Type Inference and the Hindley-Milner Type System

Armando Solar-Lezama
Computer Science and Artificial Intelligence Laboratory
M.I.T.
With slides from Arvind

September 29, 2011
Type Inference

- Consider the following expression
  - \((\lambda f: \text{int} \to \text{int}. f \ 5) \ (\lambda x: \text{int}. x + 1)\)
  - Is it well typed in \(F_1\)?

\[
\begin{array}{c}
\frac{x: \tau \in \Gamma}{\Gamma \vdash x : \tau} \\
\frac{\Gamma, x: \tau_1 \vdash e: \tau_2}{\Gamma \vdash (\lambda x: \tau_1 \ e): \tau_1 \to \tau_2} \\
\frac{\Gamma \vdash e_1: \tau' \to \tau \quad \Gamma \vdash e_2: \tau'}{\Gamma \vdash e_1e_2: \tau}
\end{array}
\]

\[
\begin{array}{c}
\frac{\Gamma \vdash e_1: \text{int} \quad \Gamma \vdash e_2: \text{int}}{\Gamma \vdash e_1 + e_2: \text{int}}
\end{array}
\]
Type Inference

- There wasn’t a single point in the derivation where we had to look at the type labels in order to know what rule to apply!
  - we could have written the derivation without the labels

- The labels helped us determine the actual types for all the $\tau$s in the typing rules.
  - we could have figured these out even without the labels
  - this is the key idea behind type inference!
Type Inference Strategy 1

- 1. Use the typing rules to define constraints on the possible types of expressions

- 2. Solve the resulting constraint system
Deducing Types

twice \( f \) \( x = f \ (f \ x) \)
What is the most "general type" for twice?

1. Assign types to every subexpression

\[
\begin{align*}
x &: t_0 \\
f &: t_1 \\
f \ x &: t_2 \\
f \ (f \ x) &: t_3
\end{align*}
\]

\( \Rightarrow \) twice \( :: \ t_1 \rightarrow t_0 \rightarrow t_3 \)

2. Set up the constraints

\[
\begin{align*}
t_1 &= t_0 \rightarrow t_2 & \text{because of } (f \ x) \\
t_1 &= t_2 \rightarrow t_3 & \text{because of } f \ (f \ x)
\end{align*}
\]

3. Resolve the constraints

\[
\begin{align*}
t_0 \rightarrow t_2 &= t_2 \rightarrow t_3 \\
\Rightarrow t_0 &= t_2 \text{ and } t_2 = t_3 \Rightarrow t_0 = t_2 = t_3 \\
\Rightarrow \text{twice} :: (t_0 \rightarrow t_0) \rightarrow t_0 \rightarrow t_0
\end{align*}
\]
The language of Equality Constraints

• Consider the following Language of Constraints

\[ C ::= \tau_1 = \tau_2 \mid C \land C \mid \exists \tau. C \]

• Constraints in this language have a lot of good properties
  - Nice and compositional
  - Linear time solution algorithm
Building Constraints from Typing Rules

- Notation

  \[\llbracket \text{Judgment} \rrbracket = \text{Constraints}\]

  - The constraints on the right ensure that the judgment on the left holds
  - This mapping is defined recursively.

- Base cases

  \[\llbracket \Gamma \vdash x : \tau \rrbracket = \Gamma(x) = \tau \quad \llbracket \Gamma \vdash N : \tau \rrbracket = \text{int} = \tau\]

- Inductive Cases

  \[\llbracket \Gamma \vdash e_1 e_2 : \tau \rrbracket = \exists a_1 a_2. ( \llbracket \Gamma \vdash e_1 : a \rightarrow \tau \rrbracket \land \llbracket \Gamma \vdash e_2 : a \rrbracket )\]

  \[\llbracket \Gamma \vdash \lambda x. e : \tau \rrbracket = \exists a_1 a_2. ( \llbracket \Gamma; x : a_1 \vdash e : a_2 \rrbracket \land \tau = a_1 \rightarrow a_2 )\]

  \[\llbracket \Gamma \vdash e_1 + e_2 : \tau \rrbracket = \llbracket \Gamma \vdash e_1 : \text{int} \rrbracket \land \llbracket \Gamma \vdash e_2 : \text{int} \rrbracket \land \tau = \text{int}\]
Back to our example

$$(\lambda f \cdot f\ 5) \ (\lambda x \cdot x + 1)$$

$$[[\Gamma \vdash x : \tau]] = \Gamma(x) = \tau$$
$$[[\Gamma \vdash N : \tau]] = \text{int} = \tau$$

$$[[\Gamma \vdash e_1 e_2 : \tau]] = \exists a ( [ [\Gamma \vdash e_1 : a \rightarrow \tau] \land [ [\Gamma \vdash e_2 : a] ] )$$

$$[[\Gamma \vdash \lambda x. e : \tau]] = \exists a_1 a_2 \cdot ([ [\Gamma ; x : a_1 \vdash e : a_2] ] \land \tau = a_1 \rightarrow a_2 )$$

$$[[\Gamma \vdash e_1 + e_2 : \tau]] = [ [\Gamma \vdash e_1 : \text{int} ] \land [ [\Gamma \vdash e_2 : \text{int} ] \land \tau = \text{int} ]$$
Equality and Unification

• What does it mean for two types $\tau_a$ and $\tau_b$ to be equal?
  – *Structural Equality*
    Suppose $\tau_a = \tau_1 \rightarrow \tau_2$
    $\tau_b = \tau_3 \rightarrow \tau_4$
    Is $\tau_a = \tau_b$?  
    iff $\tau_1 = \tau_3$ and $\tau_2 = \tau_4$

• Can two types be made equal by choosing appropriate substitutions for their type variables?
  – *Robinson’s unification algorithm*
    Suppose $\tau_a = t_1 \rightarrow \text{Bool}$
    $\tau_b = \text{Int} \rightarrow t_2$
    Are $\tau_a$ and $\tau_b$ unifiable?  
    if $t_1 = \text{Int}$ and $t_2 = \text{Bool}$

    Suppose $\tau_a = t_1 \rightarrow \text{Bool}$
    $\tau_b = \text{Int} \rightarrow \text{Int}$
    Are $\tau_a$ and $\tau_b$ unifiable?  
    No
Simple Type Substitutions
needed to define type unification

A substitution is a map

\[ S : \text{Type Variables} \to \text{Types} \]

\[ S = [\tau_1 / t_1, \ldots, \tau_n / t_n] \]

\[ \tau' = S \tau \quad \tau' \text{ is a Substitution Instance of } \tau \]

Example:

\[ S = [(t \to \text{Bool}) / t_1] \]

\[ S \ (t_1 \to t_1) = (t \to \text{Bool}) \to (t \to \text{Bool}) \]

Substitutions can be composed, i.e., \( S_2 \ S_1 \)

Example:

\[ S_1 = [(t \to \text{Bool}) / t_1] ; S_2 = [\text{Int} / t] \]

\[ S_2 \ S_1 \ (t_1 \to t_1) = S_2 ((t \to \text{Bool}) \to (t \to \text{Bool})) \]

\[ = (\text{Int} \to \text{Bool}) \to (\text{Int} \to \text{Bool}) \]
Unification
An essential subroutine for type inference

Unify(τ₁, τ₂) tries to unify τ₁ and τ₂ and returns a substitution if successful

\[
\text{def } \text{Unify}(\tau_1, \tau_2) = \\
\text{case } (\tau_1, \tau_2) \text{ of} \\
(\tau_1, t_2) = [\tau_1 / t_2] \text{ provided } t_2 \not\in \text{FV}(\tau_1) \\
(t_1, \tau_2) = [\tau_2 / t_1] \text{ provided } t_1 \not\in \text{FV}(\tau_2) \\
(\iota_1, \iota_2) = \text{if } (\text{eq? } \iota_1 \iota_2) \text{ then } [ ] \\
\text{else } \text{fail} \\
(\tau_{11} \to \tau_{12}, \tau_{21} \to \tau_{22}) \\
= \text{let } S_1 = \text{Unify}(\tau_{11}, \tau_{21}) \\
S_2 = \text{Unify}(S_1(\tau_{12}), S_1(\tau_{22})) \\
in S_2 S_1 \\
\text{otherwise } = \text{fail}
\]

Does the order matter? No
Type inference strategy 2

- Like strategy 1, but we solve the constraints as we see them
  - Build the substitution map incrementally
**Simple Inference Algorithm**

W(TE, e) returns (S, \(\tau\)) such that S (TE) \(\vdash\) e : \(\tau\)

This is just \(\Gamma\) (it's hard to write \(\Gamma\) in code)

The type environment TE records the most general type of each identifier while the substitution S records the changes in the type variables

\[
\text{Def } W(\text{TE}, \text{e}) = \\
\text{Case } \text{e of} \\
\quad \text{x} = \ldots \\
\quad n = \ldots \\
\quad \lambda x.\text{e} = \ldots \\
\quad (\text{e}_1 \text{e}_2) = \ldots \\
\ldots
\]
Simple Inference Algorithm (cont-1)

\[
\text{Def } W(TE, e) = \\
\text{Case } e \text{ of}
\]

\[
[\Gamma \vdash N : \tau] = \text{int} = \tau \quad N = (\{\}, \text{Typeof(N)})
\]

\[
[\Gamma \vdash x : \tau] = \Gamma(x) = \tau \quad x = \text{if } (x \not\in \text{Dom(TE)}) \text{ then Fail}
\]

\[
\quad \text{else let } \tau = \text{TE}(x);
\]

\[
\quad \text{in } (\{\}, \tau)
\]

\[
[\Gamma \vdash \lambda x. e : \tau] = \exists a_1 a_2. ([\Gamma; x: a_1 \vdash e : a_2] \land \tau = a_1 \to a_2)
\]

\[
\lambda x.e = \text{let } (S_1, \tau_2) = W(\text{TE} + \{ x : u \}, e)
\]

\[
\quad \text{in } (\{\}, \text{Typeof(N)})
\]

\[
\text{u’s represent new type variables}
\]

\[
[\Gamma \vdash e_1 e_2 : \tau] = \exists a ([\Gamma \vdash e_1 : a \to \tau] \land [\Gamma \vdash e_2 : a])
\]

\[
(e_1 e_2) = \text{let } (S_1, \tau_1) = W(\text{TE}, e_1);
\]

\[
(S_2, \tau_2) = W(S_1(\text{TE}), e_2);
\]

\[
S_3 = \text{Unify}(S_2(\tau_1), \tau_2 \to u);
\]

\[
\quad \text{in } (S_3 S_2 S_1, S_3(u))
\]
Simple Inference Algorithm (cont-1)

Def $W(TE, e)$ =

Case $e$ of

$c$ = ($\{\}$, $\text{Typeof}(c)$)

$x$ = if ($x \notin \text{Dom}(TE)$) then Fail

else let $\tau = TE(x)$;

in ($\{\}$, $\tau$)

$\lambda x. e$ = let $(S_1, \tau_1) = W(TE + \{x : u\}, e)$

in $(S_1, S_1(u) \rightarrow \tau_1)$

$(e_1 e_2)$ = let $(S_1, \tau_1) = W(TE, e_1)$;

$(S_2, \tau_2) = W(S_1(TE), e_2)$;

$S_3 = \text{Unify}(S_2(\tau_1), \tau_2 \rightarrow u)$;

in $(S_3 S_2 S_1, S_3(u))$

let $x = e_1$ in $e_2$

= let $(S_1, \tau_1) = W(TE + \{x : u\}, e_1)$;

$S_2 = \text{Unify}(S_1(u), \tau_1)$;

$(S_3, \tau_2) = W(S_2 S_1(TE) + \{x : \tau_1\}, e_2)$;

in $(S_3 S_2 S_1, \tau_2)$

u’s represent new type variables
Def $W(TE, e) =$

Case $e$ of

...$
\lambda x. e = \text{let } (S_1, \tau_1) = W(TE + \{ x : u \}, e)\in (S_1, S_1(u) \rightarrow \tau_1)$

$(e_1 e_2) = \text{let } (S_1, \tau_1) = W(TE, e_1); (S_2, \tau_2) = W(S_1(TE), e_2); S_3 = \text{Unify}(S_2(\tau_1), \tau_2 \rightarrow u); in (S_3 S_2 S_1, S_3(u))$

$W(\{f:u_0\}, f) = (\emptyset, u_0) \quad W(\{f:u_0\}, 5) = (\emptyset, \text{Int})$

$\text{Unify}(u_0, \text{Int} \rightarrow u_1) = \emptyset$

$W(\{f:u_0\}, f \, 5) =$

$W(\emptyset, (\lambda f. f \, 5)) =$

$W(\emptyset, (\lambda f. f \, 5)(\lambda x. x))$
Example

\[
\text{def Unify(}\tau_1, \tau_2) = \\
\text{case } (\tau_1, \tau_2) \text{ of} \\
(\tau_1, t_2) = [\tau_1 / t_2] \text{ provided } t_2 \not\in \text{FV}(\tau_1) \\
(t_1, \tau_2) = [\tau_2 / t_1] \text{ provided } t_1 \not\in \text{FV}(\tau_2) \\
(\iota_1, \iota_2) = \text{if } (\text{eq? } \iota_1 \iota_2) \text{ then } [ ] \\
\text{else fail} \\
(\tau_{11} \rightarrow \tau_{12}, \tau_{21} \rightarrow \tau_{22}) = \text{let } S_1 = \text{Unify}(\tau_{11}, \tau_{21}) \\
S_2 = \text{Unify}(S_1(\tau_{12}), S_1(\tau_{22})) \text{ in } S_2 S_1
\]

\[
W(\{f: u_0\}, f) = (\emptyset, u_0) \\
W(\{f: u_0\}, 5) = (\emptyset, \text{Int}) \\
\text{Unify}(u_0, \text{Int} \rightarrow u_1) = [(\text{Int} \rightarrow u_1)/u_0] \\
W(\{f: u_0\}, f \ 5) = \\
W(\emptyset, (\lambda f. f \ 5)) = \\
W(\emptyset, (\lambda f. f \ 5)(\lambda x. x))
\]
Example

\[
\begin{align*}
\text{Def } W(\text{TE}, e) &= \ldots \\
\text{Case } e \text{ of } \ldots
\end{align*}
\]

\[
\begin{align*}
\lambda x. e &= \text{let } (S_1, \tau_1) = W(\text{TE} + \{ x : u \}, e) \\
&\quad \text{in } (S_1, S_1(u) \rightarrow \tau_1) \\
(e_1 \, e_2) &= \text{let } (S_1, \tau_1) = W(\text{TE}, e_1); \\
&\quad (S_2, \tau_2) = W(S_1(\text{TE}), e_2); \\
&\quad S_3 = \text{Unify}(S_2(\tau_1), \tau_2 \rightarrow u); \\
&\quad \text{in } (S_3 \, S_2 \, S_1, S_3(u))
\end{align*}
\]

\[
\begin{align*}
W(\{ f : u_0 \}, f) &= (\emptyset, u_0) & W(\{ f : u_0 \}, 5) &= (\emptyset, \text{Int}) \\
\text{Unify}(u_0, \text{Int} \rightarrow u_1) &= [(\text{Int} \rightarrow u_1)/u_0] \\
W(\{ f : u_0 \}, f \, 5) &= ([\text{Int} \rightarrow u_1/u_0], u_1) \\
W(\emptyset, (\lambda f. f \, 5)) &= ([([\text{Int} \rightarrow u_1]/u_0], (\text{Int} \rightarrow u_1) \rightarrow u_1) \\
W(\emptyset, (\lambda f. f \, 5)(\lambda x. x))
\end{align*}
\]
Example

Def \( W(TE, e) \)

Case \( e \) of

... 
\( \lambda x.e = \text{let } (S_1, \tau_1) = W(TE + \{ x : u \}, e) \)
\[ \text{in } (S_1, S_1(u) \rightarrow \tau_1) \]
\( (e_1, e_2) = \text{let } (S_1, \tau_1) = W(TE, e_1); \)
\( (S_2, \tau_2) = W(S_1(TE), e_2); \)
\( S_3 = \text{Unify}(S_2(\tau_1), \tau_2 \rightarrow u); \)
\[ \text{in } (S_3 S_2 S_1, S_3(u)) \]

\[ W(\emptyset, (\lambda f. f \, 5)) = ((\text{Int} \rightarrow u_1)/u_0), (\text{Int} \rightarrow u_1) \rightarrow u_1) \]
\[ W(\emptyset, (\lambda x. x)) = (\emptyset, u_3 \rightarrow u_3) \]

\[ \text{Unify}((\text{Int} \rightarrow u_1) \rightarrow u_1, (u_3 \rightarrow u_3) \rightarrow u_4) = \]
\[ W(\emptyset, (\lambda f. f \, 5)(\lambda x. x)) \]
\textbf{Example}

\begin{align*}
\text{def } \text{Unify}(\tau_1, \tau_2) &= \text{case } (\tau_1, \tau_2) \text{ of } \\
(\tau_1, t_2) &= [\tau_1 / t_2] \text{ provided } t_2 \notin \text{FV}(\tau_1) \\
(t_1, \tau_2) &= [\tau_2 / t_1] \text{ provided } t_1 \notin \text{FV}(\tau_2) \\
(\iota_1, \iota_2) &= \text{if } (\text{eq? } \iota_1 \iota_2) \text{ then } [] \\ &\quad \text{ else } \text{ fail} \\
(\tau_{11} \rightarrow \tau_{12}, \tau_{21} \rightarrow \tau_{22}) &= \text{let } S_1 = \text{Unify}(\tau_{11}, \tau_{21}) \quad S_2 = \text{Unify}(S_1(\tau_{12}), \tau_{12}) \\
&\quad \text{in } S_2 S_1
\end{align*}

\begin{align*}
\text{Unify}((\text{Int } \rightarrow u_1), (u_3 \rightarrow u_3)) &= [\text{Int} / u_3 ; \text{Int} / u_1] \\
\text{Unify}((\text{Int } \rightarrow u_1) \rightarrow u_1, (u_3 \rightarrow u_3) \rightarrow u_4) &= [\text{Int} / u_3 ; \text{Int} / u_1 ; \text{Int} / u_4] \\
W(\emptyset, (\lambda f. f 5)(\lambda x. x))
\end{align*}
Example

\[ W(\emptyset, (\lambda f. f \ 5)) = \left( ([\text{Int} \to u_1]/u_0], (\text{Int} \to u_1) \to u_1 \right) \]

\[ W(\emptyset, (\lambda x. x)) = (\emptyset, u_3 \to u_3) \]

\[ \text{Unify}((\text{Int} \to u_1) \to u_1, (u_3 \to u_3) \to u_4) = [\text{Int}/u_3; \text{Int}/u_1; \text{Int}/u_4] \]

\[ W(\emptyset, (\lambda f. f \ 5)(\lambda x. x)) = \left( ([\text{Int} \to u_1]/u_0; \text{Int}/u_3; \text{Int}/u_1; \text{Int}/u_4], \text{Int} \right) \]
What about Let?

- \( \text{let } x = e_1 \text{ in } e_2 \)  
  
  This is Hindley Milner without polymorphism

- Typing rule

  \[ \frac{\Gamma; x: \tau' \vdash e_1: \tau' \quad \Gamma; x: \tau' \vdash e_2: \tau}{\Gamma \vdash \text{let } x = e_1 \text{ in } e_2: \tau} \]

- Constraints

  - \( \llbracket \Gamma \vdash \text{let } x = e_1 \text{ in } e_2 : \tau \rrbracket = \exists \tau', \ (\llbracket \Gamma; x: \tau' \vdash e_1: \tau' \rrbracket \land \llbracket \Gamma; x: \tau' \vdash e_2: \tau \rrbracket) \)

- Algorithm

  Case \( \text{Exp} = \text{let } x = e_1 \text{ in } e_2 \)

  \[ \Rightarrow \text{let } (S_1, \tau_1) = W(TE + \{x : u\}, e_1); \]

  \[ S_2 = \text{Unify}(S_1(u), \tau_1); \]

  \[ (S_3, \tau_2) = W(S_2 S_1(TE) + \{x : \tau_1\}, e_2); \]

  \[ \text{in } (S_3 S_2 S_1, \tau_2) \]
Polymorphism
Some observations

• A type system restricts the class of programs that are considered “legal”
• It is possible a term in the untyped $\lambda$-calculus may be reducible to a value but may not be typeable in a particular type system

```
let
  id = $\lambda$x. x
in
... (id True) ... (id 1) ...
```

*This term is not typeable in the simple type system we have discussed so far. However, it is typeable in the Hindley-Milner system*
Explicit polymorphism

- You’ve seen this before

```java
public interface List<E>
{
    void add(E x);
    E get();
}
```

List<String> ls = ...
ls.add("Hello");
String hello = ls.get(0);

- How do we formalize this?

\[
\Gamma \vdash e : \tau \\
\frac{}{\Gamma \vdash \Lambda t. e : \forall t. \tau}
\]

\[
\Gamma \vdash e : \forall t. \tau' \\
\frac{}{\Gamma \vdash e[\tau] : \tau'[\tau / t]}
\]

- Example

\[
id = \Lambda T. \lambda x : T. x
\]

\[
id[int] 5
\]
Different Styles of Polymorphism

• Impredicative Polymorphism
  \[
  \tau ::= b \mid \tau_1 \rightarrow \tau_2 \mid T \mid \forall T. \tau
  \]
  \[
  e ::= x \mid \lambda x: \tau.e \mid e_1 e_2 \mid \forall T.e \mid e[\tau]
  \]

• Very powerful
  – Although you still can’t express recursion

• Type inference is undecidable!
Different Styles of Polymorphism

- Predicative Polymorphism
  
  \[ \tau ::= b \mid \tau_1 \to \tau_2 \mid T \]
  
  \[ \sigma ::= \tau \mid \forall T. \sigma \mid \sigma_1 \to \sigma_2 \]
  
  \[ e ::= x \mid \lambda x: \sigma. e \mid e_1 e_2 \mid \Lambda T. e \mid e[\tau] \]

- Still very powerful
  - But you can no longer instantiate with a polymorphic type

- Type inference is still undecidable!
Different Styles of Polymorphism

- Prenex Predicative Polymorphism
  \[ \tau ::= b \mid \tau_1 \rightarrow \tau_2 \mid T \]
  \[ \sigma ::= \tau \mid \forall T. \sigma \]
  \[ e ::= x \mid \lambda x: \tau . e \mid e_1 e_2 \mid \Lambda T. e \mid e[\tau] \]

- Now we have decidable type inference
- But polymorphism is now very limited
  - We can’t pass polymorphic functions as arguments!!
  - \((\lambda s: \forall T. \tau \ldots s[int]x \ldots s[bool]x)(\Lambda T. code\ for\ sort)\)
Let polymorphism

- Introduce let $x = e_1$ in $e_2$
  - Just like saying $(\lambda x. e_2) e$
  - Except $x$ can be polymorphic

- Good engineering compromise
  - Enhance expressiveness
  - Preserve decidability

- This is the Hindley Milner type system
Type inference with polymorphism
Polymorphic Types

Constraints:

\[
\begin{align*}
\text{id} &: t_1 \rightarrow t_1 \\
\text{id} &: \text{Int} \rightarrow t_2 \\
\text{id} &: \text{Bool} \rightarrow t_3 \\
\end{align*}
\]

Does not unify!!

Solution: Generalize the type variable

\[
\text{id} &: \forall t_1. t_1 \rightarrow t_1
\]

Different uses of a generalized type variable may be \textit{instantiated} differently

\[
\begin{align*}
\text{id}_2 &: \text{Bool} \rightarrow \text{Bool} \\
\text{id}_1 &: \text{Int} \rightarrow \text{Int} \\
\end{align*}
\]

When can we generalize?
A mini Language
to study Hindley-Milner Types

Expressions

\[
E ::= c \quad \text{constant} \\
| \ x \quad \text{variable} \\
| \ \lambda x. \ E \quad \text{abstraction} \\
| \ (E_1 \ E_2) \quad \text{application} \\
| \ \text{let} \ x = E_1 \ \text{in} \ E_2 \quad \text{let-block}
\]

• There are no types in the syntax of the language!

• The type of each subexpression is derived by the Hindley-Milner type inference algorithm.
## A Formal Type System

<table>
<thead>
<tr>
<th><strong>Types</strong></th>
<th><strong>Type Schemes</strong></th>
<th><strong>Type Environments</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>( \tau ) ::= ( i )</td>
<td>( \sigma ) ::= ( \tau )</td>
<td>( \text{TE} ::= \text{Identifiers} \rightarrow \text{Type Schemes} )</td>
</tr>
<tr>
<td></td>
<td>( \mid t )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( \mid \tau_1 \rightarrow \tau_2 )</td>
<td></td>
</tr>
<tr>
<td>base types</td>
<td>type variables</td>
<td></td>
</tr>
<tr>
<td>type variables</td>
<td>Function types</td>
<td></td>
</tr>
</tbody>
</table>

Note, all the \( \forall \)'s occur in the beginning of a type scheme, i.e., a type \( \tau \) cannot contain a type scheme \( \sigma \)
**Instantiations**

\[
\sigma = \forall t_1 \ldots t_n. \tau
\]

- Type scheme \(\sigma\) can be *instantiated* into a type \(\tau'\) by substituting types for the *bound variables* of \(\sigma\), i.e.,

\[
\tau' = S \tau
\]

for some \(S\) s.t. \(\text{Dom}(S) \subseteq \text{BV}(\sigma)\)

- \(\tau'\) is said to be an *instance of* \(\sigma\) (\(\sigma > \tau'\))

- \(\tau'\) is said to be a *generic instance of* \(\sigma\) when \(S\) maps variables to new variables.

Example:

\[
\sigma = \forall t_1. \ t_1 \rightarrow t_2
\]

\(t_3 \rightarrow t_2\) is a generic instance of \(\sigma\)

\(\text{Int} \rightarrow t_2\) is a non generic instance of \(\sigma\)
Generalization \textit{aka} Closing

\[ \text{Gen}(\text{TE}, \tau) = \forall t_1 \ldots t_n. \ \tau \]
\[ \text{where } \{ t_1 \ldots t_n \} = \text{FV}(\tau) - \text{FV}(\text{TE}) \]

- \textit{Generalization} introduces polymorphism
- Quantify type variables that are free in $\tau$ but not \textit{free} in the type environment (TE)
- Captures the notion of \textit{new} type variables of $\tau$
HM Type Inference Rules

(App) \( \Gamma \vdash e_1 : \tau \rightarrow \tau' \quad \Gamma \vdash e_2 : \tau \) \( \Gamma \vdash (e_1 e_2) : \tau' \) Remember, \( \tau \) stands for a monotype, \( \sigma \) for a polymorphic type

(Abs) \( \Gamma ; \{x: \tau\} \vdash e : \tau' \) \( \Gamma \vdash \lambda x. e : \tau \rightarrow \tau' \) \( x \) can be considered of type \( \tau \) as long as its type as specified in the environment can be specialized to \( \tau \) (i.e. \( \tau \) is an instance of \( \sigma \))

(Var) \( (x: \sigma) \in \Gamma \) \( \sigma \geq \tau \) \( \Gamma \vdash x : \tau \) \( x \) has a different type in \( e_1 \) than in \( e_2 \). In \( e_1 \), \( x \) is not a polymorphic type, but in \( e_2 \) it gets generalized into one.

(Const) \( \) \( \text{typeof}(c) \geq \tau \) \( \Gamma \vdash c : \tau \) 

(Let) \( \Gamma ; \{x: \tau\} \vdash e_1 : \tau \quad \Gamma ; \{x: \text{Gen}(\Gamma, \tau)\} \vdash e_2 : \tau' \) \( \Gamma \vdash (\text{let } x = e_1 \text{ in } e_2) : \tau' \)
HM Inference Algorithm

\[
\text{Def } W(TE, e) = \begin{cases} 
\text{Case } e \text{ of} \\
\text{c} & = (\{\}, \text{Typeof(c)}) \\
x & = \begin{cases} 
\text{if } (x \notin \text{Dom}(TE)) \text{ then } \text{Fail} \\
\text{else let } \forall t_1...t_n.\tau = TE(x); \\
\text{in } (\{\}, [u_i \mapsto t_i] \tau) 
\end{cases} \\
\lambda x.e & = \begin{cases} 
\text{let } (S_1, \tau_1) = W(TE + \{x : u\}, e); \\
\text{in } (S_1, S_1(u) \rightarrow \tau_1) 
\end{cases} \\
(e_1 e_2) & = \begin{cases} 
\text{let } (S_1, \tau_1) = W(TE, e_1); \\
(S_2, \tau_2) = W(S_1(TE), e_2); \\
S_3 = \text{Unify}(S_2(\tau_1), \tau_2 \rightarrow u); \\
\text{in } (S_3 S_2 S_1, S_3(u)) 
\end{cases} \\
\text{let } x = e_1 \text{ in } e_2 & = \begin{cases} 
\text{let } (S_1, \tau_1) = W(TE + \{x : u\}, e_1); \\
S_2 = \text{Unify}(S_1(u), \tau_1); \\
\sigma = \text{Gen}(S_2 S_1(TE), S_2(\tau_1)); \\
(S_3, \tau_2) = W(S_2 S_1(TE) + \{x : \sigma\}, e_2); \\
\text{in } (S_3 S_2 S_1, \tau_2) 
\end{cases}
\]

u’s represent new type variables

Hindley-Milner: Example

\[
\lambda x. \text{let } f = \lambda y. x \text{ in } (f 1, f \text{ True})
\]

\[
W(\emptyset, A) = ([], u_1 \to (u_1, u_1))
\]

\[
W(\{x : u_1\}, B) = ([], (u_1, u_1))
\]

\[
W(\{x : u_1, f : u_2\}, \lambda y. x) = ([], u_3 \to u_1)
\]

\[
W(\{x : u_1, f : u_2, y : u_3\}, x) = ([], u_1)
\]

Unify(u_2, u_3 \to u_1) = [(u_3 \to u_1) / u_2]

Gen(\{x : u_1\}, u_3 \to u_1) = \forall u_3. u_3 \to u_1

TE = \{x : u_1, f : \forall u_3. u_3 \to u_1\}

W(TE, (f 1)) = ([], u_1)

W(TE, f) = ([], u_4 \to u_1)

W(TE, 1) = ([], \text{Int})

Unify(u_4 \to u_1, \text{Int} \to u_5) = [\text{Int} / u_4, u_1 / u_5]

...
Important Observations

- Do not generalize over type variables used elsewhere
- Let is the only way of defining polymorphic constructs
- Generalize the types of let-bound identifiers only after processing their definitions
Properties of HM Type Inference

• It is sound with respect to the type system. An inferred type is verifiable.

• It generates most general types of expressions. Any verifiable type is inferred.

• Complexity
  PSPACE-Hard
  Nested let blocks
Extensions

• Type Declarations
  Sanity check; can relax restrictions

• Incremental Type checking
  The whole program is not given at the same time, sound inferencing when types of some functions are not known

• Typing references to mutable objects
  Hindley-Milner system is unsound for a language with refs (mutable locations)

• Overloading Resolution
HM Limitations:
\(\lambda\)-bound vs Let-bound Variables

Only let-bound identifiers can be instantiated differently.

\[
\text{let } \quad \text{twice } f \ x = f \ (f \ x) \\quad \text{in} \\
\text{twice twice succ 4}
\]

versus

\[
\text{let } \quad \text{twice } f \ x = f \ (f \ x) \quad \text{foo } g = (g \ g \ g \ succ) \ 4 \\
\text{in} \\
\text{foo twice}
\]

foo is not type correct!

Generic vs. Non-generic type variables
### Puzzle: Another set of Inference rules

<table>
<thead>
<tr>
<th>Rule</th>
<th>Premise</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Gen)</td>
<td>( \text{TE} \vdash e : \tau \quad \tau \notin \text{FV(TE)} )</td>
<td>( \text{TE} \vdash e : \forall t.\tau )</td>
</tr>
<tr>
<td>(Spec)</td>
<td>( \text{TE} \vdash e : \forall t.\tau )</td>
<td>( \text{TE} \vdash e : \tau [u/t] )</td>
</tr>
<tr>
<td>(Var)</td>
<td>((x : \tau) \in \text{TE})</td>
<td>( \text{TE} \vdash x : \tau )</td>
</tr>
<tr>
<td>(Let)</td>
<td>( \text{TE} + {x : \tau} \vdash e_1 : \tau \quad \text{TE} + {x : \tau} \vdash e_2 : \tau' )</td>
<td>( \text{TE} \vdash (\text{let } x = e_1 \text{ in } e_2) : \tau' )</td>
</tr>
</tbody>
</table>

- Sound but no inference algorithm!

(App) and (Abs) rules remain unchanged.