# Typed Lambda Calculus (3/3)

#### Eduardo Bonelli

"There may, indeed, be other applications of the system other than its use as a logic"

Alonzo Church, 1932

February 21, 2019

#### Contents

#### Inference

Motivation Type Variables and Type Substitutions Problem Specification

#### Unification

Motivation
Definition and Algorithm
Unification Algorithm

### Inference Algorithm

Inference Algorithm Examples

### Type Inference

- Transform terms without type annotations or with partial type annotations into a typable term
- ▶ The missing type information must be inferred
- Benefits
  - programmer may omit some type declarations
  - does not affect performance: type inference takes place at compile time

### Type Inference

- Specially useful in languages with polymorphism
- We start by restricting our study to  $\lambda_V^{\mathbb{B}, o}$
- ▶ Even though  $\lambda_V^{\mathbb{B}, \to}$  is **not** polymorphic, it suffices to introduce many important notions in type inference
- ▶ We will address type inference for let-polymorphism later

# The Type Inference Problem

Modified  $\lambda_V^{\mathbb{B}, o}$  syntax – Removal of type annotations

```
M ::= x
\mid true \mid false \mid if M then P else Q
\mid 0 \mid succ(M) \mid pred(M) \mid iszero(M)
\mid \lambda x : \sigma.M \mid M N \mid
\mid fix M
```

# The Type Inference Problem

Modified  $\lambda_V^{\mathbb{B}, o}$  syntax – Removal of type annotations

```
M ::= x
\mid true \mid false \mid if M then P else Q
\mid 0 \mid succ(M) \mid pred(M) \mid iszero(M)
\mid \lambda x.M \mid M N \mid
\mid fix M
```

# **Erasing Function**

The function  $Erase(\cdot)$ , given a term in  $\lambda_V^{\mathbb{B},\to}$ , erases all type annotations

$$Erase(\lambda x : \mathbb{N}.\lambda f : \mathbb{N} \to \mathbb{N}.f x) = \lambda x.\lambda f.f x$$

Given a term U without type annotations, find a standard term (i.e. one with type annotations) M s.t.

- 1.  $\Gamma \triangleright M : \sigma$ , for some  $\Gamma$  and  $\sigma$ ; and
- 2. Erase(M) = U

### Examples

▶ For  $U = \lambda x.x + 5$  we take  $M = \lambda x : \mathbb{N}.x + 5$  (note: no other possibility)

Given a term U without type annotations, find a standard term (i.e. one with type annotations) M s.t.

- 1.  $\Gamma \triangleright M : \sigma$ , for some  $\Gamma$  and  $\sigma$ ; and
- 2. Erase(M) = U

- ▶ For  $U = \lambda x.x + 5$  we take  $M = \lambda x : \mathbb{N}.x + 5$  (note: no other possibility)
- ▶ For  $U = \lambda x.\lambda f.f.x$  we take  $M_{\sigma,\tau} = \lambda x: \sigma.\lambda f: \sigma \to \tau.f.x$  (there is a  $M_{\sigma,\tau}$  for each  $\sigma,\tau$ )

Given a term U without type annotations, find a standard term (i.e. one with type annotations) M s.t.

- 1.  $\Gamma \triangleright M : \sigma$ , for some  $\Gamma$  and  $\sigma$ ; and
- 2. Erase(M) = U

- ▶ For  $U = \lambda x.x + 5$  we take  $M = \lambda x : \mathbb{N}.x + 5$  (note: no other possibility)
- ▶ For  $U = \lambda x.\lambda f.f.x$  we take  $M_{\sigma,\tau} = \lambda x: \sigma.\lambda f: \sigma \to \tau.f.x$  (there is a  $M_{\sigma,\tau}$  for each  $\sigma,\tau$ )
- ► For  $U = \lambda x.\lambda f.f(fx)$  we take  $M_{\sigma} = \lambda x : \sigma.\lambda f : \sigma \to \sigma.f(fx)$  (there is a  $M_{\sigma}$  for each  $\sigma$ )

Given a term U without type annotations, find a standard term (i.e. one with type annotations) M s.t.

- 1.  $\Gamma \triangleright M : \sigma$ , for some  $\Gamma$  and  $\sigma$ ; and
- 2. Erase(M) = U

- ▶ For  $U = \lambda x.x + 5$  we take  $M = \lambda x : \mathbb{N}.x + 5$  (note: no other possibility)
- ▶ For  $U = \lambda x.\lambda f.f.x$  we take  $M_{\sigma,\tau} = \lambda x: \sigma.\lambda f: \sigma \to \tau.f.x$  (there is a  $M_{\sigma,\tau}$  for each  $\sigma,\tau$ )
- ► For  $U = \lambda x.\lambda f.f(fx)$  we take  $M_{\sigma} = \lambda x : \sigma.\lambda f : \sigma \to \sigma.f(fx)$  (there is a  $M_{\sigma}$  for each  $\sigma$ )
- For U = xx there is no M with the desired property

## The Type Checking Problem

#### type checking $\neq$ type inference

#### Type checking

Given a standard term M determine whether there exist  $\Gamma$  and  $\sigma$  s.t.  $\Gamma \rhd M$ :  $\sigma$  is derivable.

- ► Easier than type inference
- Simply follow the syntactic structure of M to reconstruct typing judgement
- ► Essentially equivalent to determining, given  $\Gamma$  and  $\sigma$ , if  $\Gamma \rhd M : \sigma$  is derivable.

#### Inference

Motivation

### Type Variables and Type Substitutions

**Problem Specification** 

#### Unification

Motivation Definition and Algorithm Unification Algorithm

### Inference Algorithm

Inference Algorithm Examples

### Type Variables

- ▶ Given  $\lambda x.\lambda f.f(fx)$ , for each  $\sigma$ ,  $M_{\sigma} = \lambda x : \sigma.\lambda f : \sigma \to \sigma.f(fx)$  is a possible solution
- How may we provide a unique expression that encompasses all of them? Using type variables
  - ▶ All solutions may be represented with

$$\lambda x : s.\lambda f : s \rightarrow s.f(fx)$$

▶ "s" is a type variable that models an arbitrary type expression

# Type Variables

► Type expressions of  $\lambda_V^{\mathbb{B}, \to}$  are extended with type variables  $s, t, u, \dots$ 

$$\sigma ::= \mathbf{s} \mid \mathbb{N} \mid \mathbb{B} \mid \sigma \to \tau$$

- ightharpoonup s 
  ightarrow t
- $ightharpoonup \mathbb{N} o \mathbb{N} o t$
- ightharpoonup  $\mathbb{B} o t$

# Type Substitution

Function from type variables to type expressions. We use S and T for type substitutions.

- ▶ A substitution S may be applied to:
  - 1. a type expression  $\sigma$  ( $S\sigma$ )
  - 2. a term M(SM)
  - 3. typing context  $\Gamma = \{x_1 : \sigma_1, \dots, x_n : \sigma_n\}$  ( $S\Gamma$  defined below)

$$S\Gamma \stackrel{\mathrm{def}}{=} \{x_1 : S\sigma_1, \ldots, x_n : S\sigma_n\}$$

# Type Substitution - Additional Notions

- ▶ Supporting set  $\{t \mid St \neq t\}$ 
  - ▶ Variables that *S* "affects"
- ▶ We use the notation  $\{\sigma_1/t_1, \ldots, \sigma_n/t_n\}$  for substitutions
- ▶ The substitution whose supporting set is  $\emptyset$  is the identity substitution (id)

#### Inference

Motivation
Type Variables and Type Substitutions
Problem Specification

#### Unification

Motivation Definition and Algorithm Unification Algorithm

### Inference Algorithm

Inference Algorithm Examples

### Instance of a Typing Judgement

A typing judgement  $\Gamma' \triangleright M' : \sigma'$  is an instance of  $\Gamma \triangleright M : \sigma$  if there exists a type substitution S s.t.

$$\Gamma' \supseteq S\Gamma$$
,  $M' = SM$  and  $\sigma' = S\sigma$ 

#### **Property**

If  $\Gamma \triangleright M$ :  $\sigma$  is derivable, then any of its instances too

# Inference Function $\mathbb{W}(\cdot)$

```
Define a function \mathbb{W}(\cdot) s.t. given a term U without type
annotations it enjoys:
 Correctness \mathbb{W}(U) = \Gamma \triangleright M : \sigma implies

ightharpoonup Erase(M) = U and

ightharpoonup \Gamma 
ightharpoonup M : \sigma is derivable
Completeness If \Gamma \triangleright M : \sigma is derivable and Erase(M) = U, then
                       ▶ W(U) is successful and
                      ▶ produces a judgement \Gamma' \triangleright M' : \sigma' s.t.
                          \Gamma \triangleright M : \sigma is an instance of it (we say that \mathbb{W}(\cdot)
                          computes the principal type)
```

#### Inference

#### Unification

#### Motivation

Definition and Algorithm Unification Algorithm

### Inference Algorithm

#### Unification

- ➤ The inference algorithm analyzes a term (without type annotations) through its subterms
- Once it obtains type information for them it
  - Consistency Determines if the information obtained for each subterm is consistent
  - Synthesis Synthesizes information about the original term via that of its subterms

### Example

### Consider xy + x(y+1)

- From x y we know  $x :: s \rightarrow t$  and y :: s
- ▶ From x(y+1) we know  $x :: \mathbb{N} \to u$  and  $y :: \mathbb{N}$
- Since a variable can only have one type we must unify the type information
  - ▶ Type  $s \to t$  must be unifiable with  $\mathbb{N} \to u$  since both refer to x
  - ▶ Type s must be unifiable with  $\mathbb{N}$  since both refer to y

### Unification

- ▶ Is type  $s \to t$  compatible or unifiable with  $\mathbb{N} \to u$ ? Yes
  - ▶ It suffices to take substitution  $S \stackrel{\text{def}}{=} \{ \mathbb{N}/s, u/t \}$
  - And observer that  $S(s \to t) = \mathbb{N} \to u = S(\mathbb{N} \to u)$
- ▶ Is type s compatible or unifiable with  $\mathbb{N}$ ? Yes
  - ▶ The aforementioned substitutions is s.t. Ss = SN

Unification: Process of determining whether there exists a substitution S s.t. two given type expressions  $\sigma, \tau$  are unifiable (ie.  $S\sigma = S\tau$ )

▶ We will take a closer look at unification

### Substitution Composition

Composition of S and T, denoted  $S \circ T$ , is the substitution that behaves as follows:

$$(S \circ T)(\sigma) = S(T\sigma)$$

Let 
$$S = \{u \to \mathbb{B}/t, \mathbb{N}/s\}$$
 and  $T = \{v \times \mathbb{N}/u, \mathbb{N}/s\}$ , then  $T \circ S = \{(v \times \mathbb{N}) \to \mathbb{B}/t, v \times \mathbb{N}/u, \mathbb{N}/s\}$ 

- We say S = T if they have the same support set and St = Tt for all t in the support set of S
- $\triangleright$   $S \circ id = id \circ S = S$
- $S \circ (T \circ U) = (S \circ T) \circ U$

#### Preorder on Substitutions

A substitution S is more general than T if there exists U s.t.  $T = U \circ S$ .

► *S* is more general than *T* because *T* is obtained by instantiation of *S* 

### Unifier

A substitution S is a unifier of the set of terms  $\{\sigma_1, \ldots, \sigma_n\}$  if  $S\sigma_1 = \ldots = S\sigma_n$ 

- ▶ Subst.  $\{\mathbb{B}/v, \mathbb{B} \times \mathbb{N}/u\}$  unifies  $\{v \times \mathbb{N} \to \mathbb{N}, u \to \mathbb{N}\}$
- $\{\mathbb{B} \times \mathbb{B}/v, (\mathbb{B} \times \mathbb{B}) \times \mathbb{N}/u\}$  too!
- $\{v \times \mathbb{N}/u\}$  too!
- ▶  ${\mathbb{N} \to s, t \times u}$  are not unifiable
- ▶  $\{s \to \mathbb{N}, s\}$  are not unifiable

# Most General Unifier (MGU)

Substitution S is a MGU of  $\{\sigma_1, \ldots, \sigma_n\}$  if

- 1. *S* is a unifier of  $\{\sigma_1, \ldots, \sigma_n\}$
- 2. S is more general than any other unifier of  $\{\sigma_1, \ldots, \sigma_n\}$

### Examples

- ▶ Subst.  $\{\mathbb{B}/v, \mathbb{B} \times \mathbb{N}/u\}$  unifies  $\{v \times \mathbb{N} \to \mathbb{N}, u \to \mathbb{N}\}$  but is not a MGU since it is an instance of the unifier  $\{v \times \mathbb{N}/u\}$
- ▶  $\{v \times \mathbb{N}/u\}$  is a MGU of the this set

#### **Theorem**

If  $\{\sigma_1, \ldots, \sigma_n\}$  is unifiable, then there exists a MGU and, moreover, it is unique up to renaming of variables.

## Unification Algorithm

### Unification Algorithm for Pairs of Types

- ▶ Input: Ordered pair of types  $\sigma_1 \doteq \sigma_2$
- Output:
  - 1. MGU *S* of  $\sigma_1 \doteq \sigma_2$ , if  $\sigma_1 \doteq \sigma_2$  is unifiable
  - 2. fail, otherwise

### Unification algorithm for sets of types

In order to unify  $\{\sigma_1, \ldots, \sigma_n\}$  with n > 2,

- 1. obtain MGU *S* of  $\sigma_1 \doteq \sigma_2$
- 2. then recursively compute MGU T of  $\{S\sigma_2, \ldots, S\sigma_n\}$
- 3. The MGU of  $\{\sigma_1, \ldots, \sigma_n\}$  is  $T \circ S$

### Martelli-Montanari Algorithm

- Non-deterministic algorithm
- Consists of simplification rules that simplify sets of pairs of types that must be unified (goals)

$$G_0 \mapsto G_1 \mapsto \ldots \mapsto G_n$$

- Sequences that terminate in an empty goal are successful;
   those that terminate in fail fail
- Some simplification steps carry a substitution that represents a partial solution to the problem

$$G_0 \mapsto G_1 \mapsto_{S_1} G_2 \mapsto \ldots \mapsto_{S_k} G_n$$

▶ If the sequence is successful the MGU is  $S_k \circ ... \circ S_1$ 

## Rules of the Martelli-Montanari Algorithm

1. Decomposition

$$\{\sigma_1 \to \sigma_2 \doteq \tau_1 \to \tau_2\} \cup G \mapsto \{\sigma_1 \doteq \tau_1, \sigma_2 \doteq \tau_2\} \cup G$$

$$\{\mathbb{N} \doteq \mathbb{N}\} \cup G \mapsto G$$

$$\{\mathbb{B} \doteq \mathbb{B}\} \cup G \mapsto G$$

2. Trivial Pair Elimination

$$\{s \doteq s\} \cup G \mapsto G$$

- 3. **Swap**: if  $\sigma$  is not a variable  $\{\sigma \doteq s\} \cup G \mapsto \{s \doteq \sigma\} \cup G$
- 4. **Variable Elimination**: if  $s \notin FV(\sigma)$   $\{s \doteq \sigma\} \cup G \mapsto_{\sigma/s} G[\sigma/s]$
- 5. Fail  $\{\sigma \doteq \tau\} \cup G \mapsto \text{fail}, \text{ con } (\sigma, \tau) \in T \cup T^{-1} \text{ y}$   $T = \{(\mathbb{B}, \mathbb{N}), (\mathbb{N}, \sigma_1 \to \sigma_2), (\mathbb{B}, \sigma_1 \to \sigma_1)\}$
- 6. Occur check: si  $s \neq \sigma$  y  $s \in FV(\sigma)$   $\{s \doteq \sigma\} \cup G \mapsto fail$

# Example – Successful Sequence

```
\{(\mathbb{N} \to x) \to (x \to u) \stackrel{.}{=} z \to (y \to y) \to z\}
\mapsto^{1} \qquad \{\mathbb{N} \to x \stackrel{.}{=} z, x \to u \stackrel{.}{=} (y \to y) \to z\}
\mapsto^{3} \qquad \{z \stackrel{.}{=} \mathbb{N} \to x, x \to u \stackrel{.}{=} (y \to y) \to z\}
\mapsto^{4}_{\mathbb{N} \to x/z} \qquad \{x \to u \stackrel{.}{=} (y \to y) \to (\mathbb{N} \to x)\}
\mapsto^{1} \qquad \{x \stackrel{.}{=} y \to y, u \stackrel{.}{=} \mathbb{N} \to x\}
\mapsto^{4}_{y \to y/x} \qquad \{u \stackrel{.}{=} \mathbb{N} \to (y \to y)\}
\mapsto^{4}_{\mathbb{N} \to (y \to y)/u} \qquad \emptyset
```

► The MGU is  $\{\mathbb{N} \to (y \to y)/u\} \circ \{y \to y/x\} \circ \{\mathbb{N} \to x/z\} = \{\mathbb{N} \to (y \to y)/z, y \to y/x, \mathbb{N} \to (y \to y)/u\}$ 

# Example - Failed Sequence

## Properties of the Algorithm

#### **Theorem**

- ▶ The Martelli-Montanari algorithm always terminates
- ▶ Let *G* be the set of of pairs of types
  - if G has a unifier, the algorithm will terminate successfully and return an MGU
  - ▶ if G has no unifier, the algorithm terminates with a fail

#### Inference

#### Unification

Inference Algorithm
Inference Algorithm

## Inference Algorithm

- lacktriangle We'll present the type inference algorithm for  $\lambda_V^{\mathbb{B},\mathbb{N}, o}$
- ▶ Aim: define  $\mathbb{W}(U)$  by recursion on the structure of U
- Frist we address the case for constants and variables, then we address the others
- ▶ We will make use of the unification algorithm

# Inference Algorithm (constants and variables)

```
\mathbb{W}(0) \stackrel{\text{def}}{=} \emptyset \rhd 0 : \mathbb{N}
\mathbb{W}(true) \stackrel{\text{def}}{=} \emptyset \rhd true : \mathbb{B}
\mathbb{W}(false) \stackrel{\text{def}}{=} \emptyset \rhd false : \mathbb{B}
\mathbb{W}(x) \stackrel{\text{def}}{=} \{x : s\} \rhd x : s, \quad s \text{ fresh variable}
```

# Inference Algorithm (succ)

- ▶ Let  $\mathbb{W}(U) = \Gamma \triangleright M : \tau$
- ▶ Let  $S = MGU\{\tau \doteq \mathbb{N}\}$
- ► Then

$$\mathbb{W}(\operatorname{succ}(U)) \stackrel{\operatorname{def}}{=} S\Gamma \rhd S \operatorname{succ}(M) : \mathbb{N}$$

▶ Note: Case *pred* is similar

# Inference Algorithm (iszero)

- ▶ Let  $\mathbb{W}(U) = \Gamma \triangleright M : \tau$
- ▶ Let  $S = MGU\{\tau \doteq \mathbb{N}\}$
- ► Then

$$\mathbb{W}(iszero(U)) \stackrel{\text{def}}{=} S\Gamma \rhd S iszero(M) : \mathbb{B}$$

#### Inference Algorithm (ifThenElse)

- Let
  - $\blacktriangleright$   $\mathbb{W}(U) = \Gamma_1 \triangleright M : \rho$
  - $\blacktriangleright$   $\mathbb{W}(V) = \Gamma_2 \triangleright P : \sigma$
  - $\blacktriangleright \ \mathbb{W}(W) = \Gamma_3 \triangleright Q : \tau$
  - ▶ All type variables in  $\mathbb{W}(U)$ ,  $\mathbb{W}(V)$  and  $\mathbb{W}(W)$  must be different; if they are not we rename them
- Let

$$S = MGU(\{\sigma_1 \doteq \sigma_2 \mid x : \sigma_1 \in \Gamma_i \land x : \sigma_2 \in \Gamma_j, i \neq j\} \cup \{\sigma \doteq \tau, \rho \doteq \mathbb{B}\})$$

Then

$$\mathbb{W}(\textit{if U then V else W})$$

$$\stackrel{\text{def}}{=} S\Gamma_1 \cup S\Gamma_2 \cup S\Gamma_3 \rhd S(\textit{if M then P else Q}) : S\sigma$$

#### Inference Algorithm (application)

- Let
  - $\blacktriangleright$   $\mathbb{W}(U) = \Gamma_1 \triangleright M : \tau$
  - $\blacktriangleright \mathbb{W}(V) = \Gamma_2 \triangleright N : \rho$
  - All type variables in  $\Gamma_2 \triangleright N : \rho$  must be renamed to be disjoint from those in  $\mathbb{W}(U)$
- Let

$$S = MGU\{\sigma_1 \doteq \sigma_2 \mid x : \sigma_1 \in \Gamma_1 \land x : \sigma_2 \in \Gamma_2\}$$

$$\cup$$

$$\{\tau \doteq \rho \rightarrow t\} \text{ with } t \text{ a fresh variable}$$

Then

$$\mathbb{W}(UV) \stackrel{\text{def}}{=} S\Gamma_1 \cup S\Gamma_2 \triangleright S(MN) : St$$

## Inference Algorithm (abstraction)

- ▶ Let  $\mathbb{W}(U) = \Gamma \triangleright M : \rho$
- ▶ If the context has type information on x (i.e.  $x : \tau \in \Gamma$  for some  $\tau$ ), then

$$\mathbb{W}(\lambda x. U) \stackrel{\text{def}}{=} \Gamma \setminus \{x : \tau\} \rhd \lambda x : \tau. M : \tau \to \rho$$

If the context has no information on the type of x (i.e.  $x \notin Dom(\Gamma)$ ) we choose a fresh type variable s and then

$$\mathbb{W}(\lambda x. U) \stackrel{\text{def}}{=} \Gamma \rhd \lambda x : s. M : s \to \rho$$

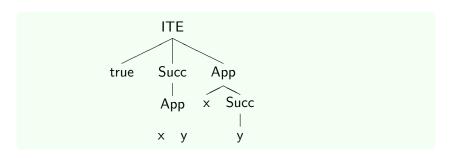
# Inference Algorithm (fix)

- ▶ Let  $\mathbb{W}(U) = \Gamma \triangleright M : \tau$
- ▶ Let  $S = MGU\{\tau \doteq t \rightarrow t\}$ , t fresh variable

$$\mathbb{W}(fix(U)) \stackrel{\text{def}}{=} S\Gamma \rhd S fix(M) : St$$

#### Example

- if true then succ(x y) else x(succ(y))
- ▶ We'll apply the algorithm, step-by-step



#### Example (1/4)

if true then succ(xy) else x(succ(y))

 $\mathbb{W}(true) = \emptyset \triangleright true : \mathbb{B}$ 

## Example (2/4)

#### if true then succ(x y) else x(succ(y))

## Example (3/4)

#### if true then succ(x y) else x(succ(y))

## Example (4/4)

$$M = if true then succ(x y) else x (succ(y))$$

```
▶ \mathbb{W}(true) = \emptyset \triangleright true : \mathbb{B}

▶ \mathbb{W}(succ(xy)) = \{x : t \to \mathbb{N}, y : t\} \triangleright succ(xy) : \mathbb{N}

▶ \mathbb{W}(x succ(y)) = \{x : \mathbb{N} \to w, y : \mathbb{N}\} \triangleright x succ(y) : w

\mathbb{W}(M) = \{x : \mathbb{N} \to \mathbb{N}, y : \mathbb{N}\} \triangleright M : \mathbb{N}

where S = MGU(\{t \to \mathbb{N} \doteq \mathbb{N} \to w, t \doteq \mathbb{N}, \mathbb{N} \doteq w\}) = \{\mathbb{N}/t, \mathbb{N}/w\}
```

#### An Example that Fails

#### M = if true then x 2 else x true

$$\mathbb{W}(x) = \{x : s\} \triangleright x : s$$

$$\mathbb{W}(\underline{2}) = \emptyset \triangleright \underline{2} : \mathbb{N}$$

$$\mathbb{W}(x\underline{2}) = \{x : \mathbb{N} \to t\} \triangleright x\underline{2} : t$$

$$\mathbb{W}(x) = \{x : u\} \triangleright x : u$$

$$\mathbb{W}(true) = \emptyset \triangleright true : \mathbb{B}$$

$$\mathbb{W}(x true) = \{x : \mathbb{B} \to v\} \triangleright x\underline{2} : v$$

$$\mathbb{W}(M) = fail$$
there is no  $MGU(\{\mathbb{N} \to t \doteq \mathbb{B} \to v\})$ 

#### Complexity

- **>** Both unification and type inference for  $\lambda_V^{\mathbb{B}, o}$  are linear
- ► The principal type associated to a term without annotations can be exponential in the size of the term

#### Consider $P^n M$ where $P: s \to s \times s$ and $M: \sigma$

- Does this contradict the statement in the first item?
- No. They may be represented as dags (in which case the size of the principal type is  $\mathcal{O}(n)$ )
- NB: In the presence for polymorphism type inference is exponential