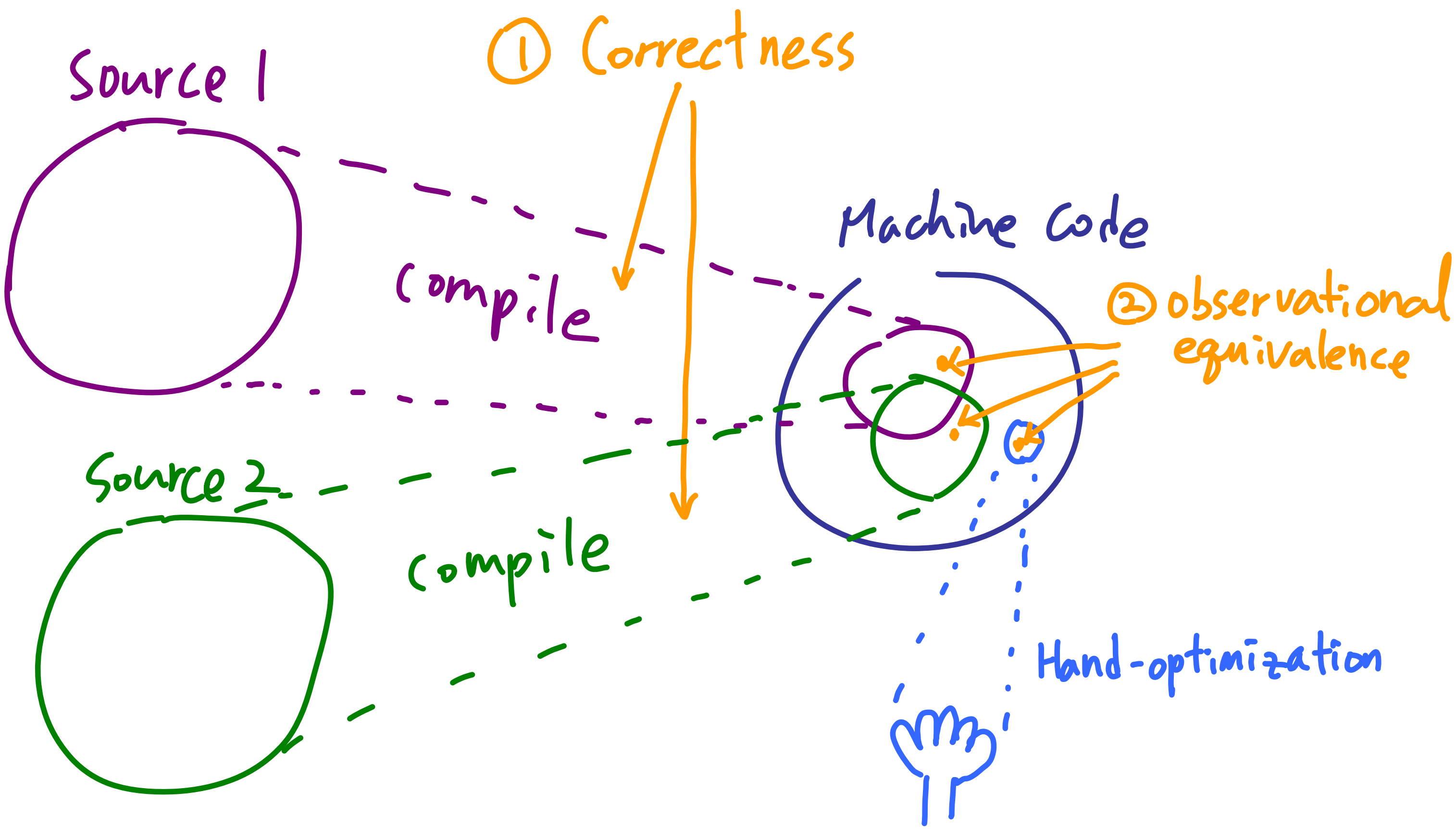
Compiler Correctness & Observational Equivalence on Machine code

Chung-Kil Hur
Jointwork with Nick Benton

16th Mar. 2009
@ University of Cambridge

Motivation: Overview



Motivation : Compiler Correctness

Compiler $\langle\!\langle \cdot \rangle\!\rangle$: Source $\to$ Machine code

Compiler correctness

$\forall\, t : T_1,\ t_1 : T_1 \to T_2,\ t_2 : T_2 \to T_3,\ \dots,\ t_n : T_n \to \mathrm{Int}$

function app

$$\big(\forall\, n{:}int,\ t_n(t_{n-1}(\cdots(t)\cdots)) \downarrow \underline{n} \Rightarrow \langle\!\langle t_n \rangle\!\rangle \circ \langle\!\langle t_{n-1} \rangle\!\rangle \circ \cdots \circ \langle\!\langle t \rangle\!\rangle \downarrow \underline{n}\big)$$

$$\wedge\ \big(t_n(t_{n-1}(\cdots(t)\cdots)) \uparrow \Rightarrow \langle\!\langle t_n \rangle\!\rangle \circ \langle\!\langle t_{n-1} \rangle\!\rangle \circ \cdots \circ \langle\!\langle t \rangle\!\rangle \uparrow\big)$$

Similarly for any base type

Relational proof technique

For each type T, $\sim^{S\,M}_{T} \subseteq$ Source $\times$ Machine code

s.t. $\big(\forall\, t : T,\ t \sim^{S\,M}_{T} \langle\!\langle t \rangle\!\rangle\big) \Rightarrow \langle\!\langle \cdot \rangle\!\rangle$ is correct

Motivation : Observational Equivalence on Machine code

Optimization

???

Hand-written code or code compiled from another compiler.

We want to show $\langle\!\langle t \rangle\!\rangle \approx^{obs} p$ and safely use P in place of $\langle\!\langle t \rangle\!\rangle$.

Naïve observational equivalence

$$P_1 \approx^{obs} P_2 \text{ iff } \forall C, \big(C[P_1]\uparrow \Leftrightarrow C[P_2]\uparrow \big) \wedge \big(\forall n{:}int, C[P_1]\downarrow \underline{n} \Leftrightarrow C[P_2]\downarrow \underline{n} \big)$$

$\approx^{obs}$ becomes the identity relation : Too strong!!!

Typeful observational equivalence

$$\{ M(T) \subseteq \text{Machine code} \}_{T \in Type}$$

???

$$P_1 \approx^{obs}_{T} P_2 \text{ iff } \forall Q \in M(T \to Int), \big(Q \circ P_1 \uparrow \Leftrightarrow Q \circ P_2 \uparrow \big) \wedge \big(\forall n{:}int, Q \circ P_1 \downarrow \underline{n} \Leftrightarrow Q \circ P_2 \downarrow \underline{n} \big)$$

$\subseteq M(T) \times M(T)$

Motivation : Summary

We want

① Types on Machine code $\{ M(T) \subseteq \text{Machine code} \}_{T:\text{Type}}$

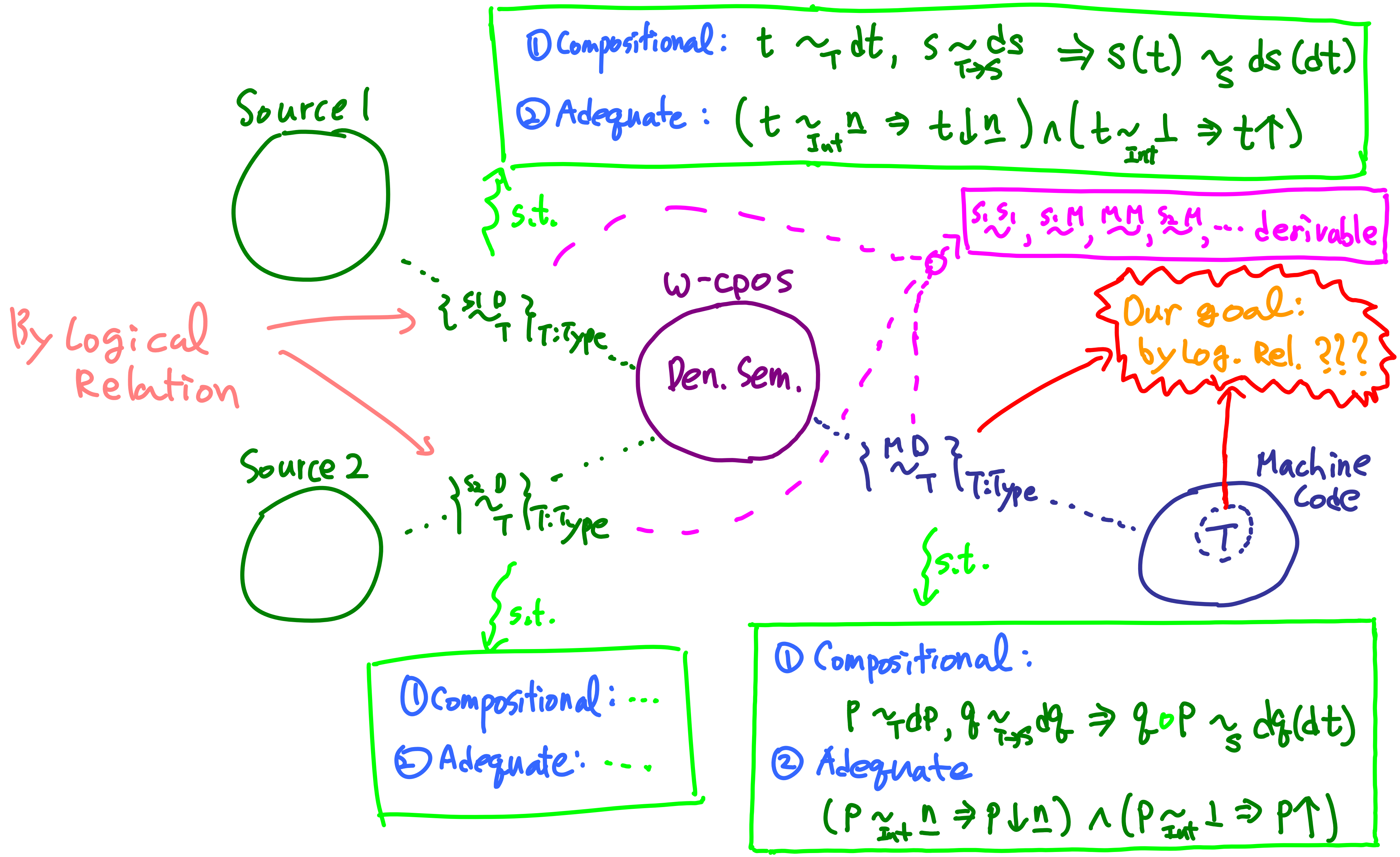
② Relations $\{ \overset{S\ \ M}{\sim}_T \subseteq \text{Source}(T) \times M(T) \}_{T:\text{Type}}$

s.t. $\left(\forall t:T,\ t \overset{S\ M}{\sim}_T \langle\langle t \rangle\rangle \right) \Rightarrow \langle\langle \cdot \rangle\rangle$ is correct

③ Relations $\{ \overset{M\ \ M}{\sim}_T \subseteq M(T) \times M(T) \}_{T \in \text{Type}}$

s.t. $P_1 \overset{M\ M}{\sim}_T P_2 \Rightarrow P_1 \overset{obs}{\approx}_T P_2$

Our approach : Overview



Negative Result

The standard Logical Relation does NOT work!!! (in the presence of recursion & Eq)

- CBV Untyped Lambda Calculus + Eq

  Val $:= x \mid \lambda x.t$ Term $:= v \mid t\,s \mid \underbrace{u \equiv_\alpha v}_{\text{Syntactic Eq test}} \mid \text{ERROR}$

  with the usual encoding of

  True, False, if - then - else - , rec -

- Theorem

  For $T \in \text{Type} := \text{Bool} \mid T \to T$ and $\{[\![T]\!] \subseteq \text{Term}\}_{T \in \text{Type}}$

  (1) $\text{True}, \text{False} \in [\![\text{Bool}]\!]$

  (2) $(\exists v_1 \in [\![A_1]\!], \cdots, v_n \in [\![A_n]\!],\ u\,v_1 \cdots v_n \overset{+}{\leadsto} \text{ERROR}) \Rightarrow u \notin [\![A_1 \to \cdots \to A_n \to B]\!]$

  (3) $(\forall v \in [\![A]\!],\ u\,v \overset{+}{\leadsto} v) \Rightarrow u \in [\![A \to A]\!]$

  Then $\lambda f.\,\text{rec}\ f \notin [\![((\text{Bool} \to \text{Bool}) \to \text{Bool} \to \text{Bool}) \to \text{Bool} \to \text{Bool}]\!]$

Logical Relation + Biorthogonality

$T \in \mathrm{Type} ::= \mathrm{Int} \mid T \to T$ Domains $[\![T]\!] = \begin{cases} \mathbb{N} & \text{if } T = \mathrm{Int} \\ [\![A]\!] \xrightarrow{c} [\![B]\!]_\perp & \text{if } T = A \to B \end{cases}$

We want $\overset{v}{\sim}_T \subseteq \text{Machine code} \times [\![T]\!]$, $\overset{c}{\sim}_{\Gamma;T} \subseteq \text{Machine code} \times ([\![\Gamma]\!] \xrightarrow{c} [\![T]\!]_\perp)$.

Types for local variables or environment

$\Rightarrow \quad \mathcal{V}(T) = \{p \mid \exists d,\ p \overset{v}{\sim}_T d\}$ $\mathcal{C}(T) = \{p \mid \exists d,\ p \overset{c}{\sim}_{\emptyset;T} d\}$

Try the standard Logical Relation

$\overset{v}{\sim}_{\mathrm{Int}}$: $\underline{n} \overset{v}{\sim}_{\mathrm{Int}} n$

$\overset{c}{\sim}_{\emptyset;\mathrm{Int}}$: $p \overset{c}{\sim}_{\emptyset;\mathrm{Int}} \perp$ iff $p\uparrow$, $p \overset{c}{\sim}_{\emptyset;\mathrm{Int}} [n]$ iff $p \downarrow \underline{n}$ (X)

$p \overset{c}{\sim}_{\emptyset;\mathrm{Int}} \perp$ iff $\forall c,\ c[p]\uparrow$

$p \overset{c}{\sim}_{\emptyset;\mathrm{Int}} [n]$ iff $\forall c,\ (c[\underline{n}]\uparrow \Rightarrow c[p]\uparrow)$
$\wedge\ (c[\underline{n}]\downarrow \Rightarrow c[p]\downarrow)$ (o.k.)

$\boxed{p \in \{n\}^{\perp\perp} \wedge p \in \{n\}^{\top\top}}$

Biorthogonality

$\mathbb{P}^\perp = \{c \mid \forall p \in \mathbb{P},\ c[p]\uparrow\}$ $\mathbb{C}^\perp = \{p \mid \forall c \in \mathbb{C},\ c[p]\uparrow\}$ ⇝ Galois connection

$\mathbb{P}^\top = \{c \mid \forall p \in \mathbb{P},\ c[p]\downarrow\}$ $\mathbb{C}^\top = \{p \mid \forall c \in \mathbb{C},\ c[p]\downarrow\}$ ⇝

Logical Relation + Biorthogonality

$\sim^{v}_{A\to B}$: $\underline{f} \sim^{v}_{A\to B} df$ iff $\forall \underline{a} \sim^{v}_{A} da,\ \underline{f} \circ \underline{a} \in \mathcal{C}_e[df(da)]$

$\Rightarrow$ $\boxed{\mathcal{C}_e(d:T) = \{ p \mid p \cancel{\sim}^{c}_{T} d }$

$\boxed{\mathcal{V}(d:T) = \{ \underline{p} \mid \underline{p} \sim^{v}_{T} d \}}$

$\sim^{c}_{\emptyset\, A\to B}$: $f \sim^{c}_{A\to B} \bot$ iff $\forall c,\ c[f]\uparrow$

$f \sim^{c}_{A\to B} [df]$ iff $f \in \mathcal{V}(df)^{\bot\bot} \wedge f \in \mathcal{V}(df)^{\top\top}$

$\sim^{c}_{\Gamma\, T}$: $f \sim^{c}_{T} df$ iff $\forall e \sim^{v}_{\Gamma} de$ $\rightsquigarrow$ $\boxed{\Gamma = (T_1, \cdots, T_n)\ e = (e_1, \cdots, e_n)\ de = (de_1, \cdots, de_n)\ \ \forall i\ \ e_i \sim^{v}_{T_i} de_i}$

$(df(de) = \bot \Rightarrow \forall c, (c \diamond e)[f]\uparrow\ )$

$\wedge\ (df(de) = [d] \Rightarrow \forall c,$

$(\forall \underline{f} \in \mathcal{V}[d], (c \diamond e)[\underline{f}]\uparrow) \Rightarrow (c \diamond e)[f]\uparrow$

$\wedge\ (\forall \underline{f} \in \mathcal{V}[d], (c \diamond e)[\underline{f}]\downarrow) \Rightarrow (c \diamond e)[f]\downarrow)$

$\leftsquigarrow$ $\boxed{P \in \mathcal{V}[d]^{\bot_e \bot_e} \wedge P \in \mathcal{V}[d]^{\top_e \top_e}}$

e-Parametrized Biorthogonality

$\mathbb{P}^{\bot_e} = \{ c \mid \forall p \in \mathbb{P}, (c \diamond e)[p]\uparrow \}$ $\quad \mathbb{C}^{\bot_e} = \{ p \mid \forall c \in \mathbb{C}, (c \diamond e)[p]\uparrow \}$

$\mathbb{P}^{\top_e} = \{ c \mid \forall p \in \mathbb{P}, (c \diamond e)[p]\downarrow \}$ $\quad \mathbb{C}^{\top_e} = \{ p \mid \forall c \in \mathbb{C}, (c \diamond e)[p]\downarrow \}$

Step-indexed logical relation

- The previous try fails due to the negative result.

- We introduce the step-indexing as an appropriate notion of approximation, but it only applies to the divergence observation, not to the termination one.

- e.g.

$$p \sim^{k\,c}_{Int} n \text{ iff } \forall C. \left(C[\underline{n}] \rightsquigarrow^{k} \Rightarrow C[p] \rightsquigarrow^{k}\right) \text{ (o.k.)}$$
$$\wedge \left(C[\underline{n}] \downarrow^{k} \Rightarrow C[p] \downarrow^{k}\right) \text{ (X)}$$

- We split the relation $\sim$ into $\lhd$ and $\rhd$ and apply the step-indexing to $\lhd$.

- Indeed, the negative rusult essentially arises from $\lhd$.

step-indexed logical relation + Biorthogonality

$\underline{n}\ {}^{k}\!\lhd^{V}_{Int}\ n$

$p\ {}^{k}\!\lhd^{C}_{Int}\ \perp$ iff $\forall c,\ c[p] \overset{k}{\rightsquigarrow}$

$\boxed{(k,p) \in \{(j,\underline{n}) \mid j \in \mathbb{N}\}^{\perp\perp}}$

$p\ {}^{k}_{\emptyset}\!\lhd^{C}_{Int}\ [n]$ iff $\forall j \le k\ \forall c,\ c[\underline{n}] \overset{j}{\rightsquigarrow} \Rightarrow c[p] \overset{j}{\rightsquigarrow}$

$\underline{f}\ {}^{k}\!\lhd^{V}_{A\to B}\ df$ iff $\forall j \le k\ \forall\ \underline{a}\ {}^{j}\!\lhd^{V}_{A}\ da\quad \underline{f} \circ \underline{a}\ {}^{j}_{\emptyset}\!\lhd^{C}_{B}\ df(da)$

$f\ {}^{k}_{\emptyset}\!\lhd^{C}_{A\to B}\ \perp$ iff $\forall c,\ c[p] \overset{k}{\rightsquigarrow}$

$\boxed{(k,f) \in \{(j,\underline{f}) \mid \underline{f}\ {}^{j}\!\lhd^{V}_{A\to B}\ df\}^{\perp\perp}}$

$f\ {}^{k}_{\emptyset}\!\lhd^{C}_{A\to B}\ [df]$ iff $\forall j \le k\ \forall c \left(\left(\forall j' \le j\ \forall\ \underline{f}\ {}^{j'}\!\lhd^{V}_{A\to B}\ df,\ c[\underline{f}] \overset{j'}{\rightsquigarrow}\right) \Rightarrow c[f] \overset{j}{\rightsquigarrow}\right)$

$f\ {}^{k}_{\Gamma}\!\lhd^{C}_{A\to B}\ df$ iff $\forall j \le k\ \forall e\ {}^{j}\!\lhd^{V}_{\Gamma}\ de\ \ldots$

$\underline{p} \lll^{V}_{T} d$ iff $\forall k\ \ \underline{p}\ {}^{k}\!\lhd^{V}_{T}\ d$

$p\ {}_{\Gamma}\!\lll^{C}_{T} d$ iff $\forall k\ \ p\ {}^{k}_{\Gamma}\!\lhd^{C}_{T}\ d$

$\rightsquigarrow$ Works well !!

Chain approximation + Biorthogonality

- same as before (just remove the condition on divergence)

  $\underline{n} \rhd^{V}_{Int} n$, $P \rhd^{C}_{\emptyset\, Int} \bot$ iff true, $\cdots$, $f \rhd^{C}_{\emptyset\, A \to B}\, df$ iff $f \in V[\![df^{\vee}]\!]^{\top\top}$, $\cdots$

- For a fixed $\underline{P}$, $\{ d \mid \underline{P} \rhd^{V}_{T} d \} \subseteq [\![T]\!]$ is NOT admissible in general because $(-)^{\top\top}$ is not closed under intersection.

  i.e. $\left(\bigcap_{k \in \mathbb{N}} V_n\right)^{\top\top} \neq \bigcap_{k \in \mathbb{N}} V_n^{\top\top}$

- Our solution: chain approximation

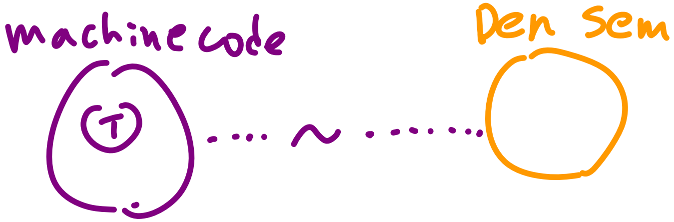
  $P \rhd\!\!\rhd^{V}_{T} d$ iff $\exists \langle d_i \rangle$ s.t. lub $\langle d_i \rangle \geq d \wedge \forall i.\ P \rhd^{V}_{T} d_i$

  $P \rhd\!\!\rhd^{C}_{\Gamma\, T} d$ iff $\exists \langle d_i \rangle$ s.t. lub$\langle d_i \rangle \gtrsim d \wedge \forall i.\ P \rhd^{C}_{\Gamma\, T} d_i$

- N.B. ① $\rhd^{V}_{T} \subseteq \rhd\!\!\rhd^{V}_{T}$ and $\rhd^{C}_{\Gamma\, T} \subseteq \rhd\!\!\rhd^{C}_{\Gamma\, T}$ $\rightsquigarrow$ $\{ d \mid \underline{P} \rhd\!\!\rhd^{V}_{T} d \}$ addmissible

  ② $\rhd^{V}_{Int} = \rhd\!\!\rhd^{V}_{Int}$ and $\rhd^{C}_{\emptyset\, Int} = \rhd^{C}_{\emptyset\, Int}$ $\rightsquigarrow$ we only observe on base types, so it does not affect observation.

Step-indexing + Chain approximation + Biorthogonality



- $\underline{p} \sim^{V}_{T} d$ iff $\underline{p} \triangleleft\!\triangleleft^{V}_{T} d \wedge \underline{p} \triangleright\!\triangleright^{V}_{T} d$

  $p \sim^{c}_{\rho\,T} d$ iff $p \triangleleft\!\triangleleft^{c}_{\rho\,T} d \wedge p \triangleright\!\triangleright^{c}_{\rho\,T} d$

- $\mathcal{V}[T] = \{ \underline{p} \mid \exists d\ \underline{p} \sim^{V}_{T} d \}$

  $\mathcal{C}[T] = \{ p \mid \exists d\ p \sim^{c}_{\emptyset\,T} d \}$

  N.B. $\mathcal{V}[T] \subseteq \mathcal{C}[T]$

- $p_1 \approx^{c}_{\rho\,T} p_2$ iff $\exists d\ p_1 \sim^{c}_{\rho\,T} d \wedge p_2 \sim^{c}_{\rho\,T} d$

- $\overset{+}{\approx}{}^{c}_{\rho\,T}$ is the transitive closure of $\approx^{c}_{\rho\,T}$.

Compositionality and Adequacy

Theorem (Compositionality)

$$f \underset{\Gamma\, A\to B}{\overset{c}{\sim}} df \;,\; a \underset{\Gamma\, A}{\overset{c}{\sim}} da \;\Rightarrow\; f \circ a \underset{B}{\overset{c}{\sim}} df(da)$$

Definition

- $P \Downarrow \underline{n}$ iff $\forall C, (C[\underline{n}]\uparrow \Rightarrow C[P]\uparrow)$
  $\wedge (C[\underline{n}]\downarrow \Rightarrow C[P]\downarrow)$
- $P \Uparrow$ iff $\forall C, \; C[P]\uparrow$

Theorem (Adequacy)

$$f \underset{\mathrm{Int}}{\overset{c}{\sim}} \bot \;\Rightarrow\; f \Uparrow$$

$$f \underset{\mathrm{Int}}{\overset{c}{\sim}} [n] \;\Rightarrow\; f \Downarrow n$$

Corollary (obs. Equv.)

$$P_1 \underset{T}{\overset{c}{\approx}} P_2 \;\Rightarrow\; P_1 \underset{T}{\overset{obs}{\approx}} P_2$$

$\forall Q : T \to \mathrm{Int}$
$(Q \circ P_1 \Downarrow \underline{n} \Leftrightarrow Q \circ P_2 \Downarrow n)$
$\wedge (Q \circ P_1 \Uparrow \Leftrightarrow Q \circ P_2 \Uparrow)$

$$\therefore\; P_1 \underset{T}{\overset{+c}{\approx}} P_2 \;\Rightarrow\; P_1 \underset{T}{\overset{obs}{\approx}} P_2 \quad (\text{as } \underset{T}{\overset{obs}{\approx}} \text{ is an equivalence relation})$$

Machine Language: SECD Machine + Eq

Inst ::= Swap | Dup | PushV n | Op $*$ | PushC c | PushRC c
| App | Ret | Sel(c_1,c_2) | Join | MkPair | Fst | Snd | Eq

Syntactic Eq test

Val ::= $\underline{n}$ | CL(e,c) | RCL(e,c) | PR(v_1,v_2)

$c \in$ Code = list Inst

$e \in$ Env = list Val

$s \in$ Stack = list Val

$d \in$ Dump = list (Code × Env × Stack)

CESD = Code × Env × Stack × Dump

- Operational Semantics (Selected)

$\langle \text{PushC } body :: c, e, s, d \rangle \mapsto \langle c, e, CL(e, body) :: s, d \rangle$

$\langle \text{PushRC } body :: c, e, s, d \rangle \mapsto \langle c, e, RCL(e, body) :: s, d \rangle$

$\langle \text{App} :: c, e, v :: CL(e', body) :: s, d \rangle \mapsto \langle body, v :: e', [], (c, e, s) :: d \rangle$

$\langle \text{App} :: c, e, v :: RCL(e', body) :: s, d \rangle \mapsto \langle body, v :: RCL(e', body) :: e', [], (c, e, s) :: d \rangle$

$\langle \text{Ret} :: c, e, v :: s, (c', e', s') :: d \rangle \mapsto \langle c', e', v :: s', d \rangle$

$\langle \text{Eq} :: c, e, v_1 :: v_2 :: s, d \rangle \mapsto \langle c, e, \underline{1} :: s, d \rangle$ (if $v_1 = v_2$)

$\langle \text{Eq} :: c, e, v_1 :: v_2 :: s, d \rangle \mapsto \langle c, e, \underline{0} :: s, d \rangle$ (if $v_1 \neq v_2$)

Types and Relations on SECD

$T \in \mathrm{Type} ::= \mathrm{Int} \mid \mathrm{Bool} \mid T_1 * T_2 \mid T_1 \to T_2$

$\mathrm{Comp} = \mathrm{Code} \times \mathrm{Stack}$

$\mathrm{VToC}(v) = (\mathrm{nil}, [v])$

$\mathrm{CToC}(\mathrm{code}) = (\mathrm{code}, \mathrm{nil})$

Conceptually $c : \Gamma \to T \xrightarrow{e:\Gamma} v : T$

$(c +\!\!+ c_0, e +\!\!+ e_0, s_0, d_0) \rightsquigarrow (c' +\!\!+ c_0, e +\!\!+ e_0, s' +\!\!+ s_0, d_0)$
$\rightsquigarrow (c_0, e +\!\!+ e_0, v :: s_0, d_0)$

$\circ : \mathrm{Val} \times \mathrm{Val} \to \mathrm{Comp} \quad : \quad \underline{f} \circ \underline{a} \triangleq (\mathrm{App}, [\underline{a}, \underline{f}])$

$\circ : \mathrm{Code} \times \mathrm{Code} \to \mathrm{Comp} \quad : \quad f \circ a \triangleq (f +\!\!+ a :: [\mathrm{App}], \mathrm{nil})$

$\diamond : \mathrm{CESD} \times \mathrm{Env} \to \mathrm{CESD} \quad : (c_0, e_0, s_0, d_0) \diamond e \triangleq (c_0, e +\!\!+ e_0, s_0, d_0)$

$[\,] : \mathrm{CESD} \times \mathrm{Comp} \to \mathrm{CESD} : (c_0, e_0, s_0, d_0)[c, s] \triangleq (c +\!\!+ c_0, e_0, s +\!\!+ s_0, d_0)$

$\sim^V_T \subseteq \mathrm{Val} \times [\![T]\!]$

$\sim^C_{\Gamma \vdash T} \subseteq \mathrm{Comp} \times ([\![\Gamma]\!] \hookrightarrow [\![T]\!]_\perp)$

Compositional and Adequate !!

Source Language : PCF$_v$

$T \in$ Types $:=$ Int | Bool | $T_1 ** T_2$ | $T_1 \to T_2$

Values:

$$[TVAR]\ \frac{}{\Gamma, x : t \vdash x : t} \qquad [TBOOL]\ \frac{}{\Gamma \vdash b : \mathtt{Bool}}\ (b \in \mathbb{B}) \qquad [TINT]\ \frac{}{\Gamma \vdash n : \mathtt{Int}}\ (n \in \mathbb{N})$$

$$[TFIX]\ \frac{\Gamma, f : t \to t', x : t \vdash M : t'}{\Gamma \vdash \mathrm{Fix}\ f\ x\ =\ M : t \to t'} \qquad [TP]\ \frac{\Gamma \vdash V_i : t_i\ (i = 1, 2)}{\Gamma \vdash \langle V_1, V_2 \rangle : t_1 \times t_2}$$

Expressions:

$$[TVAL]\ \frac{\Gamma \vdash V : t}{\Gamma \vdash [V] : t} \qquad [TLET]\ \frac{\Gamma \vdash M : t \quad \Gamma, x : t \vdash N : t'}{\Gamma \vdash \mathtt{let}\ x = M\ \mathtt{in}\ N : t'}$$

$$[TAPP]\ \frac{\Gamma \vdash V_1 : t \to t' \quad \Gamma \vdash V_2 : t}{\Gamma \vdash V_1\, V_2 : t'} \qquad [TIF]\ \frac{\Gamma \vdash V : \mathtt{Bool} \quad \Gamma \vdash M_1 : t \quad \Gamma \vdash M_2 : t}{\Gamma \vdash \mathtt{if}\ V\ \mathtt{then}\ M_1\ \mathtt{else}\ M_2 : t}$$

$$[TOP]\ \frac{\Gamma \vdash V_1 : \mathtt{Int} \quad \Gamma \vdash V_2 : \mathtt{Int}}{\Gamma \vdash V_1 \star V_2 : \mathtt{Int}} \qquad [TGT]\ \frac{\Gamma \vdash V_1 : \mathtt{Int} \quad \Gamma \vdash V_2 : \mathtt{Int}}{\Gamma \vdash V_1 > V_2 : \mathtt{Bool}}$$

$$[TFST, TSND]\ \frac{\Gamma \vdash V : t_1 \times t_2}{\Gamma \vdash \pi_i(V) : t_i\ (i = 1, 2)}$$

Figure 1. Typing rules for PCF$_v$

Compiler : PCF$_v$ to SECD

Values:

$$
\begin{aligned}
(\!| x_1 : t_1, \ldots, x_n : t_n \vdash x_i : t_i |\!) &= [\texttt{PushV}\; i] \\
(\!| \Gamma \vdash \mathit{true} : \texttt{Bool} |\!) &= [\texttt{PushN}\; 1] \\
(\!| \Gamma \vdash \mathit{false} : \texttt{Bool} |\!) &= [\texttt{PushN}\; 0] \\
(\!| \Gamma \vdash n : \texttt{Int} |\!) &= [\texttt{PushN}\; n] \\
(\!| \Gamma \vdash \langle V_1, V_2 \rangle : t_1 \times t_2 |\!) &= (\!| \Gamma \vdash V_1 : t_1 |\!) +\!\!+ (\!| \Gamma \vdash V_2 : t_2 |\!) +\!\!+ [\texttt{MkPair}] \\
(\!| \Gamma \vdash \texttt{Fix}\; f\, x \;=\; M : t \to t' |\!) &= [\texttt{PushRC}\; ((\!| \Gamma, f : t \to t', x : t \vdash M : t' |\!) +\!\!+ [\texttt{Ret}])]
\end{aligned}
$$

Expressions:

$$
\begin{aligned}
(\!| \Gamma \vdash [V] : t |\!) &= (\!| \Gamma \vdash V : t |\!) \\
(\!| \Gamma \vdash \texttt{let}\; x = M \;\texttt{in}\; N : t' |\!) &= [\texttt{PushC}\; ((\!| \Gamma, x : t \vdash N : t' |\!) +\!\!+ [\texttt{Ret}])] +\!\!+ (\!| \Gamma \vdash M : t |\!) +\!\!+ [\texttt{App}] \\
(\!| \Gamma \vdash V_1\, V_2 : t' |\!) &= (\!| \Gamma \vdash V_1 : t \to t' |\!) +\!\!+ (\!| \Gamma \vdash V_2 : t |\!) +\!\!+ [\texttt{App}] \\
(\!| \Gamma \vdash \texttt{if}\; V \;\texttt{then}\; M_1 \;\texttt{else}\; M_2 : t |\!) &= (\!| \Gamma \vdash V : \texttt{Bool} |\!) +\!\!+ [\texttt{Sel}\; (((\!| \Gamma \vdash M_1 : t |\!) +\!\!+ [\texttt{Join}]), ((\!| \Gamma \vdash M_2 : t |\!) +\!\!+ [\texttt{Join}]))] \\
(\!| \Gamma \vdash V_1 \star V_2 : \texttt{Int} |\!) &= (\!| \Gamma \vdash V_1 : \texttt{Int} |\!) +\!\!+ (\!| \Gamma \vdash V_2 : \texttt{Int} |\!) +\!\!+ [\texttt{Op}\; \star] \\
(\!| \Gamma \vdash V_1 > V_2 : \texttt{Bool} |\!) &= (\!| \Gamma \vdash V_1 : \texttt{Int} |\!) +\!\!+ (\!| \Gamma \vdash V_2 : \texttt{Int} |\!) +\!\!+ [\texttt{Op}\; (\lambda(n_1, n_2).n_1 > n_2 \supset 1 \mid 0)]
\end{aligned}
$$

Figure 3. Compiler for PCF$_v$

Compiler Correctness

- standard denotational semantics for PCF_v

$$[\![\Gamma \vdash v : T]\!] \in [\![\Gamma]\!] \hookrightarrow [\![T]\!], \quad [\![\Gamma \vdash t : T]\!] \in [\![\Gamma]\!] \hookrightarrow [\![T]\!]_\perp$$

- known to be compositional and adequate

(The proof uses the standard logical relation.)

$$\Rightarrow [\![\Gamma]\!] \xrightarrow{[\![\Gamma \vdash v : T]\!]} [\![T]\!] \xrightarrow{[\,]} [\![T]\!]_\perp$$

Theorem (Correctness)

$$\Gamma \vdash v : T \quad \Rightarrow CToC(⦇\Gamma \vdash v : T⦈) \underset{\Gamma \vdash T}{\overset{c}{\sim}} \left[[\![\Gamma \vdash v : T]\!] \right]$$

$$\Gamma \vdash t : T \quad \Rightarrow CToC(⦇\Gamma \vdash t : T⦈) \underset{\Gamma \vdash T}{\overset{c}{\sim}} [\![\Gamma \vdash t : T]\!]$$

Corollary

$$\forall\ \emptyset \vdash t : T_1,\ \emptyset \vdash t_1 : T_1 \to T_2,\ \cdots,\ \emptyset \vdash t_n : T_n \to Int$$

$$(t_n\, t_{n-1} \cdots t \Uparrow \ \Rightarrow ⦇t_n⦈ \circ ⦇t_{n-1}⦈ \circ \cdots \circ ⦇t⦈ \Uparrow)$$

$$\wedge\ (\forall n{:}int \quad " \quad \downarrow \underline{n} \ \Rightarrow \quad " \quad \Downarrow \underline{n})$$

Equational Reasoning I

① Commutativity of addition

$\text{pluscode}(\Gamma) = (\!| \Gamma \vdash \text{Fix}\ \_x.\ \text{Fix}\ \_y.\ x+y\ |\!)$

Proposition $\rightsquigarrow$ $\boxed{\exists dc_1,\ c_1 \underset{\Gamma\,\text{Int}}{\overset{c}{\sim}} dc_1}$

$\forall\ c_1 : \Gamma \vdash \text{Int} \quad c_2 : \Gamma \vdash \text{Int}$

$\text{CToC}(\text{pluscode}(\Gamma) \circ c_1 \circ c_2) \underset{\Gamma\ T}{\overset{c}{\approx}} \text{CToC}(\text{pluscode}(\Gamma) \circ c_2 \circ c_1)$

② First projection

$\text{projfstcode}(\Gamma, T_1, T_2) = (\!| \Gamma \vdash \text{Fix}\ \_x.\ \text{Fix}\ \_y.\ x : T_1 \to T_2 \to T_1\ |\!)$

Proposition

$\forall\ c_1 : \Gamma \vdash T_1\ ,\quad c_2 \underset{\Gamma\,T_2}{\overset{c}{\sim}} [dc_2]$ $\rightsquigarrow$ $\boxed{[\![\Gamma]\!] \xrightarrow{dc_2} [\![T_2]\!] \xrightarrow{[-]} [\![T_2]\!]_\perp}$

$\text{CToC}(\text{projfstcode}(\Gamma, T_1, T_2) \circ c_1 \circ c_2) \underset{\Gamma\ T_1}{\overset{c}{\approx}} c_1$

Equational Reasoning II

③ Optimizing iteration

$$\mathrm{idcode}(\Gamma) = (\!| \Gamma \vdash \mathrm{Fix}\ \_\ x.\ x : \mathrm{Int} \to \mathrm{Int} |\!)$$

$$\mathrm{appncode}(\Gamma) = (\!| \Gamma \vdash \mathrm{Fix}\ \_\ f.\ \mathrm{Fix}\ apf\ n.\ \mathrm{Fix}\ \_\ v.$$

$\rightsquigarrow \boxed{\mathrm{appn}\ f\ n\ v = f^n v}$

$$\left.\begin{array}{l} \text{if } n>0 \text{ then} \\ \quad f(apf\ (n-1)\ v) \\ \text{else} \\ \quad v \end{array}\right\}\ \text{pseudo code}$$

$$: (\mathrm{Int}\to\mathrm{Int})\to\mathrm{Int}\to\mathrm{Int}\to\mathrm{Int} |\!)$$

$$\mathrm{appnoptcode}(\Gamma) = [\mathrm{Push}(\ \cdots\ ]$$

syntactic Eq test

$\rightsquigarrow \lambda f.\lambda n.\lambda v.$ if $\mathrm{Eq}(f, \mathrm{idcode}(\Gamma))$
then v
else $\mathrm{appncode}(\Gamma)\ f\ n\ v$

Proposition

$$\mathrm{CToC}(\mathrm{appncode}(\Gamma)) \ \approx^{c}_{\Gamma\ (\mathrm{Int}\to\mathrm{Int})\to\mathrm{Int}\to\mathrm{Int}\to\mathrm{Int}}\ \mathrm{CToC}(\mathrm{appnoptcode}(\Gamma))$$

Discussion & Future work

- Discussion
  - 5000 lines in Coq
    excluding domain package & PCF_v & its denotational semantics
  - Full abstraction issue

- Future work
  - idealized assembly language
  - reference and polymorphism
  - recursive types
  - effects (input, ouput, exception, ...)