

Type inference for functional interfaces in Java 8

Martin Plümicke

Baden-Wuerttemberg Cooperative State University
Stuttgart/Horb

6. Mai 2013

Overview

Introduction

Type inference algorithm TI

- The function TYPE

 - The sub-function TYPEExpr

 - The sub-function TYPEStmt

- The function SOLVE

- The whole algorithm TI

Comparison to WTYPE for Java λ

Summary and Outlook

Two approaches [Plümicke, Bad Honnef 2012]

- ▶ **Martin's approach: Java_λ** [Project Lambda 2010, version 0.1.5; Martin Plümicke, PPPJ 2011]
 - ▶ closures (λ -expressions)
 - ▶ function types
 - ▶ higher-order functions
- ▶ **Brian's approach: Java 8** [Project Lambda 2011 version 0.4.2, Brian Goetz 2011]
 - ▶ closures (λ -expressions)
 - ▶ functional interfaces
 - ▶ method references
 - ▶ default methods

Function types vs. Functional interfaces

Function types (Java_λ)

```
void doAddition(#int(int, int) o)  
{ ... }
```

Functional interfaces (Java 8)

```
interface Operation {  
    public int op (int x, int y);  
}
```

```
void doAddition(Operation o)  
{ ... }
```

Functional interfaces as compatible target types

A lambda expression is *compatible* with a type T , if

- ▶ T is a **functional interface** type
- ▶ The lambda expression has the **same number** of parameters as T 's method, and those **parameters' types are the same**
- ▶ Each **expression returned** by the lambda body is **compatible** with T 's method's return type
- ▶ Each **exception** thrown by the lambda body is **allowed by T 's method's** throws clause

Functional interfaces as compatible target types

A lambda expression is *compatible* with a type T , if

- ▶ T is a **functional interface** type
- ▶ The lambda expression has the **same number** of parameters as T 's method, and those **parameters' types are the same**
- ▶ Each **expression returned** by the lambda body is **compatible** with T 's method's return type
- ▶ Each **exception** thrown by the lambda body is **allowed by T 's method's** throws clause

Lemma: There is an **equivalence class** of compatible target types for a lambda expression.

Canonical representation

For the **equivalence class** of the compatible target types of a lambda expression, there is a **canonical representation**

$$\text{Fun}N\langle R, T_1, \dots, T_N \rangle$$

with

```
interface FunN<R,T1, ..., TN>
  { R apply(T1 arg1 , ..., TN argN); }
```

if the type of the single method of a compatible target type is

$$(T_1, \dots, T_N) \rightarrow R$$

Matrices in Java 8

```
class Matrix extends Vector<Vector<Integer>> {  
    //Matrix -> ((Matrix, Matrix) -> Matrix) -> Matrix  
    Fun1<Fun1<Matrix, Fun2<Matrix, Matrix,Matrix>>, Matrix>  
    op = (Matrix m) -> (Fun2<Matrix, Matrix,Matrix> f) ->  
        f.apply(this, m);  
}
```


Matrices in Java 8

```
class Matrix extends Vector<Vector<Integer>> {  
    //Matrix -> ((Matrix, Matrix) -> Matrix) -> Matrix  
    Fun1<Fun1<Matrix, Fun2<Matrix, Matrix,Matrix>>, Matrix>  
    op = (Matrix m) -> (Fun2<Matrix, Matrix,Matrix> f) ->  
        f.apply(this, m);  
  
    //(Matrix, Matrix) -> Matrix  
    Fun2<Matrix, Matrix,Matrix>  
    mul = (Matrix m1, Matrix m2) -> {  
        Matrix ret = new Matrix ();  
        ...           //matrix multiplication  
        return ret; }  
}
```

Matrices in Java 8

```

class Matrix extends Vector<Vector<Integer>> {
    //Matrix -> ((Matrix, Matrix) -> Matrix) -> Matrix
    Fun1<Fun1<Matrix, Fun2<Matrix, Matrix,Matrix>>, Matrix>
    op = (Matrix m) -> (Fun2<Matrix, Matrix,Matrix> f) ->
        f.apply(this, m);

    //(Matrix, Matrix) -> Matrix
    Fun2<Matrix, Matrix,Matrix>
    mul = (Matrix m1, Matrix m2) -> {
        Matrix ret = new Matrix ();
        ... //matrix multiplication
        return ret; }

    public static void main(String[] args) {
        Matrix m1 = new Matrix(...);
        Matrix m2 = new Matrix(...);
        (m1.op.apply(m2)).apply(m1.mul);} } }

```

Goal

```
class Matrix extends Vector<Vector<Integer>> {
```

```
    op = (m) -> (f) ->  
        f.apply(this, m);
```

```
    mul = (m1, m2) -> {  
        ret = new Matrix ();  
        ... //matrix multiplication  
        return ret; }  
}
```

```
public static void main(String[] args) {  
    Matrix m1 = new Matrix(...);  
    Matrix m2 = new Matrix(...);  
    (m1.op.apply(m2)).apply(m1.mul); } } }
```

The language

Source := *class**
class := Class(*stype*, [extends(*stype*),] *FieldDecl**)
FieldDecl := Field([*type*,] *var*[, *expr*])
block := Block(*stmt**)
stmt := *block* | Return(*expr*) | While(*bexpr*, *block*)
| LocalVarDecl(*var*[, *type*]) | If(*bexpr*, *block*[, *block*])
| *stmtexpr*
lambdaexpr := Lambda(((*var*[, *type*]))*, (*stmt* | *expr*))
stmtexpr := Assign(*var*, *expr*) | New(*stype*, *expr**)
| MethodCall(*expr*, *apply*, *expr**)
vexpr := LocalVar(*var*) | InstVar(*iexpr*, *var*)
iexpr = *vexpr* | *stmtexpr* | Cast(*type*, *iexpr*) | this | super
expr := *lambdaexpr* | *iexpr* | *bexp* | *sexp*

Sketch of the algorithm TI

TYPE: Introduces fresh type variables and collects the constraints

SOLVE: Solves the constraints by type unification

Data-structures

► **Set of type assumptions:**

$v : \theta$: Assumptions for fields or local variables of the actual class.

$\tau.v : \theta$: Assumptions for fields of the class τ .

$\text{Fun } N \langle R, T_1, \dots, T_N \rangle . \text{apply} : (T_1, \dots, T_N) \rightarrow R$

Data-structures

► **Set of type assumptions:**

$v : \theta$: Assumptions for fields or local variables of the actual class.

$\tau.v : \theta$: Assumptions for fields of the class τ .

$\text{Fun } N \langle R, T1, \dots, TN \rangle . \text{apply} : (T1, \dots, TN) \rightarrow R$

► **Set of constraints:**

$$\{ \theta R \theta' \mid R \in \{ \triangleleft, \triangleleft?, \doteq \} \}$$

TYPE: $\text{TypeAssumptions} \times \text{Class} \rightarrow \text{TClass} \times \text{ConstraintsSet}$

TYPE: $\text{TypeAssumptions} \times \text{Class} \rightarrow \text{TClass} \times \text{ConstraintsSet}$

TYPE(*Ass*, $\text{Class}(\tau, \text{extends}(\tau'), \text{fdecls})$) =

let

fdecls = [Field(f_1, lexpr_1), ..., Field(f_n, lexpr_n)]

fypeass = { $\text{this}.f_i : a_i \mid a_i$ fresh type variables }

∪ { $\text{this} : \tau, \text{super} : \tau' \}$

∪ { *visible types of fields of τ'* }

AssAll = *Ass* ∪ *fypeass*

TYPE: $\text{TypeAssumptions} \times \text{Class} \rightarrow \text{TClass} \times \text{ConstraintsSet}$

TYPE(*Ass*, $\text{Class}(\tau, \text{extends}(\tau'), \text{fdecls})$) =

let

fdecls = [Field(f_1, lexpr_1), ..., Field(f_n, lexpr_n)]

fypeass = { $\text{this}.f_i : a_i \mid a_i$ fresh type variables }

∪ { $\text{this} : \tau, \text{super} : \tau' \}$

∪ { visible types of fields of $\tau' \}$

AssAll = *Ass* ∪ *fypeass*

Forall $1 \leq i \leq n$

($\text{lexp}_{i_t} : \text{rty}F_i, \text{ConSF}_i$) = **TYPEExpr**(*AssAll*, lexpr_i)

TYPE: $\text{TypeAssumptions} \times \text{Class} \rightarrow \text{TClass} \times \text{ConstraintsSet}$

TYPE(*Ass*, $\text{Class}(\tau, \text{extends}(\tau'), \text{fdecls})$) =

let

fdecls = [Field(f_1, lexpr_1), ..., Field(f_n, lexpr_n)]

fypeass = { $\text{this}.f_i : a_i \mid a_i$ fresh type variables }

∪ { $\text{this} : \tau, \text{super} : \tau' \}$

∪ { visible types of fields of $\tau' \}$

AssAll = *Ass* ∪ *fypeass*

Forall $1 \leq i \leq n$

($\text{lexpr}_{i_t} : \text{rty}F_i, \text{ConSF}_i$) = **TYPEExpr**(*AssAll*, lexpr_i)

*fdecls*_{*t*} =

[Field($a_1, f_1, \text{lexpr}_{1_t} : \text{rty}F_1$), ..., Field($a_n, f_n, \text{lexpr}_{n_t} : \text{rty}F_{n_t}$)]

TYPE: $\text{TypeAssumptions} \times \text{Class} \rightarrow \text{TClass} \times \text{ConstraintsSet}$

TYPE(*Ass*, $\text{Class}(\tau, \text{extends}(\tau'), \text{fdecls})$) =

let

fdecls = $[\text{Field}(f_1, \text{lexpr}_1), \dots, \text{Field}(f_n, \text{lexpr}_n)]$

ftypeass = $\{ \text{this}.f_i : a_i \mid a_i \text{ fresh type variables} \}$

$\cup \{ \text{this} : \tau, \text{super} : \tau' \}$

$\cup \{ \text{visible types of fields of } \tau' \}$

AssAll = *Ass* \cup *ftypeass*

Forall $1 \leq i \leq n$

$(\text{lexp}_{i_t} : \text{rty}F_i, \text{ConSF}_i) = \text{TYPEExpr}(\text{AssAll}, \text{lexpr}_i)$

*fdecls*_t =

$[\text{Field}(a_1, f_1, \text{lexpr}_{1_t} : \text{rty}F_1), \dots, \text{Field}(a_n, f_n, \text{lexpr}_{n_t} : \text{rty}F_{n_t})]$

in

$\text{Class}(\tau, \text{extends}(\tau'), \text{fdecls}_t),$

$(\bigcup_i \text{ConSF}_i \cup \{ (r\text{Ty}F_i \triangleleft a_i) \mid 1 \leq i \leq n \})$)

TYPEExpr: TypeAssumptions × Expr → TExpr × ConstraintsSet

TYPEExpr: $\text{TypeAssumptions} \times \text{Expr} \rightarrow \text{TExpr} \times \text{ConstraintsSet}$

For Lambda:

TYPEExpr(*Ass*, **Lambda**((x_1, \dots, x_N) , *expr* | *stmt*)) =

let

AssArgs = { $x_i : a_i$ | *a_i* fresh type variables }

(*expr_t* : *rty*, *ConS*) = **TYPEExpr**(*Ass* ∪ *AssArgs*, *expr*)

| (*stmt_t* : *rty*, *ConS*) = **TYPEStmt**(*Ass* ∪ *AssArgs*, *stmt*)

in

(**Lambda**($(x_1 : a_1, \dots, x_N : a_N)$, *expr_t* : *rty* | *stmt_t* : *rty*) : *a*,

ConS ∪ { (**Fun** *N* < *rty*, *a*₁, ..., *a*_{*N*} > ≤ *a*) }),

where *a* is a fresh type variable

TYPEstmt: TypeAssumptions × Stmt → TStmt × ConstraintsSet

TYPEstmt: $\text{TypeAssumptions} \times \text{Stmt} \rightarrow \text{TStmt} \times \text{ConstraintsSet}$

For Return:

TYPEstmt(*Ass*, **Return**(*e*)) =
 let
 (*e*_{*t*} : *rty*, *ConS*) = **TYPEExpr**(*Ass*, *e*)
 in
 (**Return**(*e*_{*t*} : *rty*) : *a*, *ConS* ∪ { (*rty* < *a*) })
 where a is a fresh type variable

For Block:

```
TYPEstmt( Ass, Block( s ) ) =  
  let  
    (st : rty, ConS) = TYPEstmt( Ass, s )  
  in  
    (Block( st : rty ) : rty, ConS)
```

For Block:

$$\begin{aligned} \text{TYPEStmt}(\text{Ass}, \text{Block}(s)) = \\ \text{let} \\ (s_t : rty, \text{ConS}) = \text{TYPEStmt}(\text{Ass}, s) \\ \text{in} \\ (\text{Block}(s_t : rty) : rty, \text{ConS}) \end{aligned}$$

$$\begin{aligned} \text{TYPEStmt}(\text{Ass}, \text{Block}(s_1, \dots, s_n)) = \\ \text{let} \\ (s_{1_t} : rty_1, \text{ConS}_1) = \text{TYPEStmt}(\text{Ass}, s_1) \\ (\text{Block}(s_{2_t}, \dots, s_{n_t}) : rty_2, \text{ConS}_2) = \\ \text{TYPEStmt}(\text{Ass}, \text{Block}(s_2, \dots, s_n)) \\ \text{in} \\ (\text{Block}(s_{1_t} : rty_1, s_{2_t}, \dots, s_{n_t}) : a, \\ \text{ConS}_1 \cup \text{ConS}_2 \cup \{ (rty_1 \triangleleft a), (rty_2 \triangleleft a) \}) \\ \text{where } a \text{ is a fresh type variable} \end{aligned}$$

Example

```
class Matrix extends Vector<Vector<Integer>> {  
    op = (m) -> (f) -> f.apply(this, m); }  
}
```

Example

```
class Matrix extends Vector<Vector<Integer>> {
    op = (m) -> (f) -> f.apply(this, m); }
```

The result contains:

$l_{expr1_t} =$

```
Lam(m : am,
    Lam(f : af,
        MCall(LoVar(f) : af,
            apply(this : Matrix,
                LoVar(m) : am)) : aapp) : aλf) : aλm
```

Example

```
class Matrix extends Vector<Vector<Integer>> {
    op = (m) -> (f) -> f.apply(this, m); }
```

The result contains:

$lexpr_{1_t} =$

```
Lam(m : am,
    Lam(f : af,
        MCall(LoVar(f) : af,
            apply(this : Matrix,
                LoVar(m) : am)) : aapp) : aλf) : aλm
```

and the set of constraint:

$$\{ (a_{\lambda m} \triangleleft a_{op}), (Fun1\langle a_{\lambda f}, a_m \rangle \triangleleft a_{\lambda m}), \\ (Fun1\langle a_{app}, a_f \rangle \triangleleft a_{\lambda f}), (a_f \triangleleft Fun2\langle a_3, a_1, a_2 \rangle), \\ (Matrix \triangleleft a_1), (a_m \triangleleft a_2), (a_3 \triangleleft a_{app}) \}$$

SOLVE: $\text{ConstraintsSet} \rightarrow \text{Constraints_SubstSet} \cup \{ \textit{fail} \}$

SOLVE(*cs*) =

let *subs* = **TUnify**(*cs*)

in

if (there are $\sigma \in \textit{subs}$ in solved form) **then**

$\{ \sigma \in \textit{subs} \mid \sigma \text{ is in solved form} \}$

if (there are $\sigma \in \textit{subs}$, which has the form

$\{ v R v' \mid v, v' \text{ are type vars} \} \cup \{ v \doteq \theta \mid v \text{ is a type vars} \}$)

then

$\{ v R v' \mid v, v' \text{ are type vars} \} \cup \{ v \doteq \theta \mid v \text{ is a type vars} \}$

else *fail*

The sub-function **TUnify**¹

TUnify: ConstraintsSet \rightarrow { ConstraintsSet } \cup { fail }

Input: $\{ (\theta_1 \triangleleft \theta'_1), \dots, (\theta_n \triangleleft \theta'_n) \}$

Output: $\{ \sigma_1, \dots, \sigma_m \}$

Post-condition: $\forall_j \{ (\sigma_j(\theta_1) \leq^* \sigma_j(\theta'_1)), \dots, (\sigma_j(\theta_n) \leq^* \sigma_j(\theta'_n)) \}$
where \leq^* is the sub-typing relation

¹[Pluemicke 2009]: *Java type unification with wildcards*, INAP 07

TI: $\text{TypeAssumptions} \times \text{Class} \rightarrow \{ (\text{Constraints}, \text{TClass}) \} \cup \{ \text{fail} \}$

TI: $\text{TypeAssumptions} \times \text{Class} \rightarrow \{ (\text{Constraints}, \text{TClass}) \} \cup \{ \text{fail} \}$

TI($\text{Ass}, \text{Class}(\tau, \text{extends}(\tau'), fdecls, mdecls)$) =

let

($\text{Class}(\tau, \text{extends}(\tau'), fdecls_t, mdecls_t)$, ConS) =
TYPE($\text{Ass}, \text{Class}(\tau, \text{extends}(\tau'), fdecls, mdecls)$)
 $\{ (cs_1, \sigma_1), \dots, (cs_n, \sigma_n) \} = \text{SOLVE}(\text{ConS})$

in

$\{ (cs_i, \sigma_i(\text{Class}(\tau, \text{extends}(\tau'), fdecls_t, mdecls_t))) \mid 1 \leq i \leq n \}$

Example

$$cs = \{ (a_{\lambda m} \triangleleft a_{op}), (\text{Fun1} \langle a_{\lambda f}, a_m \rangle \triangleleft a_{\lambda m}), \\ (\text{Fun1} \langle a_{app}, a_f \rangle \triangleleft a_{\lambda f}), (a_f \triangleleft \text{Fun2} \langle a_3, a_1, a_2 \rangle), \\ (\text{Matrix} \triangleleft a_1), (a_m \triangleleft a_2), (a_3 \triangleleft a_{app}) \}$$

Example

$$cs = \{ (a_{\lambda m} \triangleleft a_{op}), (\text{Fun1}\langle a_{\lambda f}, a_m \rangle \triangleleft a_{\lambda m}), \\ (\text{Fun1}\langle a_{app}, a_f \rangle \triangleleft a_{\lambda f}), (a_f \triangleleft \text{Fun2}\langle a_3, a_1, a_2 \rangle), \\ (\text{Matrix} \triangleleft a_1), (a_m \triangleleft a_2), (a_3 \triangleleft a_{app}) \}$$

The result of SOLVE:

$$\{ (\{ a_m \triangleleft a_2, a_3 \triangleleft a_{app} \} \cup \\ \{ a_{op} \doteq \text{Fun1}\langle \text{Fun1}\langle a_{app}, \text{Fun2}\langle a_3, \text{Matrix}, a_2 \rangle \rangle, a_m \rangle, \\ a_{\lambda m} \doteq \text{Fun1}\langle \text{Fun1}\langle a_{app}, \text{Fun2}\langle a_3, \text{Matrix}, a_2 \rangle \rangle, a_m \rangle, \\ a_{\lambda f} \doteq \text{Fun1}\langle a_{app}, \text{Fun2}\langle a_3, \text{Matrix}, a_2 \rangle \rangle, \\ a_f \doteq \text{Fun2}\langle a_3, \text{Matrix}, a_2 \rangle, a_1 \doteq \text{Matrix} \rangle) \\ (\{ a_m \triangleleft a_2, a_3 \triangleleft a_{app} \} \cup \\ \{ a_{op} \doteq \text{Fun1}\langle \text{Fun1}\langle a_{app}, \text{Fun2}\langle a_3, \text{Vec}\langle \text{Vec}\langle \text{Int} \rangle \rangle, a_2 \rangle \rangle, a_m \rangle, \\ a_{\lambda m} \doteq \text{Fun1}\langle \text{Fun1}\langle a_{app}, \text{Fun2}\langle a_3, \text{Vec}\langle \text{Vec}\langle \text{Int} \rangle \rangle, a_2 \rangle \rangle, a_m \rangle, \\ a_{\lambda f} \doteq \text{Fun1}\langle a_{app}, \text{Fun2}\langle a_3, \text{Vec}\langle \text{Vec}\langle \text{Int} \rangle \rangle, a_2 \rangle \rangle, \\ a_f \doteq \text{Fun2}\langle a_3, \text{Vec}\langle \text{Vec}\langle \text{Int} \rangle \rangle, a_2 \rangle, a_1 \doteq \text{Vec}\langle \text{Vec}\langle \text{Int} \rangle \rangle \rangle) \} \}$$

Example cont.

```
class Matrix<a_2, a_m extends a_2, a_app, a_3 extends a_app>  
extends Vector<Vector<Integer>> {  
    Fun1<Fun1<a_app, Fun2<a_3, Matrix,a_2>>, a_m>  
    op = (a_m m) -> (Fun2<a_3, Matrix,a_2> f) ->  
        f.apply(this, m); }
```

Example cont.

```
class Matrix<a_2, a_m extends a_2, a_app, a_3 extends a_app>
  extends Vector<Vector<Integer>> {
  Fun1<Fun1<a_app, Fun2<a_3, Matrix,a_2>>, a_m>
  op = (a_m m) -> (Fun2<a_3, Matrix,a_2> f) ->
    f.apply(this, m); }
```

more principal than ...

Example cont.

```
class Matrix
extends Vector<Vector<Integer>> {
    Fun1<Fun1<Matrix, Fun2<Matrix, Matrix,Matrix>>, Matrix>
    op = (Matrix m) -> (Fun2<Matrix, Matrix,Matrix> f) ->
        f.apply(this, m); }
```

The algorithm WTYPE for Java_λ²

WTYPE: TypeAssumptions × class → {WellTyping} ∪ {fail}

WTYPE(Ass, Class(cl, extends(τ'), fdecls, ivardecls)) =

let

({ f₁ : a₁, ..., f_n : a_n }, CoeS) =

TYPE(Ass, Class(cl, extends(τ'), fdecls, ivardecls))

(σ, AC) = **match**(CoeS)

((τ₁, ..., τ_m), AC') = **TUnify**(AC)

in

{ (AC', Ass ⊢ { f_i : τ_j ∘ σ(a_i) | 1 ≤ i ≤ n }) | 1 ≤ j ≤ m }

²[Pluemicke 2012]: *Functional implementation of well-typings in Java_λ*,
 IFL 2012

The algorithm WTYPE for Java_λ²

WTYPE: $\text{TypeAssumptions} \times \text{class} \rightarrow \{ \text{WellTyping} \} \cup \{ \text{fail} \}$

WTYPE(Ass , $\text{Class}(cl, \text{extends}(\tau'), fdecls, ivardecls)$) =

let

($\{ f_1 : a_1, \dots, f_n : a_n \}, \text{CoeS}$) =

TYPE(Ass , $\text{Class}(cl, \text{extends}(\tau'), fdecls, ivardecls)$)

(σ, AC) = **match**(CoeS)

($(\tau_1, \dots, \tau_m), AC'$) = **TUnify**(AC)

in

$\{ (AC', \text{Ass} \vdash \{ f_i : \tau_j \circ \sigma(a_i) \mid 1 \leq i \leq n \}) \mid 1 \leq j \leq m \}$

match reduces the function types

²[Pluemicke 2012]: *Functional implementation of well-typings in Java_λ*,
 IFL 2012

Summary and Outlook

Summary

- ▶ Adoption of the Java_λ type inference algorithm to Java 8
- ▶ Simplification as there are no function types

Summary and Outlook

Summary

- ▶ Adoption of the Java_λ type inference algorithm to Java 8
- ▶ Simplification as there are no function types

Outlook

- ▶ Consideration of methods with overloading and overriding in Java_λ and Java 8
- ▶ Introduction of intersection types in the code generation