# Spine-local Type Inference in Cedille

Christopher Jenkins and Aaron Stump

Computer Science University of Iowa

ICFP '18 (Cedille Tutorial)

### Outline

- Background and Motivation
  - Local Type Inference
  - Spine-local Type Inference
- 2 Detailed Example
- 3 Annotation Requirements

### Outline

- Background and Motivation
  - Local Type Inference
  - Spine-local Type Inference
- Detailed Example
- 3 Annotation Requirements

# What is "Local Type Inference"?

- Uses two main techniques
  - Bidirectional typing rules:

► Local type-argument inference:

# What is "Local Type Inference"?

- Uses two main techniques
  - Bidirectional typing rules:

```
Synthesis mode: \lambda x: Nat. x \uparrow Nat \rightarrow Nat
Checking mode: \lambda x. x \downarrow Nat \rightarrow Nat
```

Local type-argument inference:

# What is "Local Type Inference"?

- Uses two main techniques
  - Bidirectional typing rules:

```
Synthesis mode: \lambda x: Nat. x \uparrow Nat \rightarrow Nat
Checking mode: \lambda x. x \downarrow Nat \rightarrow Nat
```

► Local type-argument inference:

Let 
$$id : \forall X. X \rightarrow X$$
  
Type  $id \ 0$   $\uparrow Nat$   
Infer  $X = Nat$  from  $0$ 

Local and Synthetic

# Why use local type inference?

- It is a method of partial type inference
  - Complete type inference: no annotations ever (e.g. Damas-Hindley-Milner and ML)
  - Undecidable for System F (let alone Cedille!)

# Why use local type inference?

- It is a method of partial type inference
  - Complete type inference: no annotations ever (e.g. Damas-Hindley-Milner and ML)
  - Undecidable for System F (let alone Cedille!)
- It is user-friendly
  - Infers many type annotations
  - Predictable annotation requirements
  - Better-quality error messages

# Why use local type inference?

- It is a method of partial type inference
  - Complete type inference: no annotations ever (e.g. Damas-Hindley-Milner and ML)
  - Undecidable for System F (let alone Cedille!)
- It is user-friendly
  - Infers many type annotations
  - Predictable annotation requirements
  - Better-quality error messages
- It is implementer-friendly
  - Relatively simple implementation
  - Extensible: new features added without threatening decidability

Cedille provides a way to interrogate the two core features of LTI

- Bidirectional Typechecking gets special highlighting
- Type-argument Inference gets a dedicated buffer

Cedille provides a way to interrogate the two core features of LTI

- Bidirectional Typechecking gets special highlighting
  - ▶ In navigation mode, type C-h 3 to see bidirectional highlighting.
- Type-argument Inference gets a dedicated buffer

Cedille provides a way to interrogate the two core features of LTI

- Bidirectional Typechecking gets special highlighting
  - ▶ In navigation mode, type C-h 3 to see bidirectional highlighting.
- Type-argument Inference gets a dedicated buffer
  - ▶ In navigation mode, type m to see the meta-variables buffer.
  - all meta-variables present, where introduced, what solutions

Cedille provides a way to interrogate the two core features of LTI

- Bidirectional Typechecking gets special highlighting
  - ▶ In navigation mode, type C-h 3 to see bidirectional highlighting.
- Type-argument Inference gets a dedicated buffer
  - ▶ In navigation mode, type m to see the meta-variables buffer.
  - all meta-variables present, where introduced, what solutions

Cedille also has a <u>novel</u> inference system: *spine-local type inference* 

Why a novel system? Other local type inference systems can sometimes still require "silly" type annotations...

Assume mkpair :  $\forall X \ Y. \ X \rightarrow Y \rightarrow Pair \cdot X \cdot Y$ Type  $mkpair \ (\lambda x. x) \ 0$ 

Why a novel system? Other local type inference systems can sometimes still require "silly" type annotations...

```
Assume mkpair : \forall X \ Y. \ X \rightarrow Y \rightarrow Pair \cdot X \cdot Y Type mkpair \ (\lambda x. x) \ 0 \ \uparrow \ ???
```

We do not expect to locally synthesize a type

Why a novel system? Other local type inference systems can sometimes still require "silly" type annotations...

```
Assume mkpair : \forall X \ Y. \ X \rightarrow Y \rightarrow Pair \cdot X \cdot Y
Type mkpair \ (\lambda x. x) \ 0 \ \uparrow \ ???
```

We do not expect to locally synthesize a type

Why a novel system? Other local type inference systems can sometimes still require "silly" type annotations...

```
Assume mkpair : \forall X \ Y. \ X \rightarrow Y \rightarrow Pair \cdot X \cdot Y
Type mkpair \ (\lambda x. x) \ 0 \quad \uparrow \quad ???
Type mkpair \ (\lambda x. x) \ 0 \quad \Downarrow \quad Pair \cdot (Nat \rightarrow Nat) \cdot Nat
```

- We do not expect to locally synthesize a type
- ... but we would expect to check it against a type

Why a novel system? Other local type inference systems can sometimes still require "silly" type annotations...

```
Assume mkpair : \forall X \ Y. \ X \rightarrow Y \rightarrow Pair \cdot X \cdot Y
Type mkpair \ (\lambda x. x) \ 0 \quad \uparrow \quad ???
Type mkpair \ (\lambda x. x) \ 0 \quad \Downarrow \quad Pair \cdot (Nat \rightarrow Nat) \cdot Nat
```

- We do not expect to locally synthesize a type
- ... but we would expect to check it against a type
  - ▶ We could call this "contextual" type-argument inference.

Why a novel system? Other local type inference systems can sometimes still require "silly" type annotations...

```
Assume mkpair : \forall X \ Y. \ X \rightarrow Y \rightarrow Pair \cdot X \cdot Y
Type mkpair \ (\lambda \ x. \ x) \ 0 \quad \Uparrow \quad ???
Type mkpair \ (\lambda \ x. \ x) \ 0 \quad \Downarrow \quad Pair \cdot (Nat \rightarrow Nat) \cdot Nat
```

- We do not expect to locally synthesize a type
- ... but we would expect to check it against a type
  - ▶ We could call this "contextual" type-argument inference.
- Unfortunately, not done in the two major published LTI systems
  - Popular "unofficial" extension (used in e.g. Scala, Rust)

# Limitations of other LTI Systems (cont.)

• Usually uses "fully-uncurried" function applications

$$f(t_1, ..., t_n)$$

Maximize available info at a single application

# Limitations of other LTI Systems (cont.)

Usually uses "fully-uncurried" function applications

$$f(t_1, ..., t_n)$$

- Maximize available info at a single application
- Usually without partial type application ("all-or-nothing")

$$f[T_1, ..., T_m](t_1, ..., t_n)$$

- Precise, specificational account of this technique
- Better support function currying and partial type applications by being "spine-local."

$$f t_1^{\uparrow\uparrow} t_2^{\uparrow\uparrow} t_3^{\downarrow\downarrow}$$

- Precise, specificational account of this technique
- Better support function currying and partial type applications by being "spine-local."

- Precise, specificational account of this technique
- Better support function currying and partial type applications by being "spine-local."

- Precise, specificational account of this technique
- Better support function currying and partial type applications by being "spine-local."

$$f[S,T,V](t_1,t_2,t_3)$$

- Precise, specificational account of this technique
- Better support function currying and partial type applications by being "spine-local."

$$f \cdot S \cdot T \cdot V \ t_1 \ t_2 \ t_3$$

- Precise, specificational account of this technique
- Better support function currying and partial type applications by being "spine-local."

$$f \cdot S$$
  $t_1 t_2$ 

Type inference in Cedille will continue to be improved upon. Some things we want to address:

• Higher-order type arguments must be provided explicitly:  $(\forall F: \star \rightarrow \star)$ 

Type arguments inferred only between applications:
 nil ↓ List · Nat

Type inference in Cedille will continue to be improved upon. Some things we want to address:

- Higher-order type arguments must be provided explicitly:  $(\forall F: \star \rightarrow \star)$ 
  - ► Comming soon: second-order matching is decidable and finitary
- Type arguments inferred *only* between applications:
   nil ↓ List · Nat

Type inference in Cedille will continue to be improved upon. Some things we want to address:

- Higher-order type arguments must be provided explicitly:  $(\forall F: \star \rightarrow \star)$ 
  - ▶ Comming soon: second-order matching is decidable and finitary
- Type arguments inferred only between applications:
   nil ↓ List · Nat Must write nil · Nat

Type inference in Cedille will continue to be improved upon. Some things we want to address:

- Higher-order type arguments must be provided explicitly:  $(\forall F: \star \to \star)$ 
  - ► Comming soon: second-order matching is decidable and finitary
- Type arguments inferred only between applications:
   nil ↓ List · Nat
   Must write nil · Nat
  - **Coming soon:** polymorphic subsumption  $\forall$  *A. List* · *A* ≤ *List* · *Nat*

### Outline

- Background and Motivation
  - Local Type Inference
  - Spine-local Type Inference
- 2 Detailed Example
- 3 Annotation Requirements

• Application head: variable or abstraction

$$x$$
,  $\Lambda X$ .  $t$ ,  $\lambda x$ .  $t$ 

• Application head: variable or abstraction

$$x$$
,  $\Lambda X$ .  $t$ ,  $\lambda x$ .  $t$ 

• Application spine: head followed by seq. of term, type arguments

$$x$$
  $t_1$   $t_2$   $t_3$  vs  $(((x t_1) t_2) t_3)$ 

• Application head: variable or abstraction

$$x$$
,  $\Lambda X.t$ ,  $\lambda x.t$ 

• Application spine: head followed by seq. of term, type arguments

$$x$$
  $t_1$   $t_2$   $t_3$  vs  $(((x t_1) t_2) t_3)$ 

• Applicand: Term in the function position of an application

$$t_1$$
 in  $t_1$   $t_2$ 

Application head: variable or abstraction

$$x$$
,  $\Lambda X.t$ ,  $\lambda x.t$ 

• Application spine: head followed by seq. of term, type arguments

$$x | t_1 t_2 t_3$$
 vs  $(((x t_1) t_2) t_3)$ 

• Applicand: Term in the function position of an application

$$t_1$$
 in  $t_1$   $t_2$ 

Maximal application: spine that is not an applicand

Not max 
$$\frac{x \ t_1 \ t_2}{x \ t_1 \ t_2} \ t_3$$
  
Max  $\frac{x \ t_1 \ t_2}{x \ t_1} \ t_2$ 

• Check the spine const  $(\lambda x. x)$  zero against  $Nat \rightarrow Nat$ 

- Check the spine const  $(\lambda x. x)$  zero against  $Nat \rightarrow Nat$
- We fill in missing types to "elaborate" to const  $\cdot$  (Nat  $\rightarrow$  Nat)  $\cdot$  Nat ( $\lambda \times :$  Nat. $\times$ ) zero

- Check the spine const  $(\lambda x. x)$  zero against  $Nat \rightarrow Nat$
- We fill in missing types to "elaborate" to const  $\cdot$  (Nat  $\rightarrow$  Nat)  $\cdot$  Nat ( $\lambda \times : Nat. \times$ ) zero
- X inferred contextually, Y synthetically

- Check the spine const  $(\lambda x. x)$  zero against  $Nat \rightarrow Nat$
- We fill in missing types to "elaborate" to const  $\cdot$  (Nat  $\rightarrow$  Nat)  $\cdot$  Nat ( $\lambda$  x: Nat. x) zero
- X inferred contextually, Y synthetically In Cedille:  $const(\lambda x. x)$  zero
  - $ightharpoonup ?X: \star \triangleleft Nat \rightarrow Nat$
  - $\triangleright$  ?Y:  $\star$  = Nat

- Check the spine const  $(\lambda x. x)$  zero against  $Nat \rightarrow Nat$
- We fill in missing types to "elaborate" to const  $\cdot$  (Nat  $\rightarrow$  Nat)  $\cdot$  Nat ( $\lambda x$ : Nat.x) zero
- X inferred contextually, Y synthetically In Cedille:  $const(\lambda x. x)$  zero
  - $ightharpoonup ?X : \star \triangleleft Nat \rightarrow Nat$
- And term argument  $\lambda x. x$  checked against  $Nat \rightarrow Nat$

$$const\ (\lambda\,x.\,x)\ zero\ \Downarrow\ {\sf Nat} o {\sf Nat}$$

• Big idea: locality for type-argument inference is the spine

$$\boxed{\textit{const } (\lambda \, \textit{x}.\, \textit{x}) \; \textit{zero}} \Downarrow \textit{Nat} \rightarrow \textit{Nat}$$

- Big idea: locality for type-argument inference is the spine
  - cage meta-variables here never "escape" the spine or "descend" into the arguments

$$f t_1...t_n$$

$$const (\lambda x. x) zero \Downarrow Nat \rightarrow Nat$$

- Big idea: locality for type-argument inference is the spine
  - ► cage meta-variables here
    never "escape" the spine or "descend" into the arguments
  - some meta-vars inferred contextually
     Using the expected result type



$$const\ (\lambda\,x.\,x)\ zero\ \downarrow \ \mathsf{Nat} o \ \mathsf{Nat}$$

- Big idea: locality for type-argument inference is the spine
  - cage meta-variables here never "escape" the spine or "descend" into the arguments
  - some meta-vars inferred contextually
     Using the expected result type
- **Consequence:** you know where to look when type-argument inference fails!



```
Assume const : \forall X \ Y. X \rightarrow Y \rightarrow X
```

Assume zero : Nat

Type  $const(\lambda x. x) zero \Downarrow Nat \rightarrow Nat$ 

Synthesize type for head

```
Assume const : \forall X \ Y. \ X \rightarrow Y \rightarrow X
Assume zero : Nat

Type const \ (\lambda x. x) \ zero \Downarrow Nat \rightarrow Nat

Match X \lhd_X \ Nat \rightarrow Nat
```

Match head return with expected return type

```
Assume const : \forall X \ Y. \ X \rightarrow Y \rightarrow X
Assume zero : Nat

Type const \ (\lambda x. x) \ zero \Downarrow Nat \rightarrow Nat = [Nat \rightarrow Nat/X]X
Match X \bowtie_X Nat \rightarrow Nat
```

Get a contextual instantiation

Type first argument

```
Assume const : \forall X \ Y. \ X \to Y \to X
Assume zero : Nat

Type const \ (\lambda x. x) \ zero \ \Downarrow \ Nat \to Nat = [Nat \to Nat/X]X
Match X \ \lhd_X \ Nat \to Nat

Type \lambda x. x \ \Downarrow \ [Nat \to Nat/X]X
```

Type first argument in checking mode

```
\forall X Y.X \rightarrow Y \rightarrow X
Assume
                       const
Assume
                                                 Nat
                        zero
             const (\lambda x. x) zero
 Type
                                            \Downarrow Nat \rightarrow Nat = [Nat \rightarrow Nat/X]X
                                           \triangleleft_X Nat \rightarrow Nat
Match
                         Χ
                                            \Downarrow [Nat \rightarrow Nat/X]X
 Type
                      \lambda x.x
 Type
                        zero
```

Type second argument

```
\forall X Y. X \rightarrow Y \rightarrow X
Assume
                     const
                                         : Nat
Assume
                      zero
 Type const (\lambda x. x) zero
                                         \Downarrow Nat \rightarrow Nat = [Nat \rightarrow Nat/X]X
                                        \triangleleft_X Nat \rightarrow Nat
Match
                       Χ
 Type
                     \lambda x.x
                                         \Downarrow [Nat \rightarrow Nat/X]X
                                            [Nat/Y]Y
 Type
                      zero
```

Type second argument in synthesis mode

```
\forall X Y.X \rightarrow Y \rightarrow X
Assume
                     const
                                         : Nat
Assume
                      zero
 Type const (\lambda x. x) zero
                                         \Downarrow Nat \rightarrow Nat = [Nat \rightarrow Nat/X]X
                                        \triangleleft_X Nat \rightarrow Nat
Match
                       Χ
 Type
                     \lambda x.x
                                         \Downarrow [Nat \rightarrow Nat/X]X
                                            [Nat/Y]Y
 Type
                      zero
```

Conclude the spine has the expected type!

#### Outline

- Background and Motivation
  - Local Type Inference
  - Spine-local Type Inference
- 2 Detailed Example
- 3 Annotation Requirements

Saw how type inference works. Now – where does it need help?

Term and type abstractions λx.t, ΛX.t
 Bare abstractions can only be checked; require type / kind annotations to synthesize

- Term and type abstractions  $\lambda x.t$ ,  $\Lambda X.t$ Bare abstractions can only be checked; require type / kind annotations to synthesize
- But: when is a term's type checked? In a spine?

- Term and type abstractions  $\lambda x. t$ ,  $\Lambda X. t$ Bare abstractions can only be checked; require type / kind annotations to synthesize
- But: when is a term's type checked? In a spine?

$$f$$
  $t_1$   $t_2$   $t_3 \Downarrow T$ 

- Term and type abstractions  $\lambda x.t$ ,  $\Lambda X.t$ Bare abstractions can only be checked; require type / kind annotations to synthesize
- But: when is a term's type checked? In a spine?
  - when the checked type of the spine

$$f$$
  $t_1$   $t_2$   $t_3 \downarrow T$ 

- Term and type abstractions  $\lambda x.t$ ,  $\Lambda X.t$ Bare abstractions can only be checked; require type / kind annotations to synthesize
- But: when is a term's type checked? In a spine?
  - when the checked type of the spine
  - and the synthesized type of the head

$$f^{\uparrow}$$
  $t_1$   $t_2$   $t_3 \downarrow T$ 

- Term and type abstractions  $\lambda x.t$ ,  $\Lambda X.t$ Bare abstractions can only be checked; require type / kind annotations to synthesize
- But: when is a term's type checked? In a spine?
  - when the checked type of the spine
  - ▶ and the synthesized type of the head
  - and any synthesized types from earlier args

$$f^{\uparrow\uparrow} t_1^{\uparrow\uparrow} t_2 t_3 \Downarrow T$$

- Term and type abstractions  $\lambda x.t$ ,  $\Lambda X.t$ Bare abstractions can only be checked; require type / kind annotations to synthesize
- But: when is a term's type checked? In a spine?
  - when the checked type of the spine
  - and the synthesized type of the head
  - and any synthesized types from earlier args
  - tells us the complete type to check

$$f^{\uparrow\uparrow} t_1^{\uparrow\uparrow} t_2^{\Downarrow} t_3 \Downarrow T$$

Saw how type inference works. Now – where does it need help?

Type arguments (FO only)
 Must provide type args explicitly when not informed by

- Type arguments (FO only)
   Must provide type args explicitly when not informed by
  - ▶ Checked type for spine:  $const(\lambda x. x) zero \Downarrow [Nat \rightarrow Nat/X]X$

- Type arguments (FO only)
   Must provide type args explicitly when not informed by
  - ▶ Checked type for spine:  $const(\lambda x.x)$   $zero \Downarrow [Nat \rightarrow Nat/X]X$
  - ▶ Synthesized type for args:  $zero \Uparrow [Nat/Y]Y$

- Type arguments (FO only)
   Must provide type args explicitly when not informed by
  - ▶ Checked type for spine:  $const (\lambda x. x) zero \Downarrow [Nat \rightarrow Nat/X]X$
  - ▶ Synthesized type for args:  $zero \Uparrow [Nat/Y]Y$
- Example:  $\forall Y X.X \rightarrow X$ 
  - Y doesn't occur in result or arg position
  - ⇒ must be instantiated explicitly.

Saw how type inference works. Now – where does it need help?

 Type of a function in a term or type application must reveal resp. a type quantifier or arrow

$$\begin{array}{ccc} f \cdot T & \Longrightarrow & f : \forall X.S \\ f t & \Longrightarrow & f : \forall \overline{X}.S \to T \end{array}$$

Saw how type inference works. Now – where does it need help?

 Type of a function in a term or type application must reveal resp. a type quantifier or arrow

$$\begin{array}{ccc} f \cdot T & \Longrightarrow & f : \forall X.S \\ f t & \Longrightarrow & f : \forall \overline{X}.S \to T \end{array}$$

• **Example:** assume absurd :  $\forall X : \star . X$ 

Saw how type inference works. Now - where does it need help?

 Type of a function in a term or type application must reveal resp. a type quantifier or arrow

$$\begin{array}{ccc}
f \cdot T & \Longrightarrow & f : \forall X.S \\
f t & \Longrightarrow & f : \forall \overline{X}.S \to T
\end{array}$$

- **Example:** assume absurd :  $\forall X : \star . X$ 
  - ▶ Will **not** type absurd zero
  - ▶ Will type  $absurd \cdot (Nat \rightarrow Nat)$  zeroOr even  $absurd \cdot (\forall X.X \rightarrow X)$  zero

Saw how type inference works. Now – where does it need help?

 Type of a function in a term or type application must reveal resp. a type quantifier or arrow

$$\begin{array}{ccc} f \cdot T & \Longrightarrow & f : \forall X.S \\ f t & \Longrightarrow & f : \forall \overline{X}.S \to T \end{array}$$

- **Example:** assume absurd :  $\forall X : \star . X$ 
  - ▶ Will **not** type absurd zero
  - ▶ Will type  $absurd \cdot (Nat \rightarrow Nat)$  zero Or even  $absurd \cdot (\forall X.X \rightarrow X)$  zero
- Advantage: meta-variables are one-to-one with quantified type variables in functions!
  - It's always easy to understand why a meta-var was introduced!

#### Thanks!

- language-overview/type-inference.ced
   For practical examples
- Paper can be found on arXiv (To appear in proceedings of IFL 2018)