$$\frac{f_1 \sim_L f_2 : \sigma \rightarrow \tau}{f_1 \overline{\sim_L} f_2 : \sigma \rightarrow \tau} \qquad \frac{c \in \llbracket \mathcal{T}_{\text{base}} \rrbracket}{c \overline{\sim_L} c : \mathcal{T}_{\text{base}}}$$

$$\frac{v_1 \overline{\sim_L} v_2 : \tau \quad v'_1 \overline{\sim_L} v'_2 : \tau'}{\langle v_1, v'_1 \rangle \overline{\sim_L} \langle v_2, v'_2 \rangle : \tau \times \tau'}$$

$$\frac{v_1 \overline{\sim_L} v_2 : \tau[\mu\alpha. \tau/\alpha]}{\text{roll } v_1 \overline{\sim_L} \text{roll } v_2 : \mu\alpha. \tau}$$

$$v_1 \approx v_2 : \tau \stackrel{\text{def}}{=} \exists \sim_L . v_1 \sim_L v_2 : \tau \\ \wedge \textit{consistent}(\sim_L)$$

If you want to prove v_1 equivalent to v_2 ,

- 1 Find a “local knowledge” \sim_L relating v_1 and v_2
- 2 Show that \sim_L is consistent

$$v_1 \approx v_2 : \tau \stackrel{\text{def}}{=} \exists \sim_L . v_1 \overline{\sim_L} v_2 : \tau \\ \wedge \textit{consistent}(\sim_L)$$

If you want to prove v_1 equivalent to v_2 ,

- 1 Find a “local knowledge” \sim_L relating v_1 and v_2
- 2 Show that \sim_L is consistent

Coinductive approach (similar to bisimulations)

$$v_1 \approx v_2 : \tau \stackrel{\text{def}}{=} \exists \sim_L . v_1 \overline{\sim_L} v_2 : \tau \wedge \boxed{\text{consistent}(\sim_L)}$$

If you want to prove v_1 equivalent to v_2 ,

- 1 Find a “local knowledge” \sim_L relating v_1 and v_2
- 2 Show that \sim_L is consistent

Definition of *consistent*(\sim_L)

$$\lambda x. e_1 \sim_L \lambda x. e_2 : \sigma \rightarrow \tau$$

$$\implies$$

$$\forall v_1 \overset{?}{\sim}_1 v_2 : \sigma.$$

$$e_1[v_1/x] \overset{?}{\sim}_2 e_2[v_2/x] : \tau$$

What should $\sim_1^?$ be?

$$\lambda x. e_1 \sim_L \lambda x. e_2 : \sigma \rightarrow \tau$$

$$\implies$$

$$\forall v_1 \sim_1^? v_2 : \sigma.$$

$$e_1[v_1/x] \sim_2^? e_2[v_2/x] : \tau$$

$$\sim_1^? = \overline{\sim_L} : \text{Unsound}$$

Because v_1, v_2 come from the context

$$\lambda x. e_1 \sim_L \lambda x. e_2 : \sigma \rightarrow \tau$$

$$\implies$$

$$\forall v_1 \sim_1^? v_2 : \sigma.$$

$$e_1[v_1/x] \sim_2^? e_2[v_2/x] : \tau$$

$\sim_1^?$ should be a **global** notion of equivalence $\overline{\sim_G}$

$$\lambda x. e_1 \sim_L \lambda x. e_2 : \sigma \rightarrow \tau$$

$$\implies$$

$$\forall v_1 \overline{\sim_G} v_2 : \sigma.$$

$$e_1[v_1/x] \sim_2^? e_2[v_2/x] : \tau$$

Intuition: Global vs. local knowledge

\sim_L represents local knowledge

- Functions **our proof/module** says are equivalent

\sim_G represents global knowledge

- Functions **the whole program** says are equivalent

Intuition: Global vs. local knowledge

\sim_L represents **local** knowledge

- Functions **our proof/module** says are equivalent

\sim_G represents **global** knowledge

- Functions **the whole program** says are equivalent

Defining \sim_G “**semantically**” is hard!

- It's as hard as the original problem of finding a good relational model of ML!

So existing H-O bisimulations all define \sim_G as some variation on **syntactic identity**

- Applicative, environmental, normal form bisim's

What is \sim_G ?

What is \sim_G ? Who cares?

Idea: **Parameterize** our whole model over \sim_G !

- We will make *some* assumptions about it ($\sim_G \supseteq \sim_L$), but \sim_G may relate **any two values** at function type.
- \sim_G can even contain “junk” like ($4 \sim_G \text{true} : \text{int} \rightarrow \text{int}$)!
- Highly reminiscent of the Girard/Reynolds method for reasoning about parametricity of ADTs

What is ...? Who cares?

Takehome #1

- **Girard/Reynolds**: Clients of ADT are parametric w.r.t. relational interpretation of **abstract types**
- **Our method**: Equivalence proofs are parametric w.r.t. relational interpretation of **function types**

Instead of defining $\sim_G \dots$

$$\lambda x. e_1 \sim_L \lambda x. e_2 : \sigma \rightarrow \tau$$

$$\implies$$

$$\forall v_1 \overline{\sim_G} v_2 : \sigma.$$

$$e_1[v_1/x] \overset{?}{\sim}_2 e_2[v_2/x] : \tau$$

Definition of $\text{consistent}(\sim_L)$

... we parameterize over $\sim_G!$

$$\lambda x. e_1 \sim_L \lambda x. e_2 : \sigma \rightarrow \tau$$

$$\implies$$

$$\boxed{\forall \sim_G \supseteq \sim_L. \forall v_1 \overline{\sim_G} v_2 : \sigma.}$$

$$e_1[v_1/x] \overset{?}{\sim}_2 e_2[v_2/x] : \tau$$

What should $\sim_2^?$ be?

$$\lambda x. e_1 \sim_L \lambda x. e_2 : \sigma \rightarrow \tau$$

$$\implies$$

$$\boxed{\forall \sim_G \supseteq \sim_L. \forall v_1 \overline{\sim_G} v_2 : \sigma.}$$

$$e_1[v_1/x] \sim_2^? e_2[v_2/x] : \tau$$

Both diverge or both converge to related values?

$$\lambda x. e_1 \sim_L \lambda x. e_2 : \sigma \rightarrow \tau$$

$$\implies$$

$$\boxed{\forall \sim_G \supseteq \sim_L. \forall v_1 \overline{\sim_G} v_2 : \sigma.}$$

$$e_1[v_1/x] \overset{?}{\sim}_2 e_2[v_2/x] : \tau$$

Both diverge or both converge to related values?

$$\lambda f. f(0) \sim_L \lambda f. f(0) : (\text{int} \rightarrow \text{int}) \rightarrow \text{int}$$

$$\implies$$

$$\boxed{\forall \sim_G \supseteq \sim_L. \forall v_1 \overline{\sim}_G v_2 : \text{int} \rightarrow \text{int}.}$$

$$v_1(0) \overset{?}{\sim}_2 v_2(0) : \text{int}$$

Both diverge or both converge to related values?

$$\lambda f. f(0) \sim_L \lambda f. f(0) : (\text{int} \rightarrow \text{int}) \rightarrow \text{int}$$



$$\boxed{\forall \sim_G \supseteq \sim_L. \lambda \text{true} : \text{int} \rightarrow \text{int}.}$$

$$\lambda \text{true}(0) \stackrel{?}{\sim}_2 \lambda \text{true}(0) : \text{int}$$

$\sim_2^?$ should be “local term equivalence” \sim_G^{exp}

$$\lambda x. e_1 \sim_L \lambda x. e_2 : \sigma \rightarrow \tau$$

$$\implies$$

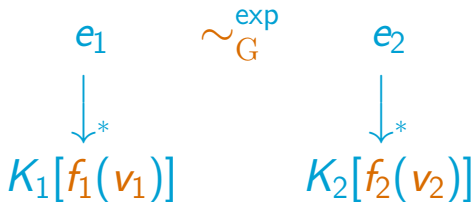
$$\boxed{\forall \sim_G \supseteq \sim_L. \forall v_1 \overline{\sim}_G v_2 : \sigma.}$$

$$e_1[v_1/x] \sim_G^{\text{exp}} e_2[v_2/x] : \tau$$

To show your terms are locally equivalent

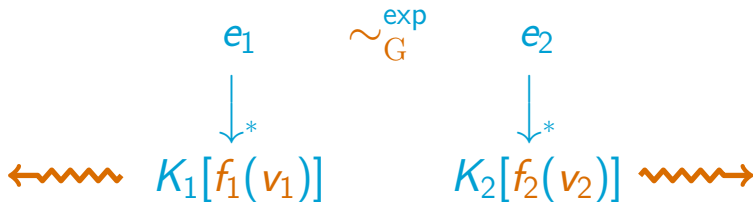
$$e_1 \quad \underset{G}{\sim}^{\text{exp}} \quad e_2$$

Execute them “locally” until...

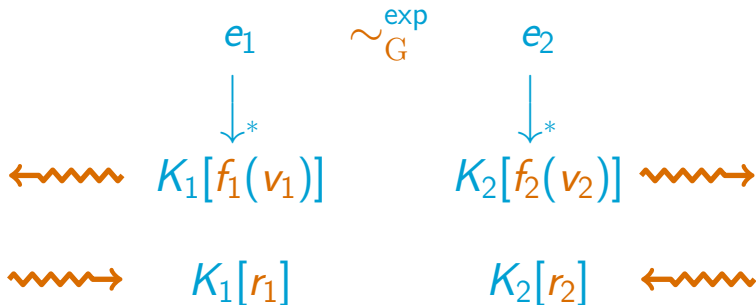


Intuition: Local term equivalence

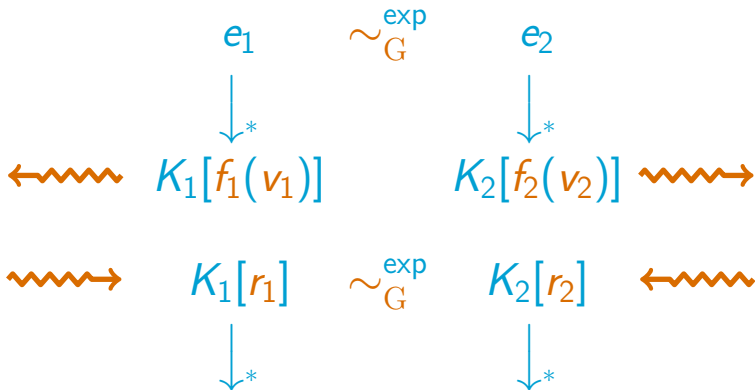
... they pass control to “external” functions



Assume you get back control
with related return values



Show your continuations are locally equivalent



Definition of local term equivalence \sim_G^{exp}

- Derive \sim_G^{exp} from \sim_G by coinduction

$$e_1 \quad \sim_G^{\text{exp}} \quad e_2 \quad : \mathcal{T}$$

- Derive \sim_G^{exp} from \sim_G by coinduction

$$\begin{array}{ccc}
 e_1 & \sim_G^{\text{exp}} & e_2 \\
 \downarrow \omega & & \downarrow \omega \\
 & &
 \end{array}
 : \mathcal{T}$$

Case 1: Both diverge

- Derive \sim_G^{exp} from \sim_G by coinduction

$$\begin{array}{ccccc}
 e_1 & \sim_G^{\text{exp}} & e_2 & : \mathcal{T} \\
 \downarrow^* & & \downarrow^* & \\
 v_1 & \overline{\sim_G} & v_2 & : \mathcal{T}
 \end{array}$$

Case 2: Both terminate

Definition of local term equivalence \sim_G^{exp}

- Derive \sim_G^{exp} from \sim_G by coinduction

$$\begin{array}{ccc} e_1 & \sim_G^{\text{exp}} & e_2 & : \tau \\ \downarrow^* & & \downarrow^* & \\ K_1[f_1(v_1)] & & K_2[f_2(v_2)] & : \tau \end{array}$$

- $f_1 \sim_G f_2 : \tau_{\text{arg}} \rightarrow \tau_{\text{ret}}$
- $v_1 \overline{\sim}_G v_2 : \tau_{\text{arg}}$
- $\forall r_1 \overline{\sim}_G r_2 : \tau_{\text{ret}}. K_1[r_1] \sim_G^{\text{exp}} K_2[r_2] : \tau$

Case 3: Both call a function

Definition of local term equivalence \sim_G^{exp}

- Derive \sim_G^{exp} from \sim_G by coinduction

$$\begin{array}{ccc}
 e_1 & \sim_G^{\text{exp}} & e_2 & : \tau \\
 \downarrow^* & & \downarrow^* & \\
 K_1[f_1(v_1)] & & K_2[f_2(v_2)] & : \tau
 \end{array}$$

$$\textcircled{1} \quad f_1 \sim_G f_2 : \tau_{\text{arg}} \rightarrow \tau_{\text{ret}}$$

$$\textcircled{2} \quad v_1 \overline{\sim}_G v_2 : \tau_{\text{arg}}$$

$$\textcircled{3} \quad \forall r_1 \overline{\sim}_G r_2 : \tau_{\text{ret}}. K_1[r_1] \sim_G^{\text{exp}} K_2[r_2] : \tau$$

Case 3: Both call a function

Definition of local term equivalence \sim_G^{exp}

- Derive \sim_G^{exp} from \sim_G by coinduction

$$\begin{array}{ccc}
 e_1 & \sim_G^{\text{exp}} & e_2 & : \tau \\
 \downarrow^* & & \downarrow^* & \\
 K_1[f_1(\mathbf{v}_1)] & & K_2[f_2(\mathbf{v}_2)] & : \tau
 \end{array}$$

① $f_1 \sim_G f_2 : \tau_{\text{arg}} \rightarrow \tau_{\text{ret}}$

② $\boxed{v_1 \sim_G v_2 : \tau_{\text{arg}}}$

③ $\forall r_1 \sim_G r_2 : \tau_{\text{ret}}. K_1[r_1] \sim_G^{\text{exp}} K_2[r_2] : \tau$

Case 3: Both call a function

Definition of local term equivalence \sim_G^{exp}

- Derive \sim_G^{exp} from \sim_G by coinduction

$$\begin{array}{ccc}
 e_1 & \sim_G^{\text{exp}} & e_2 & : \mathcal{T} \\
 \downarrow^* & & \downarrow^* & \\
 K_1[f_1(v_1)] & & K_2[f_2(v_2)] & : \mathcal{T}
 \end{array}$$

① $f_1 \sim_G f_2 : \tau_{\text{arg}} \rightarrow \tau_{\text{ret}}$

② $v_1 \overline{\sim}_G v_2 : \tau_{\text{arg}}$

③ $\forall r_1 \overline{\sim}_G r_2 : \tau_{\text{ret}}. K_1[r_1] \sim_G^{\text{exp}} K_2[r_2] : \mathcal{T}$

Case 3: Both call a function

Definition of local term equivalence \sim_G^{exp}

- Derive \sim_G^{exp} from \sim_G by coinduction

$$\begin{array}{ccc}
 e_1 & \sim_G^{\text{exp}} & e_2 & : \mathcal{T} \\
 \downarrow^* & & \downarrow^* & \\
 K_1[f_1(v_1)] & & K_2[f_2(v_2)] & : \mathcal{T}
 \end{array}$$

① $f_1 \sim_G f_2 : \mathcal{T}_{\text{arg}} \rightarrow \mathcal{T}_{\text{ret}}$

② $v_1 \overline{\sim}_G v_2 : \mathcal{T}_{\text{arg}}$

③ $\forall r_1 \overline{\sim}_G r_2 : \mathcal{T}_{\text{ret}}. K_1[r_1] \sim_G^{\text{exp}} K_2[r_2] : \mathcal{T}$

Case 3: Both call a function

- Derive \sim_G^{exp} from \sim_G by coinduction

Takehome #2

Since our proofs are **parametric** w.r.t. \sim_G ,
we **CAN** and we **MUST** reason locally!

Case 3: Both call a function

Summary: The benefits of “proof parametricity”

- 1 Horizontal compositionality (aka congruence)
 - The less proofs about different modules assume about \sim_G , the easier they are to link together
- 2 Vertical compositionality (aka transitivity)
 - Since equivalence proofs must use “local” reasoning, their structure is highly constrained, making them easier to compose transitively

“Normal form” (or “open”) bisimulations

- Related fn arguments represented by a fresh variable x
- Hence, bisimulation must account for **terms getting stuck**
- Definition very similar to our “local term equivalence”

Mendler-style coinduction

- L is a “robustly post-fixed point (rpfp)” of an endofunction F if $\forall G \geq L. L \leq F(G)$
- Rpfp’s are closed under joins even for **non-monotone** F
- Our consistency condition is a variant of this

What else is in the paper?

- Generalization to open terms
 - Requires parameterizing \sim_L over \sim_G
- Extension of model with abstract types
 - Based on [Sumii-Pierce '05]
- Extension of model with higher-order state
 - Based on [Dreyer-Neis-Birkedal '10]
- Transitivity proved for pure fragment
 - Proof for full model currently under submission
- All results mechanized in Coq
- Future work:
 - Inter-language reasoning (certified ML/C compilers with FFI)
 - Supporting refined type system (e.g., effect system)
 - Supporting concurrency