

# Type soundness proofs with big-step operational semantics of OOL

Davide Ancona

DIBRIS, Università di Genova

Shonan Meeting on Semantics and Verification of Object-Oriented Languages  
September 21-25, 2015

# Motivations

- Proof of type soundness with big-step operational semantics
- Previous proposals use extra rules  
[LeroyGrall09,ErnstOstermann06,KusmierrekBono10]
- Proof techniques for proving type soundness with the standard rules

# Outline

- background on inference system
- type soundness with small-step and big-step operational semantics
- coinductive big-step semantics
- approximating semantics
- future directions

# Inference systems

- a formalism for recursively defining predicates (or relations)
- strongly related to logic programming
- a simple example
  - ▶ definition of predicate **leq**
  - ▶ domain: terms over constant 0 and unary operator  $s$

$$\text{(rule 1)} \frac{}{\mathbf{leq}(0, N)} \quad \text{(rule 2)} \frac{\mathbf{leq}(N_1, N_2)}{\mathbf{leq}(s(N_1), s(N_2))}$$

- standard inductive semantics:  
the equation

$$S = \{\mathbf{leq}(0, t) \mid t \text{ term}\} \cup \{\mathbf{leq}(s(t_1), s(t_2)) \mid \mathbf{leq}(t_1, t_2) \in S\}$$

has only one solution

$$S^{in} = \{\mathbf{leq}(s^n(0), s^m(0)) \mid n \leq m\}$$

# Inference systems and coinductive semantics

$$\text{(rule 1)} \frac{}{\mathbf{leq}(0, N)} \quad \text{(rule 2)} \frac{\mathbf{leq}(N_1, N_2)}{\mathbf{leq}(s(N_1), s(N_2))}$$

- let us extend the domain with  $t$  s.t.  $t = s(t)$  (that is,  $t = s^\omega$ )
- equation

$$S = \{\mathbf{leq}(0, t) \mid t \text{ term}\} \cup \{\mathbf{leq}(s(t_1), s(t_2)) \mid \mathbf{leq}(t_1, t_2) \in S\}$$

has two different solutions

- ▶ least solution:  $S^{in} = \{\mathbf{leq}(s^n(0), s^m(0)) \mid n \leq m\} \cup \{\mathbf{leq}(s^n(0), s^\omega)\}$
- ▶ greatest solution:  $S^{co} = S^{in} \cup \{\mathbf{leq}(s^\omega, s^\omega)\}$

# Fixed-point semantics

- $HU$ : Herbrand Universe of terms
- Herbrand model: a subset of  $HB = \{\mathbf{leq}(t_1, t_2) \mid t_1, t_2 \in HU\}$
- from the recursive definition of  $\mathbf{leq}$  the following one step inference function  $\mathcal{F}$  can be derived

$$\mathcal{F} : \mathcal{P}(HB) \rightarrow \mathcal{P}(HB)$$

$$\mathcal{F}(S) = \{\mathbf{leq}(0, t) \mid t \in HU\} \cup \{\mathbf{leq}(s(t_1), s(t_2)) \mid \mathbf{leq}(t_1, t_2) \in S\}$$

- Tarski-Knaster theorem  
 $\mathcal{F}$  monotone over a complete lattice  $\Rightarrow$  there exist  $lfp(\mathcal{F})$ , and  $gfp(\mathcal{F})$

# Inductive and coinductive fixed-point semantics

- inductive semantics

- ▶  $HU =$  terms over  $0$  and  $s$
- ▶ model:  $lfp(\mathcal{F}) = \{\mathbf{leq}(s^n(0), s^m(0)) \mid n \leq m\}$

- coinductive semantics

- ▶  $HU =$  finite and **infinite** terms over  $0$  and  $s$  (non well-founded trees)
- ▶ model:  $gfp(\mathcal{F}) = lfp(\mathcal{F}) \cup \{\mathbf{leq}(s^n(0), s^\omega)\} \cup \{\mathbf{leq}(s^\omega, s^\omega)\}$

# Proof tree semantics

- $\mathbf{leq}(t_1, t_2)$  is derivable iff there exists a proof tree (a.k.a. derivation) s.t.
  - ▶ all nodes are labeled over  $HB$
  - ▶ the root is labeled with  $\mathbf{leq}(t_1, t_2)$
  - ▶ all nodes with their children constitute a valid instantiation of a rule

example:

$$\begin{array}{c} \text{(rule 1)} \overline{\mathbf{leq}(0, s(0))} \\ \text{(rule 2)} \overline{\mathbf{leq}(s(0), s(s(0)))} \end{array}$$

- inductive semantics
  - ▶  $HU =$  terms over 0 and  $s$
  - ▶ proof trees can only be finite
- coinductive semantics
  - ▶  $HU =$  finite and **infinite** terms over 0 and  $s$  (non well-founded trees)
  - ▶ proof trees are allowed to be **infinite** (non well-founded) as well



# Equivalence of the two semantics

- fixed-point and proof tree semantics are equivalent
- the result holds for both inductive and coinductive semantics
- published proofs can be found in
  - ▶ inductive semantics: *P. Aczel, An introduction to inductive definitions, Handbook of Mathematical Logic, Vol. 90 of Studies in Logics and the Foundations of Mathematics, 1977*
  - ▶ coinductive semantics: *X. Leroy and H. Grall. Coinductive big-step operational semantics. Information and Computation, 207:284304, 2009*
- example:  
 $\mathbf{leq}(s^\omega, s^\omega) \notin \mathit{lfp}(\mathcal{F}), \mathbf{leq}(s^\omega, s^\omega) \in \mathit{gfp}(\mathcal{F})$   
the proof tree for  $\mathbf{leq}(s^\omega, s^\omega)$  is infinite

$$\begin{array}{c} \vdots \\ \text{(rule 2)} \frac{\quad}{\mathbf{leq}(s^\omega, s^\omega)} \\ \text{(rule 2)} \frac{\quad}{\mathbf{leq}(s^\omega, s^\omega)} \end{array}$$

# A Featherweight Java-like language

$$\begin{aligned} p &::= \overline{cd}^n e \\ cd &::= \mathbf{class} \ c_1 \ \mathbf{extends} \ c_2 \ \{ \overline{fd}^n \ \overline{md}^k \} \quad (c_1 \neq \text{Object}) \\ fd &::= \tau \ f; \\ md &::= \tau_0 \ m(\overline{\tau x}^n) \ \{e\} \quad x_i \neq \mathbf{this} \ \forall i = 1..n \\ \tau &::= \mathbf{c} \mid \mathbf{bool} \\ e &::= \mathbf{new} \ c(\overline{e}^n) \mid x \mid e.f \mid e_0.m(\overline{e}^n) \mid \mathbf{if} \ (e) \ e_1 \ \mathbf{else} \ e_2 \\ &\quad \mid \mathbf{false} \mid \mathbf{true} \end{aligned}$$

*assumptions:*  $n, k \geq 0$ , inheritance is acyclic, names of declared classes in a program, methods and fields in a class, and parameters in a method are distinct.

# Small-step operational semantics (call by value)

values

$$v ::= \mathbf{new} \ c(\bar{v}^n) \mid \mathbf{false} \mid \mathbf{true}$$

contexts

$$C[] ::= \square \mid \mathbf{new} \ c(\bar{v}^n, \square, \bar{e}^k) \mid \square.f \mid \square.m(\bar{e}^n) \mid v.m(\bar{v}^n, \square, \bar{e}^k) \mid \mathbf{if} \ (\square) \ e_1 \ \mathbf{else} \ e_2$$

rewriting rules (with usual auxiliary functions *fields* and *meth*)

$$\text{(fld)} \frac{\text{fields}(c) = \bar{\tau}^n \bar{f}^n, \quad 1 \leq i \leq n}{\mathbf{new} \ c(\bar{v}^n).f_i \rightarrow v_i}$$

$$\text{(inv)} \frac{\text{meth}(c, m) = \bar{\tau}^n \bar{x}^n.e:\tau}{\mathbf{new} \ c(\bar{v}^k).m(\bar{v}'^n) \rightarrow e[\mathbf{this} \mapsto \mathbf{new} \ c(\bar{v}^k), \bar{x}^n \mapsto \bar{v}'^n]}$$

$$\text{(ift)} \frac{}{\mathbf{if} \ (\mathbf{true}) \ e_1 \ \mathbf{else} \ e_2 \rightarrow e_1}$$

$$\text{(iff)} \frac{}{\mathbf{if} \ (\mathbf{false}) \ e_1 \ \mathbf{else} \ e_2 \rightarrow e_2}$$

$$\text{(ctx)} \frac{e \rightarrow e'}{C[e] \rightarrow C[e']}$$

# Big-step operational semantics (call by value)

- Values (for the moment let us consider well-founded values)

$$v, u ::= \text{obj}(c, [\bar{f}^n \mapsto \bar{v}^n]) \mid \text{false} \mid \text{true}$$

- Main judgment

$$\Pi \vdash e \Rightarrow v, \text{ where } \Pi = \bar{x}^n \mapsto \bar{v}^n$$

# Rules

$$\text{(VAR)} \frac{\Pi(x) = \mathbb{v}}{\Pi \vdash x \Rightarrow \mathbb{v}}$$

$$\text{(FAL)} \frac{}{\Pi \vdash \mathbf{false} \Rightarrow \mathit{false}}$$

$$\text{(TRU)} \frac{}{\Pi \vdash \mathbf{true} \Rightarrow \mathit{true}}$$

$$\text{(NEW)} \frac{\forall i = 1..n \Pi \vdash e_i \Rightarrow \mathbb{v}_i \quad \mathit{fields}(c) = \bar{\tau}^n \bar{f}^n}{\Pi \vdash \mathbf{new} \ c(\bar{e}^n) \Rightarrow \mathit{obj}(c, [\bar{f}^n \mapsto \bar{\mathbb{v}}^n])}$$

$$\text{(IFT)} \frac{\Pi \vdash e \Rightarrow \mathit{true} \quad \Pi \vdash e_1 \Rightarrow \mathbb{v}}{\Pi \vdash \mathbf{if} \ (e) \ e_1 \ \mathbf{else} \ e_2 \Rightarrow \mathbb{v}}$$

$$\text{(IFF)} \frac{\Pi \vdash e \Rightarrow \mathit{false} \quad \Pi \vdash e_2 \Rightarrow \mathbb{v}}{\Pi \vdash \mathbf{if} \ (e) \ e_1 \ \mathbf{else} \ e_2 \Rightarrow \mathbb{v}}$$

$$\text{(FLD)} \frac{\Pi \vdash e \Rightarrow \mathit{obj}(c, [\bar{f}^n \mapsto \bar{\mathbb{v}}^n]) \quad 1 \leq i \leq n}{\Pi \vdash e.f_i \Rightarrow \mathbb{v}_i}$$

$$\text{(INV)} \frac{\forall i = 0..n \Pi \vdash e_i \Rightarrow \mathbb{v}_i \quad \mathbf{this} \mapsto \mathbb{v}_0, \bar{x}^n \mapsto \bar{\mathbb{v}}^n \vdash e \Rightarrow \mathbb{v} \quad \mathbb{v}_0 = \mathit{obj}(c, [\dots]) \quad \mathit{meth}(c, m) = \bar{\tau}^n \bar{x}^n.e:\tau}{\Pi \vdash e_0.m(\bar{e}^n) \Rightarrow \mathbb{v}}$$

# Type system (well-typed programs and declarations)

$$\text{(pro)} \frac{\forall i = 1..n \vdash cd_i : \diamond \quad \emptyset \vdash e : \tau}{\vdash \overline{cd}^n e : \diamond}$$

$$\text{(cla)} \frac{\forall i = 1..k \ c \vdash md_i : \diamond \quad \text{fields}(c) \text{ defined}}{\vdash \mathbf{class} \ c \ \mathbf{extends} \ c' \ \{ \overline{fd}^n \ \overline{md}^k \} : \diamond}$$

$$\text{(met)} \frac{\mathbf{this} : c, \overline{x}^n : \overline{\tau}^n \vdash e : \tau \quad \tau \leq \tau_0 \quad \text{override}(c, m, \overline{\tau}^n, \tau_0)}{c \vdash_{\tau_0} m(\overline{\tau}^n \ \overline{x}^n) \{e\} : \diamond}$$

# Type system (well-typed expressions)

$$\begin{array}{c} \text{(var)} \\ \hline \Gamma \vdash x:\tau \quad \Gamma(x) = \tau \end{array} \quad \begin{array}{c} \text{(fal)} \\ \hline \Gamma \vdash \mathbf{false}:\mathit{bool} \end{array} \quad \begin{array}{c} \text{(tru)} \\ \hline \Gamma \vdash \mathbf{true}:\mathit{bool} \end{array}$$

$$\text{(new)} \frac{\forall i = 1..n \Gamma \vdash e_i:\tau_i \quad \mathit{fields}(c) = \overline{\tau'}^n \overline{f}^n \quad \forall i = 1..n \tau_i \leq \tau'_i}{\Gamma \vdash \mathbf{new} \ c(\overline{e}^n):c}$$

$$\text{(fld)} \frac{\Gamma \vdash e:c \quad \mathit{fields}(c) = \overline{\tau}^n \overline{f}^n \quad 1 \leq i \leq n}{\Gamma \vdash e.f_i:\tau_i}$$

$$\text{(inv)} \frac{\forall i = 0..n \Gamma \vdash e_i:\tau_i \quad \mathit{meth}(\tau_0, m) = \overline{\tau'}^n \overline{x}^n.e:\tau \quad \forall i = 1..n \tau_i \leq \tau'_i}{\Gamma \vdash e_0.m(\overline{e}^n):\tau}$$

$$\text{(if)} \frac{\Gamma \vdash e:\mathit{bool} \quad \Gamma \vdash e_1:\tau_1 \quad \Gamma \vdash e_2:\tau_2}{\Gamma \vdash \mathbf{if} \ (e) \ e_1 \ \mathbf{else} \ e_2:\forall(\tau_1, \tau_2)}$$

# Small-step and big-step operational semantics

## Small-step operational semantics

- rewriting system approach
- suited for
  - ▶ type soundness proofs
  - ▶ concurrency

## Big-step operational semantics

- more abstract approach
- suited for
  - ▶ interpreter specification
  - ▶ imperative languages



# Type soundness

- general claim: if  $P$  is well-typed, then  $P$  cannot crash
- more in detail: if  $e$  is well-typed, then any possible evaluation of  $e$  either diverges, or returns a value
- with small-step semantics: if  $e$  is well-typed,  $e \rightarrow e'$ , and  $e'$  is in normal form, then  $e'$  is a value

# Type soundness and big-step semantics

## Program 1

```
true.m() // main expression
```

## Program 2

```
class C extends Object {bool m() {this.m()}}  
new C().m() // main expression
```

- program 1 is **not** well-typed, program 2 is well-typed
- program 1 crashes, program 2 does not terminate

- there is no value  $v$  s.t.  
 $\emptyset \vdash \mathbf{true}.m() \Rightarrow v$   
 $\emptyset \vdash \mathbf{new} C().m() \Rightarrow v$

# Type soundness and big-step semantics

## Program 1

```
true.m() // main expression
```

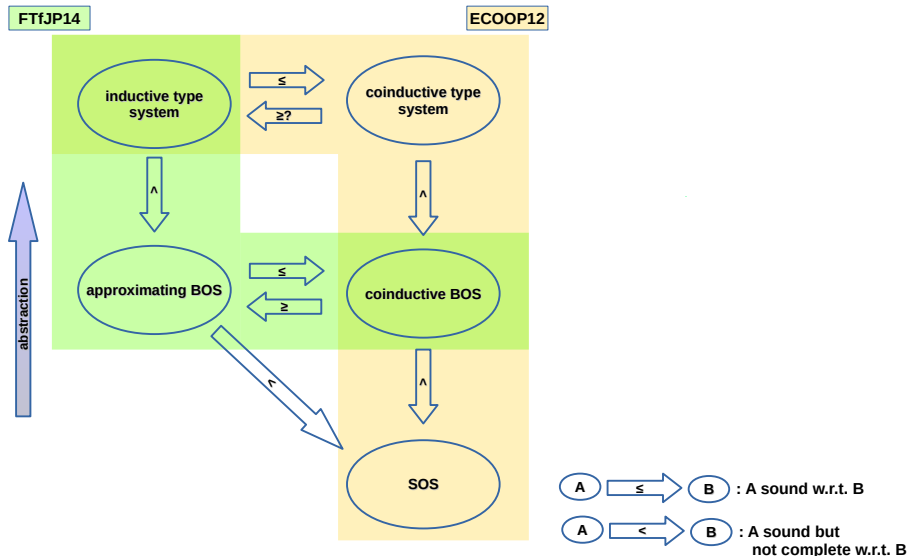
## Program 2

```
class C extends Object {bool m() {this.m()}}  
new C().m() // main expression
```

- program 1 is **not** well-typed, program 2 is well-typed
- program 1 crashes, program 2 does not terminate
- there is no value  $v$  s.t.  
 $\emptyset \vdash \mathbf{true}.m() \Rightarrow v$   
 $\emptyset \vdash \mathbf{new} C().m() \Rightarrow v$

Big-step semantics cannot distinguish programs that crash from programs that do not terminate

# Overview of the main results



# Coinductive big-step semantics (values)

## Definition

$v, u ::= \text{obj}(c, [\bar{f}^n \mapsto \bar{v}^n]) \mid \text{false} \mid \text{true}$  (**non-well-founded**)

## Examples of valid values

$\text{obj}(C, [n \mapsto \text{obj}(C, [n \mapsto \text{obj}(C, [n \mapsto \dots])])])$

that is, the unique value  $v$  s.t.  $v = \text{obj}(C, [n \mapsto v])$

$\text{obj}(L, [e \mapsto \text{obj}(Z, [ ]), n \mapsto \text{obj}(L, [e \mapsto \text{obj}(N, [p \mapsto \text{obj}(Z, [ ])]), n \mapsto \dots])])$

# Coinductive big-step semantics (rules)

$$\begin{array}{c} \Pi(x) = \mathbb{v} \\ \text{(VAR)} \frac{}{\Pi \vdash x \Rightarrow \mathbb{v}} \end{array} \quad \begin{array}{c} \text{(FAL)} \frac{}{\Pi \vdash \mathbf{false} \Rightarrow \mathit{false}} \end{array} \quad \begin{array}{c} \text{(TRU)} \frac{}{\Pi \vdash \mathbf{true} \Rightarrow \mathit{true}} \end{array}$$
$$\begin{array}{c} \forall i = 1..n \Pi \vdash e_i \Rightarrow \mathbb{v}_i \quad \mathit{fields}(c) = \bar{\tau}^n \bar{f}^n \\ \text{(NEW)} \frac{}{\Pi \vdash \mathbf{new} \ c(\bar{e}^n) \Rightarrow \mathit{obj}(c, [\bar{f}^n \mapsto \bar{v}^n])} \end{array} \quad \begin{array}{c} \Pi \vdash e \Rightarrow \mathit{true} \quad \Pi \vdash e_1 \Rightarrow \mathbb{v} \\ \text{(IFT)} \frac{}{\Pi \vdash \mathbf{if} \ (e) \ e_1 \ \mathbf{else} \ e_2 \Rightarrow \mathbb{v}} \end{array}$$
$$\begin{array}{c} \Pi \vdash e \Rightarrow \mathit{false} \quad \Pi \vdash e_2 \Rightarrow \mathbb{v} \\ \text{(IFF)} \frac{}{\Pi \vdash \mathbf{if} \ (e) \ e_1 \ \mathbf{else} \ e_2 \Rightarrow \mathbb{v}} \end{array} \quad \begin{array}{c} \Pi \vdash e \Rightarrow \mathit{obj}(c, [\bar{f}^n \mapsto \bar{v}^n]) \quad 1 \leq i \leq n \\ \text{(FLD)} \frac{}{\Pi \vdash e.f_i \Rightarrow \mathbb{v}_i} \end{array}$$
$$\begin{array}{c} \forall i = 0..n \Pi \vdash e_i \Rightarrow \mathbb{v}_i \quad \mathit{this} \mapsto \mathbb{v}_0, \bar{x}^n \mapsto \bar{v}^n \vdash e \Rightarrow \mathbb{v} \\ \mathbb{v}_0 = \mathit{obj}(c, [\dots]) \quad \mathit{meth}(c, m) = \bar{\tau}^n \bar{x}^n.e:\tau \\ \text{(INV)} \frac{}{\Pi \vdash e_0.m(\bar{e}^n) \Rightarrow \mathbb{v}} \end{array}$$

# Inductive and coinductive semantics

- inductive semantics  $\subseteq$  coinductive semantics
- the only difference concerns non terminating programs
- if  $e$  is well-typed and does not terminate, then
  - ▶ there exists **no**  $v$  s.t.  $\emptyset \vdash e \Rightarrow v$
  - ▶ there exists  $v$  s.t.  $\emptyset \Vdash e \Rightarrow v$
- valid soundness claim:  $\Gamma \vdash e:\tau \Rightarrow \exists v. \Pi \Vdash e \Rightarrow v$

# Example 1

## Program

```
class C extends Object {bool m() {this.m()}}  
new C().m() // main expression
```



# Example 1

## Program

```
class C extends Object {bool m() {this.m()}}  
new C().m() // main expression
```

## Proof tree

$$\frac{\frac{\frac{}{\emptyset \vdash \mathbf{new} C () \Rightarrow u} \quad \frac{\frac{}{\Pi \vdash \text{this} \Rightarrow u} \quad \frac{\frac{}{\Pi \vdash \text{this.m} () \Rightarrow v}}{\vdots}}{\Pi \vdash \text{this.m} () \Rightarrow v}}{\Pi \vdash \text{this} \Rightarrow u}}{\emptyset \vdash \mathbf{new} C () \Rightarrow v}}{\emptyset \vdash \mathbf{new} C ().m () \Rightarrow v}$$

- $\emptyset \vdash e \Rightarrow v$  for any  $v$
- $u = \text{obj}(C, [ ])$ ,  $\Pi = \text{this} \mapsto u$

## Example 2

### Program

```
class M extends Object {L m(){new L(this.m())}}  
class L extends Object {L n;}  
new M().m() // main expression
```

# Example 2

## Program

```
class M extends Object {L m(){new L(this.m())}}  
class L extends Object {L n;}  
new M().m() // main expression
```

## Proof tree

$$\frac{\frac{\frac{\emptyset \vdash \mathbf{new} M() \Rightarrow \mathit{obj}(M, [])}{\vdots}}{\emptyset \vdash \mathbf{new} L(\text{this.m}()) \Rightarrow v}}{\emptyset \vdash \mathbf{new} M().m() \Rightarrow \mathit{obj}(L, [n \mapsto v])}}{\emptyset \vdash \mathbf{new} M().m() \Rightarrow \mathit{obj}(L, [n \mapsto v])}}$$

- $v = \mathit{obj}(L, [n \mapsto v]), \Pi = \text{this} \mapsto \mathit{obj}(M, [])$

# Example 3

## Program

```
class Nat extends Object { }  
class Z extends Nat { }  
class NZ extends Nat {Nat p;}  
class M extends Object {L m(Nat e){new L(e,this.m(new NZ(e)))}}}  
class L extends Object {Nat e; L n;}  
new M().m(new Z()) // main expression
```



## Example 4

### Program

```
class C extends Object {bool m() {this.m()}}  
if (new C().m()) true.m() else true.m() // main expression
```

- there exists **no**  $\forall$  s.t.  $\emptyset \Vdash \mathbf{if}(\mathbf{new} C().m()) \mathbf{true}.m() \mathbf{else} \mathbf{true}.m() \Rightarrow \forall$
- $\mathbf{if}(\mathbf{new} C().m()) \mathbf{true}.m() \mathbf{else} \mathbf{true}.m()$  not well-typed

# Type soundness

## Main claim

$$\emptyset \vdash e:\tau \Rightarrow \exists v. \emptyset \Vdash e \Rightarrow v$$

## Technique

the proof uses a coinductive version of the standard type system

- $\Gamma \vdash e:\tau$  (standard system, inductive semantics)
- $\Gamma \Vdash e:\tau$  (non standard system, coinductive semantics)

## Proof

the claim can be split as follows

- 1  $\emptyset \vdash e:\tau \Rightarrow \emptyset \Vdash e:\tau$  (direct proof)
- 2  $\emptyset \Vdash e:\tau \Rightarrow \exists v. \emptyset \Vdash e \Rightarrow v$

proof of claim 2

- the proof requires to define a complete metric space of proof trees
- based on a concretization relation  $\mathcal{R}_\gamma$  (defined coinductively)

# Switching from the inductive to the coinductive TS

$$(pro) \frac{\forall i = 1..n \vdash cd_i : \diamond \quad \emptyset \vdash e : \tau}{\vdash \overline{cd}^n e : \diamond} \quad (cla) \frac{\forall i = 1..k \ c \vdash md_i : \diamond \quad \text{fields}(c) \text{ defined}}{\vdash \mathbf{class } c \text{ extends } c' \{ \overline{fd}^n \overline{md}^k \} : \diamond}$$

$$(met) \frac{\text{this} : c, \overline{x}^n : \overline{\tau}^n \vdash e : \tau \quad \tau \leq \tau_0 \quad \text{override}(c, m, \overline{\tau}^n, \tau_0)}{c \vdash_{\tau_0} m(\overline{\tau}^n \overline{x}^n) \{e\} : \diamond}$$

$$(var) \frac{\Gamma(x) = \tau}{\Gamma \vdash x : \tau} \quad (fal) \frac{}{\Gamma \vdash \mathbf{false} : \mathit{bool}} \quad (tru) \frac{}{\Gamma \vdash \mathbf{true} : \mathit{bool}}$$

$$(new) \frac{\forall i = 1..n \ \Gamma \vdash e_i : \tau_i \quad \text{fields}(c) = \overline{\tau'}^n \overline{f}^n \quad \forall i = 1..n \ \tau_i \leq \tau'_i}{\Gamma \vdash \mathbf{new } c(\overline{e}^n) : c}$$

$$(fld) \frac{\Gamma \vdash e : c \quad \text{fields}(c) = \overline{\tau}^n \overline{f}^n \quad 1 \leq i \leq n}{\Gamma \vdash e.f_i : \tau_i} \quad (if) \frac{\Gamma \vdash e : \mathit{bool} \quad \Gamma \vdash e_1 : \tau_1 \quad \Gamma \vdash e_2 : \tau_2}{\Gamma \vdash \mathbf{if } (e) e_1 \mathbf{else } e_2 : \vee(\tau_1, \tau_2)}$$

$$(inv) \frac{\forall i = 0..n \ \Gamma \vdash e_i : \tau_i \quad \text{meth}(\tau_0, m) = \overline{\tau'}^n \overline{x}^n.e : \tau \quad \forall i = 1..n \ \tau_i \leq \tau'_i}{\Gamma \vdash e_0.m(\overline{e}^n) : \tau}$$



# Switching from the inductive to the coinductive TS

$$\begin{array}{c} \text{(pro)} \frac{\forall i = 1..n \Gamma \vdash : cd_i \diamond \quad \emptyset \vdash e : \tau}{\Gamma \vdash : \overline{cd}^n e \diamond} \qquad \text{(cla)} \frac{\forall i = 1..k \Gamma \vdash c : md_i \diamond \quad \text{fields}(c) \text{ defined}}{\Gamma \vdash : \mathbf{class } c \text{ extends } c' \{ \overline{fd}^n \overline{md}^k \} \diamond} \end{array}$$

$$\text{(met)} \frac{\Gamma \vdash \text{this} : c, \overline{x}^n : \overline{\tau}^n : e : \tau \quad \tau \leq \tau_0 \quad \text{override}(c, m, \overline{\tau}^n, \tau_0)}{\Gamma \vdash c : \tau_0 \quad m(\overline{\tau}^n \overline{x}^n) \{e\} \diamond}$$

$$\begin{array}{c} \Gamma(x) = \tau \\ \text{(var)} \frac{}{\Gamma \vdash x : \tau} \qquad \text{(fal)} \frac{}{\Gamma \vdash \mathbf{false} : \mathit{bool}} \qquad \text{(tru)} \frac{}{\Gamma \vdash \mathbf{true} : \mathit{bool}} \end{array}$$

$$\text{(new)} \frac{\forall i = 1..n \Gamma \vdash e_i : \tau_i \quad \text{fields}(c) = \overline{\tau}^n \overline{f}^n \quad \forall i = 1..n \tau_i \leq \tau'_i}{\Gamma \vdash \mathbf{new } c(\overline{e}^n) : c}$$

$$\begin{array}{c} \text{(fld)} \frac{\Gamma \vdash e : c \quad \text{fields}(c) = \overline{\tau}^n \overline{f}^n \quad 1 \leq i \leq n}{\Gamma \vdash e.f_i : \tau_i} \qquad \text{(if)} \frac{\Gamma \vdash e : \mathit{bool} \quad \Gamma \vdash e_1 : \tau_1 \quad \Gamma \vdash e_2 : \tau_2}{\Gamma \vdash \mathbf{if } (e) e_1 \mathbf{ else } e_2 : \forall(\tau_1, \tau_2)} \end{array}$$

$$\begin{array}{c} \forall i = 0..n. \Gamma \vdash e_i : \tau_i \quad \text{this} : \tau_0, \overline{x}^n : \overline{\tau}^n \vdash e : \tau' \\ \text{meth}(\tau_0, m) = \overline{\tau}^n \overline{x}^n . e : \tau \quad \forall i = 1..n. \tau_i \leq \tau'_i, \tau' \leq \tau \\ \text{(co-iv)} \frac{}{\Gamma \vdash e_0 . m(\overline{e}^n) : \tau} \end{array}$$

# Relation between the coinductive TS and BOS

$$\begin{array}{c} \vdots \\ \hline \text{this:M} \Vdash \text{this:M} \quad \text{this:M} \Vdash \mathbf{new} \ L(\text{this.m}):L \\ \hline \text{this:M} \Vdash \text{this.m():L} \\ \hline \text{this:M} \Vdash \mathbf{new} \ L(\text{this.m}):L \\ \hline \end{array} \quad \text{---} \quad \begin{array}{c} \hline \emptyset \Vdash \mathbf{new} \ M():M \\ \hline \end{array} \quad \text{---} \quad \begin{array}{c} \hline \emptyset \Vdash \mathbf{new} \ M().m():L \\ \hline \end{array}$$

$\Vdash =$

# Relation between the coinductive TS and BOS

$$\Rightarrow \nabla = \frac{\frac{\frac{\frac{\frac{\frac{\frac{\Pi \Vdash \text{this} \Rightarrow \text{obj}(M, [])}{\Pi \Vdash \text{new } L(\text{this}.m()) \Rightarrow \mathbb{V}}{\vdots}}{\Pi \Vdash \text{this}.m() \Rightarrow \mathbb{V}}}{\Pi \Vdash \text{new } L(\text{this}.m()) \Rightarrow \text{obj}(L, [n \mapsto \mathbb{V}]}}{\frac{\frac{\emptyset \Vdash \text{new } M() \Rightarrow \text{obj}(M, [])}{\emptyset \Vdash \text{new } M() \Rightarrow \text{obj}(M, [])}}{\emptyset \Vdash \text{new } M().m() \Rightarrow \text{obj}(L, [n \mapsto \mathbb{V}]}}}}}}}}}}$$

# Relation between the coinductive TS and BOS

$$\begin{array}{c}
 \vdots \\
 \hline
 \Pi \Vdash \text{this} \Rightarrow \text{obj}(M, []) \quad \Pi \Vdash \text{new } L(\text{this.m}()) \Rightarrow v \\
 \hline
 \Pi \Vdash \text{this.m}() \Rightarrow v \\
 \hline
 \emptyset \Vdash \text{new } M() \Rightarrow \text{obj}(M, []) \quad \Pi \Vdash \text{new } L(\text{this.m}()) \Rightarrow \text{obj}(L, [n \mapsto v]) \\
 \hline
 \Rightarrow \nabla = \frac{}{\emptyset \Vdash \text{new } M().\text{m}() \Rightarrow \text{obj}(L, [n \mapsto v])}
 \end{array}$$

$\nabla \mathcal{R}_\gamma \Rightarrow \nabla$ : derivation  $\Rightarrow \nabla$  is a concretization of derivation  $\nabla$

# Relationship with the small-step semantics

## Coinductive BOS is sound w.r.t. SOS

If  $\emptyset \Vdash e \Rightarrow v$ ,  $e \xrightarrow{*} e'$ , and  $e'$  is in normal form, then  $e'$  is a value, and  $\emptyset \Vdash e' \Rightarrow v$

Proof: progress + subject reduction

- progress: if  $\emptyset \Vdash e \Rightarrow v$ , then either  $e$  is a value, or there exists  $e'$  s.t.  $e \rightarrow e'$
- subject reduction: if  $\emptyset \Vdash e \Rightarrow v$ , and  $e \rightarrow e'$ , then  $\emptyset \Vdash e' \Rightarrow v$

## Coinductive BOS is not complete w.r.t. SOS

Counter-example: example 4

```
class C extends Object {bool m() {this.m()}}  
if (new C().m()) true.m() else true.m() // main expression
```

there exists **no**  $v$  s.t.  $\emptyset \Vdash \mathbf{if} (\mathbf{new} C().m()) \mathbf{true}.m() \mathbf{else} \mathbf{true}.m() \Rightarrow v$

# Approximating big-step semantics

## Drawbacks of coinductive big-step semantics

Involved proof of soundness based on

- additional equivalent coinductive type system
- complete metric space of big-step semantics proof trees

## Approximating big-step semantics

Support for simpler soundness proof

- no additional type system
- fully inductive type soundness proof
- no coinduction, infinite values, and metric spaces!

# Approximating semantics

$$\begin{aligned} \mathbb{V}, \mathbb{W} &::= \text{obj}(c, [\bar{f}^n \mapsto \bar{v}^n]) \mid \text{false} \mid \text{true} \\ \Pi &::= \bar{x}_i^n \mapsto \bar{v}^n \quad \bar{x}_i^n \text{ distinct} \end{aligned}$$

$$\begin{aligned} (\text{VAR}) \frac{\Pi(x) = \mathbb{V}}{\Pi \vdash x \Rightarrow \mathbb{V}} \quad (\text{FAL}) \frac{}{\Pi \vdash \text{false} \Rightarrow \text{false}} \quad (\text{TRU}) \frac{}{\Pi \vdash \text{true} \Rightarrow \text{true}} \end{aligned}$$

$$\begin{aligned} (\text{NEW}) \frac{\forall i = 1..n. \Pi \vdash e_i \Rightarrow \mathbb{V}_i \quad \text{fields}(c) = \bar{\tau}^n \bar{f}^n}{\Pi \vdash \text{new } c(\bar{e}^n) \Rightarrow \text{obj}(c, [\bar{f}^n \mapsto \bar{v}^n])} \quad (\text{IFT}) \frac{\Pi \vdash e \Rightarrow \text{true} \quad \Pi \vdash e_1 \Rightarrow \mathbb{V}}{\Pi \vdash \text{if } (e) e_1 \text{ else } e_2 \Rightarrow \mathbb{V}} \end{aligned}$$

$$\begin{aligned} (\text{IFF}) \frac{\Pi \vdash e \Rightarrow \text{false} \quad \Pi \vdash e_2 \Rightarrow \mathbb{V}}{\Pi \vdash \text{if } (e) e_1 \text{ else } e_2 \Rightarrow \mathbb{V}} \quad (\text{FLD}) \frac{\Pi \vdash e \Rightarrow \text{obj}(c, [\bar{f}^n \mapsto \bar{v}^n]) \quad 1 \leq i \leq n}{\Pi \vdash e.f_i \Rightarrow \mathbb{V}_i} \end{aligned}$$

$$\begin{aligned} (\text{INV}) \frac{\forall i = 0..n. \Pi \vdash e_i \Rightarrow \mathbb{V}_i \quad \text{this} \mapsto \mathbb{V}_0, \bar{x}^n \mapsto \bar{v}^n \vdash e \Rightarrow \mathbb{V} \\ \mathbb{V}_0 = \text{obj}(c, [\dots]) \quad \text{meth}(c, m) = \bar{\tau}^n \bar{x}^n.e:\tau}{\Pi \vdash e_0.m(\bar{e}^n) \Rightarrow \mathbb{V}} \end{aligned}$$

# Approximating semantics

$$\begin{aligned} \mathbb{V}, \mathbb{W} &::= \text{obj}(c, [\bar{f}^n \mapsto \bar{v}^n]) \mid \text{false} \mid \text{true} \\ \Pi &::= \bar{x}_i^n \mapsto \bar{v}^n \quad \bar{x}_i^n \text{ distinct} \end{aligned}$$

$$\begin{array}{l} \text{(VAR)} \frac{\Pi(x) = \mathbb{V}}{\Pi \vdash x \Rightarrow \mathbb{V}} \qquad \text{(FAL)} \frac{}{\Pi \vdash \text{false} \Rightarrow \text{false}} \qquad \text{(TRU)} \frac{}{\Pi \vdash \text{true} \Rightarrow \text{true}} \end{array}$$

$$\begin{array}{l} \text{(NEW)} \frac{\forall i = 1..n. \Pi \vdash e_i \Rightarrow \mathbb{V}_i \quad \text{fields}(c) = \bar{\tau}^n \bar{f}^n}{\Pi \vdash \text{new } c(\bar{e}^n) \Rightarrow \text{obj}(c, [\bar{f}^n \mapsto \bar{v}^n])} \qquad \text{(IFT)} \frac{\Pi \vdash e \Rightarrow \text{true} \quad \Pi \vdash e_1 \Rightarrow \mathbb{V}}{\Pi \vdash \text{if } (e) e_1 \text{ else } e_2 \Rightarrow \mathbb{V}} \end{array}$$

$$\begin{array}{l} \text{(IFF)} \frac{\Pi \vdash e \Rightarrow \text{false} \quad \Pi \vdash e_2 \Rightarrow \mathbb{V}}{\Pi \vdash \text{if } (e) e_1 \text{ else } e_2 \Rightarrow \mathbb{V}} \qquad \text{(FLD)} \frac{\Pi \vdash e \Rightarrow \text{obj}(c, [\bar{f}^n \mapsto \bar{v}^n]) \quad 1 \leq i \leq n}{\Pi \vdash e.f_i \Rightarrow \mathbb{V}_i} \end{array}$$

$$\begin{array}{l} \forall i = 0..n. \Pi \vdash e_i \Rightarrow \mathbb{V}_i \quad \text{this} \mapsto \mathbb{V}_0, \bar{x}^n \mapsto \bar{v}^n \vdash e \Rightarrow \mathbb{V} \\ \mathbb{V}_0 = \text{obj}(c, [\dots]) \quad \text{meth}(c, m) = \bar{\tau}^n \bar{x}^n.e:\tau \\ \text{(INV)} \frac{}{\Pi \vdash e_0.m(\bar{e}^n) \Rightarrow \mathbb{V}} \qquad \text{(APPROX)} \frac{}{\Pi \vdash e \Rightarrow \mathbb{V}} \end{array}$$

Derived judgment  $\Pi \vdash \overset{\approx}{n} e \Rightarrow \mathbb{V}$

a finite derivation for  $\Pi \vdash e \Rightarrow \mathbb{V}$  where (APPROX) is allowed only at depth  $\geq n$ .



## Example 2 revisited

```
class M extends Object {L m() {new L(this.m())}}
class L extends Object {L n;}
```

## Coinductive semantics

$$\frac{\frac{\frac{\overline{\emptyset \vdash \mathbf{new} M() \Rightarrow \mathit{obj}(M, [ ])}}{\overline{\Pi \vdash \mathbf{new} M() \Rightarrow \mathit{obj}(M, [ ])}} \quad \frac{\frac{\frac{\overline{\Pi \vdash \mathbf{this} \Rightarrow \mathit{obj}(M, [ ])} \quad \frac{\overline{\Pi \vdash \mathbf{new} L(\mathbf{this}.m()) \Rightarrow v_0}}{\vdots}}{\overline{\Pi \vdash \mathbf{this}.m() \Rightarrow v_0}}}{\overline{\Pi \vdash \mathbf{new} L(\mathbf{this}.m()) \Rightarrow v_0}}}{\overline{\emptyset \vdash \mathbf{new} M().m() \Rightarrow v_0}}}$$

where  $\Pi = \mathbf{this} \mapsto \mathit{obj}(M, [ ])$      $v_0 = \mathit{obj}(L, [n \mapsto v_0])$

## Approximating semantics

$$\begin{aligned} \emptyset \vdash_0^{\approx} \mathbf{new} M().m() &\Rightarrow v_{any} \\ \emptyset \vdash_1^{\approx} \mathbf{new} M().m() &\Rightarrow v_{any} \end{aligned}$$



## Example 2 revisited

```
class M extends Object {L m() {new L(this.m()) }}  
class L extends Object {L n;}
```

## Coinductive semantics

$$\frac{\frac{\frac{}{\emptyset \Vdash \mathbf{new}\ M() \Rightarrow \mathit{obj}(M, [])}{} \quad \frac{\frac{\frac{\frac{}{\Pi \Vdash \mathbf{this} \Rightarrow \mathit{obj}(M, [])}{} \quad \frac{\frac{\frac{}{\Pi \Vdash \mathbf{new}\ L(\mathbf{this.m}()) \Rightarrow v_0}{} \quad \vdots}{}{\Pi \Vdash \mathbf{this.m}() \Rightarrow v_0}}{}{\Pi \Vdash \mathbf{new}\ L(\mathbf{this.m}()) \Rightarrow v_0}}{}{\emptyset \Vdash \mathbf{new}\ M().\mathbf{m}() \Rightarrow v_0}}{}}{}}{}}{\emptyset \Vdash \mathbf{new}\ M() \Rightarrow \mathit{obj}(M, [])}$$

where  $\Pi = \mathbf{this} \mapsto \mathit{obj}(M, [])$      $v_0 = \mathit{obj}(L, [n \mapsto v_0])$

## Approximating semantics

$$\emptyset \Vdash_4^{\approx} \mathbf{new}\ M().\mathbf{m}() \Rightarrow \mathit{obj}(L, [n \mapsto \mathit{obj}(L, [n \mapsto v_{\mathit{any}}])])$$
$$\emptyset \Vdash_5^{\approx} \mathbf{new}\ M().\mathbf{m}() \Rightarrow \mathit{obj}(L, [n \mapsto \mathit{obj}(L, [n \mapsto v_{\mathit{any}}])])$$

## Example 2 revisited

```

class M extends Object {L m() {new L(this.m())}}
class L extends Object {L n;}
  
```

## Coinductive semantics

$$\frac{\frac{\frac{\overline{\emptyset \vdash \mathbf{new} M() \Rightarrow \mathit{obj}(M, [])}{\emptyset \vdash \mathbf{new} M() \Rightarrow \mathit{obj}(M, [])} \quad \frac{\frac{\overline{\Pi \vdash \mathbf{new} L(\mathit{this.m}()) \Rightarrow v_0} \quad \overline{\Pi \vdash \mathbf{this} \Rightarrow \mathit{obj}(M, [])}}{\Pi \vdash \mathit{this.m}() \Rightarrow v_0}}{\Pi \vdash \mathbf{new} L(\mathit{this.m}()) \Rightarrow v_0}}{\emptyset \vdash \mathbf{new} M().m() \Rightarrow v_0}}{\emptyset \vdash \mathbf{new} M().m() \Rightarrow v_0}$$

where  $\Pi = \mathit{this} \mapsto \mathit{obj}(M, [])$        $v_0 = \mathit{obj}(L, [n \mapsto v_0])$

## Approximating semantics

$$\begin{aligned}
 \emptyset \vdash_n \mathbf{new} M().m() &\Rightarrow \mathit{obj}(L, [n \mapsto \mathit{obj}(L, [n \mapsto \dots \mapsto v_{\text{any}}])]) \\
 \emptyset \vdash_{n+1} \mathbf{new} M().m() &\Rightarrow \mathit{obj}(L, [n \mapsto \mathit{obj}(L, [n \mapsto \dots \mapsto v_{\text{any}}])])
 \end{aligned}$$

# Relationship with the other semantics

## Equivalence of approx. and coind. big-step semantics

- if  $\emptyset \Vdash e \Rightarrow v$ , then  $\emptyset \Vdash_i^{\approx} e \Rightarrow v$  for all  $i \in \mathbb{N}$  (by def.)
- if  $\emptyset \Vdash_i^{\approx} e \Rightarrow v_i$  for all  $i \in \mathbb{N}$ , then there exists  $v$  s.t.  $\emptyset \Vdash e \Rightarrow v$

the proof exploits the compactness of the metric space of values (standard distance between non-well-founded trees)

## Approximating semantics is sound w.r.t. SOS

if  $\emptyset \Vdash_i^{\approx} e \Rightarrow v_i$  for all  $i \in \mathbb{N}$ ,  $e \xrightarrow{*} e'$ , and  $e'$  is in normal form, then  $e'$  is a value

directly from the soundness of BOS

# Type soundness

## Claim

if  $\emptyset \vdash e : \tau$ , then for all  $k \in \mathbb{N}$  there exists  $v_k$  s.t.  $\emptyset \vdash_k^{\approx} e \Rightarrow v_k$

## Proof

By induction on  $k$  and case analysis on  $e$ .

- $k = 0$ : immediate, since all types are inhabited
- $k > 0$ : sketch for method invocation

$$(inv) \frac{\forall i = 0..n \Gamma \vdash e_i : \tau_i \quad \text{meth}(\tau_0, m) = \overline{\tau'}^n \overline{x}^n . e : \tau \quad \forall i = 1..n \tau_i \leq \tau'_i}{\Gamma \vdash e_0 . m(\overline{e}^n) : \tau}$$

$$(INV) \frac{\forall i = 0..n. \Pi \vdash_{k-1}^{\approx} e_i \Rightarrow v_i \quad \text{this} \mapsto v_0, \overline{x}^n \mapsto \overline{v}^n \vdash_{k-1}^{\approx} e \Rightarrow v \quad v_0 = \text{obj}(c, [\dots]) \quad \text{meth}(c, m) = \overline{\tau}^n \overline{x}^n . e : \tau}{\Pi \vdash_k^{\approx} e_0 . m(\overline{e}^n) \Rightarrow v}$$

# Ongoing and future work

- extension to an imperative oo language
- non determinism
- complete approximating semantics

Questions?

Thank you!