

Type Soundness Proofs with Definitional Interpreters: From $F_{<}$ to Scala

Tiark Rompf (Purdue), Nada Amin (EPFL)

October 3, 2015

IBM

twitter™

Linked in

SONY
make.believe



SIEMENS

xerox 

Novell

 outside.in

 EDF



foursquare™

 UBS

TOMTOM 

amazon™

Office
DEPOT.
*Taking Care
of Business*

guardian.co.uk

OPower
Energy & Environmental Solutions

Bank of America



yammer

Thatcham

Autodesk™

CREDIT SUISSE 

- ▶ Formal description of the Scala type system?
- ▶ Type soundness proof?

People have tried ...

- ▶ 2003: ν Obj
- ▶ 2006: Featherweight Scala
- ▶ 2008: Scalina
- ▶ 2008-2010: ScalaClassic
- ▶ 2012: DOT (Dependent Object Types)

No mechanized soundness proof.

People have tried ...

- ▶ 2003: ν Obj
- ▶ 2006: Featherweight Scala
- ▶ 2008: Scalina
- ▶ 2008-2010: ScalaClassic
- ▶ 2012: DOT (Dependent Object Types)

- ▶ 2014: μ DOT
- ▶ 2015: D2

Why is (was) it so hard?

- ▶ Scala is a rich language
- ▶ We haven't fully understood some of its core features
- ▶ Maybe we aren't doing this right ...

DOT (FOOL'12)

Syntax			
x, y, z	Variable	$L ::=$	Type label
l	Value label	L_c	class label
m	Method label	L_a	abstract type label
$v ::=$	Value	$S, T, U, V, W ::=$	Type
x	variable	$p.L$	type selection
$t ::=$	Terms	$T \{z \Rightarrow \overline{D}\}$	refinement
v	value	$T \wedge T'$	intersection type
$\text{val } x = \text{new } c; t$	new instance	$T \vee T'$	union type
$t.l$	field selection	\top	top type
$t.m(t)$	method invocation	\perp	bottom type
$p ::=$	Path	$S_n, T_n ::=$	Concrete type
x	variable	$p.L :: T_n \{z \Rightarrow \overline{D}\} \mid T_n \wedge T_n \mid \top$	Declaration
$p.l$	selection	$D ::=$	Type declaration
$c ::= T_n \{ \overline{a} \}$	Constructor	$L : S, U$	type declaration
$d ::=$	Initialization	$l : T$	field initialization
$l := v$	field initialization	$m : S \rightarrow U$	method declaration
$m(x) = t$	method initialization	$\Gamma ::= x : T$	Environment
$s ::= \overline{x} \mapsto \overline{v}$	Store		
Reduction			
$t \mapsto t'$			$\{t \mapsto t'\}$
$y \mapsto T_n \{ \overline{t} = \overline{v} \mid m(x) = t \} \in s$	(MSEL)	$\text{val } x = \text{new } c; t \mapsto t \mid s, x \mapsto c$	(NEW)
$y.m(v) \mid s \mapsto [v/x]_s \mid s$			
$y \mapsto T_n \{ \overline{t} = \overline{v} \mid m(x) = t \} \in s$	(SEL)	$t \mid s \mapsto t' \mid s'$	(CONTEXT)
$p.l \mid s \mapsto v_l \mid s$		$e \mid [s] \mapsto e' \mid [s']$	
where evaluation context $e ::= [] \mid e.m(t) \mid t.m(e) \mid e.l$			
Type Assignment			
$\Gamma \vdash t : T$			
$x : T \in \Gamma$	(VAR)	$\Gamma \vdash t \geq t' : T'$	(SEL)
$\Gamma \vdash x : T$		$\Gamma \vdash l.l : T'$	
$\Gamma \vdash t \geq m : S \rightarrow T$		$y \notin \text{fn}(T')$	
$\Gamma \vdash t' : T', T' < S$	(MSEL)	$\Gamma \vdash T \text{ wfe } T_c \prec_v S, U, \overline{D}$	
$\Gamma \vdash t.m(t') : T$		$\Gamma, y : T_n \vdash S < U, \overline{D}, T_n : T'$	
		$\Gamma \vdash \text{val } y = \text{new } T_n \{ \overline{a} \} : T'$	(NEW)
Declaration Assignment			
$\Gamma \vdash d : D$			
$\Gamma \vdash v : V', V' < V$	(VDECL)	$\Gamma, x : S \vdash T : S \text{ wfe } T'$	
$\Gamma \vdash (l = v) : (l : V)$		$\Gamma \vdash m(x) = t : (m : S \rightarrow T)$	(MDECL)

Figure 1. The DOT Calculus : Syntax, Reduction, Type / Declaration Assignment

Well-formed types		$\Gamma \vdash T \text{ wfe}$	
$\Gamma \vdash T \text{ wfe}$		$\Gamma \vdash T \text{ wfe}$	(T-WF)
$\Gamma, z : T \{z \Rightarrow \overline{D}\} \vdash T \text{ wfe}$	(RPN-WF)	$\Gamma \vdash \perp \text{ wfe}$	(L-WF)
$\Gamma \vdash T \{z \Rightarrow \overline{D}\} \text{ wfe}$		$\Gamma \vdash p.L \text{ wfe}$	(TSEL-WF ₁)
$\Gamma \vdash p \geq L : S, U' : S, U' < U, \text{ wfe } U$	(TSEL-WF ₁)	$\Gamma \vdash p.L \text{ wfe}$	(TSEL-WF ₂)
$\Gamma \vdash p.L \text{ wfe}$		$\Gamma \vdash T \text{ wfe}, T' < T$	(V-WF)
$\Gamma \vdash T \wedge T' \text{ wfe}$	(A-WF)	$\Gamma \vdash T \wedge T' \text{ wfe}$	
Well-formed declarations		$\Gamma \vdash D \text{ wfe}$	
$\Gamma \vdash S \text{ wfe}, U \text{ wfe}$	(TDECL-WF)	$\Gamma \vdash S \text{ wfe}, U \text{ wfe}$	(MDECL-WF)
$\Gamma \vdash L : S, U \text{ wfe}$		$\Gamma \vdash m : S \rightarrow U \text{ wfe}$	
$\Gamma \vdash T \text{ wfe}$	(VDECL-WF)		
$\Gamma \vdash l : T \text{ wfe}$			

$\text{dom}(\overline{D} \wedge \overline{D}') = \text{dom}(\overline{D}) \cap \text{dom}(\overline{D}')$	
$\text{dom}(\overline{D} \vee \overline{D}') = \text{dom}(\overline{D}) \cup \text{dom}(\overline{D}')$	
$(D \wedge D')(L) = L : (S \vee S') \rightarrow (U \wedge U')$	if $(L : S, U) \in \overline{D}$ and $(L : S', U') \in \overline{D}'$
$(D \wedge D')(L) = D'(L)$	if $L \notin \text{dom}(\overline{D})$
$(D \wedge D')(m) = m : (S \vee S') \rightarrow (U \wedge U')$	if $(m : S \rightarrow U) \in \overline{D}$ and $(m : S' \rightarrow U') \in \overline{D}'$
$(D \wedge D')(l) = D(m)$	if $m \notin \text{dom}(\overline{D})$
$(D \wedge D')(l) = D'(m)$	if $m \notin \text{dom}(\overline{D})$
$(D \wedge D')(l) = l : T \wedge T'$	if $(l : T) \in \overline{D}$ and $(l : T') \in \overline{D}'$
$(D \wedge D')(l) = D(l)$	if $l \notin \text{dom}(\overline{D})$
$(D \wedge D')(l) = D'(l)$	if $l \notin \text{dom}(\overline{D})$
$(D \vee D')(L) = L : (S \wedge S') \rightarrow (U \vee U')$	if $(L : S, U) \in \overline{D}$ and $(L : S', U') \in \overline{D}'$
$(D \vee D')(m) = m : (S \wedge S') \rightarrow (U \vee U')$	if $(m : S \rightarrow U) \in \overline{D}$ and $(m : S' \rightarrow U') \in \overline{D}'$
$(D \vee D')(l) = l : T \vee T'$	if $(l : T) \in \overline{D}$ and $(l : T') \in \overline{D}'$

Sets of declarations form a lattice with the given meet \wedge and join \vee , the empty set of declarations as the top element, and the bottom element \overline{D}_\perp . Here \overline{D}_\perp is the set of declarations that contains for every term label l the declaration $l : \perp$, for every type label L the declaration $L : \perp$, and for every method label m the declaration $m : \top \rightarrow \perp$.

Figure 2. The DOT Calculus : Declaration Lattice

Membership		$\Gamma \vdash t \geq D$	
$\Gamma \vdash p : T, T \prec \overline{D}$	(PATH- \geq)	$z \notin \text{fn}(D_1)$	$\Gamma \vdash t : T, T \prec \overline{D}$
$\Gamma \vdash p \geq [p/s]_{D_1}$		$\Gamma \vdash t \geq D_1$	(TERM- \geq)
Expansion		$\Gamma \vdash T \prec \overline{D}$	
$\Gamma \vdash T \prec \overline{D}$		$\Gamma \vdash p \geq L : S, U : U \prec \overline{D}$	(TSEL- \prec)
$\Gamma \vdash T \{z \Rightarrow \overline{D}\} \prec \overline{D} \wedge \overline{D}'$	(RPN- \prec)	$\Gamma \vdash T_1 \prec \overline{D}_1, T_2 \prec \overline{D}_2$	(A- \prec)
$\Gamma \vdash T_1 \prec \overline{D}_1, T_2 \prec \overline{D}_2$		$\Gamma \vdash T_1 \wedge T_2 \prec \overline{D}_1 \wedge \overline{D}_2$	(V- \prec)
$\Gamma \vdash T_1 \wedge T_2 \prec \overline{D}_1 \wedge \overline{D}_2$		$\Gamma \vdash T_1 \prec \overline{D}_1, T_2 \prec \overline{D}_2$	(L- \prec)
$\Gamma \vdash T \prec \{\}$	(T- \prec)	$\Gamma \vdash \perp \prec \overline{D}_\perp$	

Figure 3. The DOT Calculus : Membership and Expansion

Subtyping		$\Gamma \vdash S < T$	
$\Gamma \vdash T \text{ wfe}$		$\Gamma \vdash T \text{ wfe}$	(L- $<$)
$\Gamma \vdash T < T$	(REFL)	$\Gamma \vdash \perp < T$	
$\Gamma \vdash T \{z \Rightarrow \overline{D}\} \text{ wfe}, S < T, T \prec \overline{D}'$		$\Gamma \vdash T \{z \Rightarrow \overline{D}\} \text{ wfe}, T < T'$	(RPN- $<$)
$\Gamma, z : S \vdash \overline{D} < \overline{D}'$	(A-RPN)	$\Gamma \vdash T \{z \Rightarrow \overline{D}\} < T'$	
$\Gamma \vdash S < T \{z \Rightarrow \overline{D}\}$		$\Gamma \vdash T \{z \Rightarrow \overline{D}\} < T'$	
$\Gamma \vdash p \geq L : S, U' : S < U, S' < S$	(A-TSEL)	$\Gamma \vdash S < p.L$	(TSEL- $<$)
$\Gamma \vdash S < p.L$		$\Gamma \vdash T_1 < T_1, T_2 < T_2$	(V-WF)
$\Gamma \vdash T < T_1, T < T_2$		$\Gamma \vdash T_1 < T_1 \wedge T_2$	(A- $<$)
$\Gamma \vdash T_1 < T_1 \wedge T_2$		$\Gamma \vdash T_1 < T_1, T_2 < T_2$	(V- $<$)
$\Gamma \vdash T_2 \text{ wfe}, T < T_1$		$\Gamma \vdash T_1 \wedge T_2 < T$	(A- $<$)
$\Gamma \vdash T < T_1 \vee T_2$		$\Gamma \vdash T_1 < T_1, T_2 < T_2$	(A- $<$)
$\Gamma \vdash T_1 \text{ wfe}, T < T_2$		$\Gamma \vdash T_1 \wedge T_2 < T$	(A- $<$)
$\Gamma \vdash T < T_1 \vee T_2$		$\Gamma \vdash T_1 < T_1, T_2 < T_2$	(A- $<$)
$\Gamma \vdash T \text{ wfe}$		$\Gamma \vdash T < T$	(A- $<$)
$\Gamma \vdash T < T$		$\Gamma \vdash T_1 < T_1, T_2 < T_2$	(A- $<$)

Declaration subsumption

$\Gamma \vdash D < D'$

What worked:

- ▶ Bottom up instead of top down
- ▶ Focus on the core:
path-dependent types and objects with (recursive) type members
- ▶ Look beyond rewriting semantics

Types in Scala

Modules, Objects, Functions

```
object listModule {  
  trait List[Elem] = {  
    def head(): Elem  
    def tail(): List[Elem]  
  }  
  def nil() = new List[Nothing] {  
    def head() = error()  
    def tail() = error()  
  }  
  def cons[T](hd: T)(tl: List[T]) = new List[T] {  
    def head() = hd  
    def tail() = tl  
  }  
}
```

Types in Scala

'modular' named type `scala.collection.BitSet`
 compound type `Channel with Logged`
 refined type `Channel { def close(): Unit }`

'functional' parameterized type `List[String]`
 existential type `List[T] forSome { type T }`
 higher-kinded type `List`

Reducing Functional to Modular

- ▶ type parameter to type member

```
class List[Elem] {} /*vs*/ class List { type Elem }
```

- ▶ parameterized type to refined type

```
List[String] /*vs*/ List { type Elem = String }
```

- ▶ existential type?

```
List[T] forSome { type T } /*vs*/ List
```

- ▶ higher-kinded type?

```
List /*vs*/ List
```

Modules, Objects, Functions

```
val listModule = new { m =>
  type List = { this =>
    type Elem
    def head(): this.Elem
    def tail(): m.List & { type Elem <: this.Elem }
  }
  def nil() = new { this =>
    type Elem = Nothing
    def head() = error()
    def tail() = error()
  }
  def cons[T](hd:T)(tl:m.List & { type Elem <: T }) = new { this =>
    type Elem <: T
    def head() = hd
    def tail() = tl
  }
}
```

Nominality by Ascription

```
type ListAPI = { m =>
  type List <: { this =>
    type Elem
    def head(): this.Elem
    def tail(): m.List & { type Elem <: this.Elem }
  }
  def nil(): List & { type Elem = Bot }
  def cons[T]: T =>
    m.List & { type Elem <: T } =>
      m.List & { type Elem <: T }
}
```

Types in D2 (inspired by DOT)

types S, T, U

path-dependent type p.L

recursive self type { z => T }

intersection T & T

union T | T

top Any

bottom Nothing

type declaration type L: S .. U

field declaration val l: U

method declaration def m(x: S): U

Subtyping

$$\boxed{\Gamma \vdash S <: U}$$

$$\frac{\Gamma \vdash x : (\mathbf{type} \ L : S..U) , \ S' <: S}{\Gamma \vdash S' <: x.L} \quad (<:-\text{TSEL})$$

$$\frac{\Gamma \vdash x : (\mathbf{type} \ L : S..U) , \ U <: U'}{\Gamma \vdash x.L <: U'} \quad (\text{TSEL-}<:)$$

$$\frac{\Gamma, z : T \vdash T <: T'}{\Gamma \vdash \{z \Rightarrow T\} <: \{z \Rightarrow T'\}} \quad (\text{REC-}<:-\text{REC})$$

$$\frac{\Gamma \vdash S' <: S , \ U <: U'}{\Gamma \vdash (\mathbf{type} \ L : S..U) <: (\mathbf{type} \ L : S'..U')} \quad (\text{TMEM-}<:-\text{TMEM})$$

Challenge: Type Preservation

Trouble: Type Preservation

```
trait Brand {  
  type Hidden  
  def pack(x: Int): Hidden  
  def unpack(x: Hidden): Int  
}  
  
val brand: Brand = new Brand {  
  type Hidden = Int  
  def pack(x: Int): Hidden = x  
  def unpack(x: Hidden): Int = x  
}  
  
brand.unpack(brand.pack(7)) // ok  
brand.unpack(7) // not ok -- but occurs during reduction!
```

Trouble: \perp and Intersections

Type \perp is a subtype of all other types, including `{ type E = Int }` and `{ type E = String }`.

So if $p: \perp$ we have $\text{Int} <: p.E$ and $p.E <: \text{String}$.

Transitivity would give us $\text{Int} <: p.E <: \text{String}$!

Subtyping lattice collapses.

Adding intersection types is equivalent to bottom (bad bounds!)

Key Observation

- ▶ Bottom types do not occur at runtime!
- ▶ It is enough to have transitivity and narrowing on runtime environments
- ▶ Have a (restrictive) static type system and a (lenient) one during evaluation

Definitional Interpreters

STLC Static Semantics

Syntax

$$\begin{aligned} T &::= X \mid \top \mid T \rightarrow T \\ t &::= x \mid \lambda x : T. t \mid t \ t \\ \Gamma &::= \emptyset \mid \Gamma, x : T \end{aligned}$$

Subtyping

$\Gamma \vdash S <: U$

$$\frac{\Gamma \vdash S <: T \quad \Gamma \vdash T_1 <: S_1, S_2 <: T_2}{\Gamma \vdash S_1 \rightarrow S_2 <: T_1 \rightarrow T_2}$$

Type assignment

$\Gamma \vdash t : T$

$$\frac{\Gamma \ni x : T}{\Gamma \vdash x : T}$$
$$\frac{\Gamma, x : T_1 \vdash t_2 : T_2}{\Gamma \vdash \lambda x : T_1. t_2 : T_1 \rightarrow T_2}$$
$$\frac{\Gamma \vdash t_1 : T_1 \rightarrow T_2, t_2 : T_1}{\Gamma \vdash t_1 t_2 : T_2}$$
$$\frac{\Gamma \vdash t : S, S <: T}{\Gamma \vdash t : T}$$

STLC

```
Fixpoint eval(n: nat)(env: venv)(t: tm){struct n}:
option (option vl) :=
  DO n1 <== FUEL n;                                (* totality *)
  match t with
  | tcst c      => DONE VAL (vcst c)                 (* constant *)
  | tvar x      => DONE (lookup x env)               (* variable *)
  | tabs y      => DONE VAL (vabs env x ey)          (* lambda *)
  | tapp ef ex  =>                                    (* application *)
    DO vf <== eval n1 env ef;
    DO vx <== eval n1 env ex;
    match vf with
    | (vabs env2 x ey) =>
      eval n1 ((x,vx)::env2) ey
    | _ => ERROR
    end
  end.
```

Soundness

$$\frac{\Gamma \vdash e : T \quad \Gamma \models H \quad \text{eval } n \ H \ e = \text{Done } r}{r = \text{Val } v \quad v : T}$$

$F_{<}$: Static Semantics

Syntax

$$\begin{aligned} X &::= Y \mid Z \\ T &::= X \mid \top \mid T \rightarrow T \mid \forall Z <: T. T^Z \\ t &::= x \mid \lambda x : T. t \mid \Lambda Y <: T. t \mid t \ t \mid t \ [T] \\ \Gamma &::= \emptyset \mid \Gamma, x : T \mid \Gamma, X <: T \end{aligned}$$

Subtyping

$$\boxed{\Gamma \vdash S <: U}$$

$$\begin{aligned} &\Gamma \vdash S <: T \\ &\Gamma \vdash X <: X \\ \hline &\Gamma \ni X <: U \quad \Gamma \vdash U <: T \\ &\Gamma \vdash X <: T \\ \hline &\Gamma \vdash T_1 <: S_1, S_2 <: T_2 \\ &\Gamma \vdash S_1 \rightarrow S_2 <: T_1 \rightarrow T_2 \\ \hline &\Gamma \vdash T_1 <: S_1 \\ &\Gamma, Z <: T_1 \vdash S_2^Z <: T_2^Z \\ \hline &\Gamma \vdash \forall Z <: S_1. S_2^Z <: \forall Z <: T_1. T_2^Z \end{aligned}$$

Type assignment

$$\boxed{\Gamma \vdash t : T}$$

$$\begin{aligned} &\Gamma \ni x : T \\ \hline &\Gamma \vdash x : T \\ \hline &\Gamma, x : T_1 \vdash t_2 : T_2 \\ \hline &\Gamma \vdash \lambda x : T_1. t_2 : T_1 \rightarrow T_2 \\ \hline &\Gamma \vdash t_1 : T_1 \rightarrow T_2, t_2 : T_1 \\ \hline &\Gamma \vdash t_1 t_2 : T_2 \\ \hline &\Gamma, Y <: T_1 \vdash t_2 : T_2^Y \\ \hline &\Gamma \vdash \Lambda Y <: T_1. t_2 : \forall Z <: T_1. T_2^Z \\ \hline &\Gamma \vdash t_1 : \forall Z <: T_{11}. T_{12}^Z, T_2 <: T_{11} \\ \hline &\Gamma \vdash t_1 [T_2] : T_{12}^{T_2} \\ \hline &\Gamma \vdash t : S, S <: T \\ \hline &\Gamma \vdash t : T \end{aligned}$$

$F_{<} \text{ — ?}$

$$(\Lambda X <: T. t)[T] \longrightarrow t[T/X]$$

$F_{<}: \text{--- ?}$

```
Fixpoint eval(n: nat)(env: venv)(t: tm){struct n}:
option (option vl) :=
  DO n1 <== FUEL n;                                (* totality *)
  match t with
  | tcst c      => DONE VAL (vcst c)                  (* constant *)
  | tvar x      => DONE (lookup x env)                (* variable *)
  | tabs y      => DONE VAL (vabs env x ey)           (* lambda *)
  | ttabs y     => DONE VAL (vtabs env x ey)          (* forall *)
  | tapp ef ex  =>                                    (* application *)
    DO vf <== eval n1 env ef;
    DO vx <== eval n1 env ex;
    match vf with
    | (vabs env2 x ey) =>
      eval n1 ((x,vx)::env2) ey
    | _ => ERROR
    end
  | ttapp ef T =>
    DO vf <== eval n1 env ef;
    match vf with
    | (vtabs env2 x ey) => eval n1 env2 (substitute ey x T)
    | _ => ERROR
    end
  end
end.
```

$F_{<}$: — ?

```
Fixpoint eval(n: nat)(env: venv)(t: tm){struct n}:
option (option vl) :=
  DO n1 <== FUEL n;                                (* totality *)
  match t with
  | tcst c      => DONE VAL (vcst c)                  (* constant *)
  | tvar x      => DONE (lookup x env)                (* variable *)
  | tabs y      => DONE VAL (vabs env x ey)           (* lambda *)
  | ttabs y     => DONE VAL (vtabs env x ey)          (* forall *)
  | tapp ef ex  =>                                    (* application *)
    DO vf <== eval n1 env ef;
    DO vx <== eval n1 env ex;
    match vf with
    | (vabs env2 x ey) =>
      eval n1 ((x,vx)::env2) ey
    | _ => ERROR
    end
  | ttapp ef T =>
    DO vf <== eval n1 env ef;
    match vf with
    | (vtabs env2 x ey) => eval n1 ((x,T)::env2) ey
    | _ => ERROR
    end
  end
end.
```

F_<:

```
Fixpoint eval(n: nat)(env: venv)(t: tm){struct n}:
option (option vl) :=
  DO n1 <== FUEL n;                                (* totality *)
  match t with
  | tcst c      => DONE VAL (vcst c)                 (* constant *)
  | tvar x      => DONE (lookup x env)                (* variable *)
  | tabs y      => DONE VAL (vabs env x ey)           (* lambda *)
  | ttabs y     => DONE VAL (vtabs env x ey)          (* forall *)
  | tapp ef ex  =>                                   (* application *)
    DO vf <== eval n1 env ef;
    DO vx <== eval n1 env ex;
    match vf with
    | (vabs env2 x ey) =>
      eval n1 ((x,vx)::env2) ey
    | _ => ERROR
    end
  | ttapp ef T =>
    DO vf <== eval n1 env ef;
    match vf with
    | (vtabs env2 x ey) => eval n1 ((x,vty env T)::env2) ey
    | _ => ERROR
    end
  end
end.
```

F_<: Dynamics

Syntax

$$\begin{aligned}
v &::= \langle H, \lambda x : T. t \rangle \mid \langle H, \Lambda Y <: T. t \rangle \\
H &::= \emptyset \mid H, x : v \mid H, Y = \langle H, T \rangle \\
J &::= \emptyset \mid J, Z <: \langle H, T \rangle
\end{aligned}$$

Runtime subtp.

$$J \vdash H_1 T_1 <: H_2 T_2$$

$$J \vdash H_1 T <: H_2 \top$$

$$\frac{J \vdash H_2 T_1 <: H_1 S_1 \quad J \vdash H_1 S_2 <: H_2 T_2}{J \vdash H_1 S_1 \rightarrow S_2 <: H_2 T_1 \rightarrow T_2}$$

$$\frac{
\begin{array}{l}
J \vdash H_2 T_1 <: H_1 S_1 \\
J, Z <: \langle H_2, T_1 \rangle \vdash H_1 S_2^Z <: H_2 T_2^Z
\end{array}
}{J \vdash H_1 \forall Z <: S_1. S_2^Z <: H_2 \forall Z <: T_1. T_2^Z}$$

Abstract type variables

$$\frac{
\begin{array}{l}
J \vdash H_1 Z <: H_2 Z \\
J \ni Z <: \langle H, U \rangle \quad J \vdash H U <: H_2 T
\end{array}
}{J \vdash H_1 Z <: H_2 T}$$

Concrete type variables

$$\frac{H_1 \ni Y_1 = \langle H, T \rangle \quad H_2 \ni Y_2 = \langle H, T \rangle}{J \vdash H_1 Y_1 <: H_2 Y_2}$$

$$\frac{H_1 \ni Y = \langle H, U \rangle \quad J \vdash H U <: H_2 T}{J \vdash H_1 Y <: H_2 T}$$

$$\frac{J \vdash H_1 T <: H L \quad H_2 \ni Y = \langle H, L \rangle}{J \vdash H_1 T <: H_2 Y}$$

Transitivity

$$\frac{J \vdash H_1 T_1 <: H_2 T_2 \quad H_2 T_2 <: H_3 T_3}{J \vdash H_1 T_1 <: H_3 T_3}$$

F_<: Dynamics

Value type assignment

$$H \vdash v : T$$

Consistent environments

$$\Gamma \models H J$$

$$\frac{\Gamma \models H \emptyset \quad \Gamma, x : T_1 \vdash t : T_2}{H \vdash \langle H, \lambda x : T_1. t \rangle : T_1 \rightarrow T_2}$$

$$\frac{\Gamma \models H \emptyset \quad \Gamma, Y <: T_1 \vdash t : T_2^Y}{H \vdash \langle H, \lambda Y <: T_1. t \rangle : \forall Z <: T_1. T_2^Z}$$

$$\frac{H_1 \vdash v : T_1 \quad \emptyset \vdash H_1 T_1 <: H_2 T_2}{H_2 \vdash v : T_2}$$

$$\emptyset \models \emptyset \emptyset$$

$$\frac{\Gamma \models H J \quad H \vdash v : T}{\Gamma, x : T \models (H, x : v) J}$$

$$\frac{\Gamma \models H J \quad J \vdash H_1 T_1 <: H T}{\Gamma, Y <: T \models (H, Y = \langle H_1, T_1 \rangle) J}$$

$$\frac{\Gamma \models H J \quad J \vdash H_1 T_1 <: H T}{\Gamma, Z <: T \models H (J, Z <: \langle H_1, T_1 \rangle)}$$

Challenge: Transitivity

$$\frac{J \vdash H_1 T_1 <: H Y, J \vdash H Y <: H_2 T_2}{J \vdash H_1 T_1 <: H_2 T_2} \quad (<:-\text{TRANS})$$

Challenge: Transitivity

- ▶ Attempt 1: induction on subtyping derivation
problem: contravariant positions
- ▶ Attempt 2: induction on middle type in $T_1 <: T_2 <: T_3$ (like $F_{<:}$)
problem: $T_1 <: Y <: T_2$
- ▶ Attempt 3: induction on a well-formedness witness
problem: cyclicity
- ▶ Attempt 4: add transitivity axiom
problem: inversion lemma
- ▶ Solution: transitivity axiom plus “push-back”: eliminate uses of axiom after the fact

$F_{<}$: Extended to Path-Dependent Types

Syntax

$$\begin{aligned} T &::= x.\text{Type} \mid \top \mid (z : T) \rightarrow T^z \mid \{\text{Type} = T\} \\ t &::= x \mid \lambda x : T.t \mid t \ t \mid \{\text{Type} = T\} \\ v &::= \langle H, \lambda x : T.t \rangle \mid \langle H, T \rangle \\ G &::= \emptyset \mid G, x : T \\ H &::= \emptyset \mid H, x : v \\ J &::= \emptyset \mid H, z : \langle H, T \rangle \end{aligned}$$

Runtime Subtp.

$J \vdash H_1 \ T_1 <: H_2 \ T_2$

$$\frac{J \vdash H_1 \ T_1 <: H_2 \ T_2 \quad J \vdash H_2 \ T_2 <: H_1 \ T_1}{J \vdash H_1 \ \{\text{Type} = T_1\} <: H_2 \ \{\text{Type} = T_2\}}$$

Subtyping

$\Gamma \vdash S <: U$

$$\frac{\Gamma \vdash T_1 <: T_2 \quad \Gamma \vdash T_2 <: T_1}{\Gamma \vdash \{\text{Type} = T_1\} <: \{\text{Type} = T_2\}}$$

$$\frac{\Gamma(x) = U \quad \Gamma \vdash U <: \{\text{Type} = T\}}{\Gamma \vdash x.T <: T}$$

$$\frac{J(z) = \langle H, U \rangle \quad J \vdash H \ U <: H_2 \ \{\text{Type} = T\}}{J \vdash H_1 \ z.\text{Type} <: H_2 \ T}$$

$$\frac{H_1(x) = v \quad H \vdash v : U \quad J \vdash H \ U <: H_2 \ \{\text{Type} = T\}}{J \vdash H_1 \ x.\text{Type} <: H_2 \ T}$$

$F_{<}$: Extended to Path-Dependent Types

Type assignment

$$\boxed{\Gamma \vdash t : T}$$

$$\frac{\begin{array}{c} \Gamma \vdash \{\text{Type} = T\} : \{\text{Type} = T\} \\ \Gamma, x : T_1 \vdash t_2 : T_2^x \\ \hline \Gamma \vdash \lambda x : T_1. t_2 : (z : T_1) \rightarrow T_2^z \\ \Gamma \vdash t_1 : (z : T_1) \rightarrow T_2^z, \quad t_2 : T_1 \\ \hline \Gamma \vdash t_1 t_2 : T_2 \end{array}}$$

Value type assignment

$$\boxed{H \vdash v : T}$$

$$\frac{\Gamma \models H \emptyset \quad \Gamma \vdash \{\text{Type} = T\} : \{\text{Type} = T\}}{H \vdash \langle H, T \rangle : \{\text{Type} = T\}}$$

$$\frac{\Gamma \models H \emptyset \quad \Gamma, x : T_1 \vdash t : T_2^x}{H \vdash \langle H, \lambda x : T_1. t \rangle : (z : T_1) \rightarrow T_2^z}$$

Further Implemented Extensions

- ▶ Subtyping lattice (intersections and unions)
- ▶ Objects with multiple type and method members
- ▶ Recursive types $\{ x \Rightarrow T \}$
- ▶ Key technique: more intricate transitivity pushback

Thank you!

Type Inference

Type Inference in Scala (Least Upper Bound?)

```
trait A { type T <: A }
trait B { type T <: B }
trait C extends A with B { type T <: C }
trait D extends A with B { type T <: D }
// in Scala, lub(C, D) is an infinite sequence
A with B { type T <: A with B { type T <: ... } }
```



```
// in Scala REPL
> val o = if (true) (new C{}) else (new D{})
o: A with B{type T <: A with B} = ...
> val o:A with B{type T<:A with B {type T<:A with B}} =
    if (true) (new C{}) else (new D{})
o: A with B{type T <: A with B{type T <: A with B}} = ...
```


Type Inference in Scala (Working too Hard too Soon?)

[illegible]

Type Inference in Scala (Working too Hard too Soon?)

```
import scala.collection.mutable.{Map => MMap, Set => MSet}
val ms: MMap[Int, MSet[Int]] = MMap.empty

// in Scala REPL
> if (!ms.contains(1)) ms += 1 -> MSet(1) else ms(1) += 1
res0: ... (796 characters) ...
> :t res0
: ... (21481 characters) ...
```

- ▶ Inspired by a bug report
SI-5862: very slow compilation due to humonguous LUB.
- ▶ The character lengths reported are for Scala 2.11 (*after* the fix).
- ▶ In Dotty, type inference can be lazy thanks to the native unions (for least upper bounds) and intersections (for greatest lower bounds) of the core calculus (DOT).