Lecture 15 — Automatic Type Inference (3)

CSE307: Programming Languages

Minseok Jeon

2026 Spring

Finding a Solution of Type Equations

Find the values of type variables that make all the equations true.

$$\underbrace{\texttt{proc}\ (\underbrace{f}_{t_f})\ \underbrace{\texttt{proc}\ (\underbrace{x}_{t_x})\ \underbrace{(\underbrace{(f\ 3)}_{t_3} - \underbrace{(f\ x)}_{t_4})}_{t_2}}_{t_1}}_{t_0}$$

Equations			Solution		
t_0	$=$	$t_f \to t_1$	t_0	$=$	$(\text{int} \to \text{int}) \to (\text{int} \to \text{int})$
t_1	$=$	$t_x \to t_2$	t_1	$=$	$\text{int} \to \text{int}$
t_3	$=$	int	t_2	$=$	int
t_4	$=$	int	t_3	$=$	int
t_2	$=$	int	t_4	$=$	int
t_f	$=$	$\text{int} \to t_3$	t_f	$=$	$\text{int} \to \text{int}$
t_f	$=$	$t_x \to t_4$	t_x	$=$	int

Static type systems find such a solution using *unification algorithm*.

Example 1

The calculation is split into equations to be solved and substitution found so far. Initially, the substitution is empty:

Equations	Substitution
$t_0 = t_f \to t_1$	
$t_1 = t_x \to t_2$	
$t_3 = \text{int}$	
$t_4 = \text{int}$	
$t_2 = \text{int}$	
$t_f = \text{int} \to t_3$	
$t_f = t_x \to t_4$	

Example 1

Consider each equation in turn. If the equation's left-hand side is a variable, move it to the substitution:

Equations	Substitution
$t_1 = t_x \to t_2$	$t_0 = t_f \to t_1$
$t_3 = \text{int}$	
$t_4 = \text{int}$	
$t_2 = \text{int}$	
$t_f = \text{int} \to t_3$	
$t_f = t_x \to t_4$	

Example 1

Move the next equation to the substitution and propagate the information to the existing substitution (i.e., substitute the right-hand side for each occurrence of t_1):

Equations	Substitution
$t_3 = \text{int}$	$t_0 = t_f \to (t_x \to t_2)$
$t_4 = \text{int}$	$t_1 = t_x \to t_2$
$t_2 = \text{int}$	
$t_f = \text{int} \to t_3$	
$t_f = t_x \to t_4$	

Example 1

Same for the next three equations:

Equations			Substitution		
t_4	$=$	int	t_0	$=$	$t_f \to (t_x \to t_2)$
t_2	$=$	int	t_1	$=$	$t_x \to t_2$
t_f	$=$	$\text{int} \to t_3$	t_3	$=$	int
t_f	$=$	$t_x \to t_4$			

Equations			Substitution		
t_2	$=$	int	t_0	$=$	$t_f \to (t_x \to t_2)$
t_f	$=$	$\text{int} \to t_3$	t_1	$=$	$t_x \to t_2$
t_f	$=$	$t_x \to t_4$	t_3	$=$	int
			t_4	$=$	int

Example 1

Same for the next three equations:

Equations	Substitution
$t_f = \text{int} \to t_3$	$t_0 = t_f \to (t_x \to \text{int})$
$t_f = t_x \to t_4$	$t_1 = t_x \to \text{int}$
	$t_3 = \text{int}$
	$t_4 = \text{int}$
	$t_2 = \text{int}$

Example 1

Consider the next equation $t_f = \text{int} \to t_3$. The equation contains t_3, which is already bound to int in the substitution. Substitute int for t_3 in the equation. This is called *applying* the substitution to the equation.

Equations	Substitution
$t_f = \text{int} \to \text{int}$	$t_0 = t_f \to (t_x \to \text{int})$
$t_f = t_x \to t_4$	$t_1 = t_x \to \text{int}$
	$t_3 = \text{int}$
	$t_4 = \text{int}$
	$t_2 = \text{int}$

Move the resulting equation to the substitution and update it.

Equations	Substitution
$t_f = t_x \to t_4$	$t_0 = (\text{int} \to \text{int}) \to (t_x \to \text{int})$
	$t_1 = t_x \to \text{int}$
	$t_3 = \text{int}$
	$t_4 = \text{int}$
	$t_2 = \text{int}$
	$t_f = \text{int} \to \text{int}$

Example 1

Apply the substitution to the equation:

Equations			Substitution		
$\text{int} \to \text{int}$	$=$	$t_x \to \text{int}$	t_0	$=$	$(\text{int} \to \text{int}) \to (t_x \to \text{int})$
			t_1	$=$	$t_x \to \text{int}$
			t_3	$=$	int
			t_4	$=$	int
			t_2	$=$	int
			t_f	$=$	$\text{int} \to \text{int}$

If neither side of the equation is a variable, simplify the equation by yielding two new equations:

Equations			Substitution		
int	$=$	t_x	t_0	$=$	$(\text{int} \to \text{int}) \to (t_x \to \text{int})$
int	$=$	int	t_1	$=$	$t_x \to \text{int}$
			t_3	$=$	int
			t_4	$=$	int
			t_2	$=$	int
			t_f	$=$	$\text{int} \to \text{int}$

Example 1

Switch the sides of the first equation and move it to the substitution:

Equations	Substitution
int $=$ int	$t_0 = (\text{int} \to \text{int}) \to (\text{int} \to \text{int})$
	$t_1 = \text{int} \to \text{int}$
	$t_3 = \text{int}$
	$t_4 = \text{int}$
	$t_2 = \text{int}$
	$t_f = \text{int} \to \text{int}$
	$t_x = \text{int}$

The final substitution is the solution of the original equations.

Example 2

$$\underbrace{\texttt{proc}\ (\underbrace{f}_{t_f})\ (\underbrace{f\ 11}_{t_1})}_{t_0}$$

$$\begin{aligned} t_0 &= t_f \to t_1 \\ t_f &= \text{int} \to t_1 \end{aligned}$$

Example 2

1.

Equations	Substitution
$t_0 = t_f \to t_1$	
$t_f = \text{int} \to t_1$	

2.

Equations	Substitution
$t_f = \text{int} \to t_1$	$t_0 = t_f \to t_1$

3.

Equations	Substitution
	$t_0 = (\text{int} \to t_1) \to t_1$
	$t_f = \text{int} \to t_1$

The type is *polymorphic* in t_1.

Example 3

$$\underbrace{\texttt{if}\ \underbrace{x}_{t_x}\ \texttt{then}\ \underbrace{(\underbrace{x}_{t_x} - \underbrace{1}_{t_2})}_{t_1}\ \texttt{else}\ \underbrace{0}_{t_3}}_{t_0}$$

$$\begin{aligned}
t_x &= \text{bool} \\
t_1 &= t_0 \\
t_3 &= t_0 \\
t_1 &= \text{int} \\
t_x &= \text{int} \\
t_2 &= \text{int} \\
t_3 &= \text{int}
\end{aligned}$$

Example 3

The equations have no solutions because, during the unification algorithm, we encounter the following contradictory state:

Equations	Substitution
$\text{bool} = \text{int}$	$\boldsymbol{t}_x = \text{bool}$
$\boldsymbol{t}_2 = \text{int}$	$\boldsymbol{t}_1 = \boldsymbol{t}_0$
$\boldsymbol{t}_3 = \text{int}$	$\boldsymbol{t}_3 = \boldsymbol{t}_0$
	$\boldsymbol{t}_0 = \text{int}$

Because bool and int cannot be equal, there is no solution to the equations.

Example 4

$$\underbrace{\texttt{proc}\ (\underbrace{f}_{t_f})\ (\underbrace{\texttt{iszero}\ (\underbrace{(\underbrace{f}_{t_f}\ \underbrace{f}_{t_f})}_{t_2})}_{t_1})}_{t_0}$$

$$\begin{aligned} t_0 &= t_f \to t_1 \\ t_1 &= \text{bool} \\ t_2 &= \text{int} \\ t_f &= t_f \to t_2 \end{aligned}$$

Example 4

Solving as usual, we encounter a problem:

Equations	Substitution
$t_f = t_f \to \text{int}$	$t_0 = t_f \to \text{bool}$
	$t_1 = \text{bool}$
	$t_2 = \text{int}$

- There is no type t_f that satisfies the equation, because the right-hand side of the equation is always larger than the left.
- If we ever deduce an equation of the form $t = \ldots t \ldots$ where the type variable t occurs in the right-hand side, we must conclude that there is no solution. This is called *occurrence check*.

Unification Algorithm

For each equation in turn,

- Apply the current substitution to the equation.
- If the equation is always true (e.g. $\mathsf{int} = \mathsf{int}$), discard it.
- If the left- and right-hand sides are contradictory (e.g. $\mathsf{bool} = \mathsf{int}$), the algorithm fails.
- If neither side is a variable (e.g. $\mathsf{int} \to \boldsymbol{t}_1 = \boldsymbol{t}_2 \to \mathsf{bool}$), simplify the equation, which eventually generates an equation whose left- or right-hand side is a variable.
- If the left-hand side is not a variable, switch the sides.
- If the left-hand side variable occurs in the right-hand side, the algorithm fails.
- Otherwise, move it to the substitution and substitute the right-hand side for each occurrence of the variable in the substitution.

Exercise 1

$$\underbrace{\texttt{let}\ \underbrace{x}_{t_x} = \underbrace{4}_{t_1}\ \texttt{in}\ \underbrace{(\underbrace{x}_{t_x}\ \underbrace{3}_{t_3})}_{t_2}}_{t_0}$$

Equations	Substitution
$t_x = t_1$	
$t_0 = t_2$	
$t_1 = \text{int}$	
$t_x = t_3 \to t_2$	
$t_3 = \text{int}$	

Exercise 1: Unification

Apply the unification algorithm to the equations. Initially, the substitution is empty:

Equations	Substitution
$t_x = t_1$	
$t_0 = t_2$	
$t_1 = \text{int}$	
$t_x = t_3 \to t_2$	
$t_3 = \text{int}$	

Exercise 1: Unification

The first three equations all have a variable on the left. Move them to the substitution in turn; when $\boldsymbol{t}_1 = \text{int}$ is moved, propagate int for $\boldsymbol{t}_1$ into the existing entry $\boldsymbol{t}_x = \boldsymbol{t}_1$:

Equations	Substitution
$\boldsymbol{t}_x = \boldsymbol{t}_3 \to \boldsymbol{t}_2$	$\boldsymbol{t}_x = \text{int}$
$\boldsymbol{t}_3 = \text{int}$	$\boldsymbol{t}_0 = \boldsymbol{t}_2$
	$\boldsymbol{t}_1 = \text{int}$

Exercise 1: Unification

The next equation is $\boldsymbol{t_x} = \boldsymbol{t}_3 \to \boldsymbol{t}_2$, but $\boldsymbol{t_x}$ is already bound to int in the substitution. Apply the substitution to the equation:

Equations			Substitution		
int	$=$	$\boldsymbol{t}_3 \to \boldsymbol{t}_2$	$\boldsymbol{t_x}$	$=$	int
$\boldsymbol{t}_3$	$=$	int	$\boldsymbol{t}_0$	$=$	$\boldsymbol{t}_2$
			$\boldsymbol{t}_1$	$=$	int

Exercise 1: Unification (Failure)

The equation int $= \boldsymbol{t}_3 \to \boldsymbol{t}_2$ has a base type on the left and a function type on the right — they are contradictory and cannot be unified.

Unification fails. The expression `let` $x = 4$ `in` $(x\ 3)$ is *not typable*: applying the integer 4 to the argument 3 is a type error.

Exercise 2

$$\underbrace{\texttt{let}\ \underbrace{f}_{t_f} = \underbrace{\texttt{proc}\ (\underbrace{z}_{t_z})\ \underbrace{z}_{t_z}}_{t_1}\ \texttt{in}\ \underbrace{\texttt{proc}\ (\underbrace{x}_{t_x})\ \underbrace{(\underbrace{(\underbrace{f}_{t_f}\ \underbrace{x}_{t_x})}_{t_4} - \underbrace{1}_{t_5})}_{t_3}}_{t_2}}_{t_0}$$

Exercise 2

Equations			Substitution
t_0	$=$	t_2	
t_f	$=$	t_1	
t_1	$=$	$t_z \to t_z$	
t_2	$=$	$t_x \to t_3$	
t_3	$=$	int	
t_4	$=$	int	
t_5	$=$	int	
t_f	$=$	$t_x \to t_4$	

Exercise 2: Unification

Apply the unification algorithm to the equations. Initially, the substitution is empty:

Equations			Substitution
t_0	$=$	t_2	
t_f	$=$	t_1	
t_1	$=$	$t_z \to t_z$	
t_2	$=$	$t_x \to t_3$	
t_3	$=$	int	
t_4	$=$	int	
t_5	$=$	int	
t_f	$=$	$t_x \to t_4$	

Exercise 2: Unification

The first two equations have a variable on the left; move them to the substitution:

Equations			Substitution		
t_1	$=$	$t_z \to t_z$	t_0	$=$	t_2
t_2	$=$	$t_x \to t_3$	t_f	$=$	t_1
t_3	$=$	int			
t_4	$=$	int			
t_5	$=$	int			
t_f	$=$	$t_x \to t_4$			

Exercise 2: Unification

Move $t_1 = t_z \to t_z$ to the substitution and propagate (t_1 appears in the entry $t_f = t_1$):

Equations	Substitution
$t_2 = t_x \to t_3$	$t_0 = t_2$
$t_3 = \text{int}$	$t_f = t_z \to t_z$
$t_4 = \text{int}$	$t_1 = t_z \to t_z$
$t_5 = \text{int}$	
$t_f = t_x \to t_4$	

Exercise 2: Unification

Move $t_2 = t_x \to t_3$ to the substitution and propagate (t_2 appears in the entry $t_0 = t_2$):

Equations			Substitution		
t_3	$=$	int	t_0	$=$	$t_x \to t_3$
t_4	$=$	int	t_f	$=$	$t_z \to t_z$
t_5	$=$	int	t_1	$=$	$t_z \to t_z$
t_f	$=$	$t_x \to t_4$	t_2	$=$	$t_x \to t_3$

Exercise 2: Unification

Move $t_3 = \text{int}$ (propagate to t_0 and t_2), then $t_4 = \text{int}$ and $t_5 = \text{int}$:

Equations	Substitution
$t_f = t_x \to t_4$	$t_0 = t_x \to \text{int}$
	$t_f = t_z \to t_z$
	$t_1 = t_z \to t_z$
	$t_2 = t_x \to \text{int}$
	$t_3 = \text{int}$
	$t_4 = \text{int}$
	$t_5 = \text{int}$

Exercise 2: Unification

Apply the substitution to $t_f = t_x \to t_4$: since $t_f \mapsto t_z \to t_z$ and $t_4 \mapsto \text{int}$, the equation becomes $t_z \to t_z = t_x \to \text{int}$. Neither side is a variable; simplify by decomposing the function types:

Equations	Substitution
$t_z = t_x$	$t_0 = t_x \to \text{int}$
$t_z = \text{int}$	$t_f = t_z \to t_z$
	$t_1 = t_z \to t_z$
	$t_2 = t_x \to \text{int}$
	$t_3 = \text{int}$
	$t_4 = \text{int}$
	$t_5 = \text{int}$

Exercise 2: Unification

Move $t_z = t_x$ to the substitution and propagate (t_z appears in entries t_f and t_1):

Equations	Substitution
$t_z = \text{int}$	$t_0 = t_x \to \text{int}$
	$t_f = t_x \to t_x$
	$t_1 = t_x \to t_x$
	$t_2 = t_x \to \text{int}$
	$t_3 = \text{int}$
	$t_4 = \text{int}$
	$t_5 = \text{int}$
	$t_z = t_x$

Exercise 2: Unification

Apply the substitution to $t_z = \text{int}$: $t_z \mapsto t_x$, so the equation becomes $t_x = \text{int}$. Move it to the substitution and propagate everywhere:

Equations	Substitution
	$t_0 = \text{int} \to \text{int}$
	$t_f = \text{int} \to \text{int}$
	$t_1 = \text{int} \to \text{int}$
	$t_2 = \text{int} \to \text{int}$
	$t_3 = \text{int}$
	$t_4 = \text{int}$
	$t_5 = \text{int}$
	$t_z = \text{int}$
	$t_x = \text{int}$

The whole expression has type $t_0 = \text{int} \to \text{int}$, so
$\texttt{let } f = \texttt{proc } (z)\ z \texttt{ in proc } (x)\ ((f\ x) - 1)$ is typable with type $\text{int} \to \text{int}$.

Exercise 3

$$\underbrace{\texttt{let } \underbrace{\boldsymbol{p}}_{t_p} = \underbrace{\texttt{iszero } \underbrace{\texttt{1}}_{t_3}}_{t_1} \texttt{ in } \underbrace{\texttt{if } \underbrace{\boldsymbol{p}}_{t_p} \texttt{ then } \underbrace{\texttt{88}}_{t_4} \texttt{ else } \underbrace{\texttt{99}}_{t_5}}_{t_2}}_{t_0}$$

Exercise 3

Equations	Substitution
$t_0 = t_2$	
$t_p = t_1$	
$t_1 = \text{bool}$	
$t_3 = \text{int}$	
$t_p = \text{bool}$	
$t_4 = t_2$	
$t_5 = t_2$	
$t_4 = \text{int}$	
$t_5 = \text{int}$	

Exercise 4

$$\underbrace{\texttt{let}\ \underbrace{f}_{t_f} = \underbrace{\texttt{proc}\ (\underbrace{x}_{t_x})\ \underbrace{x}_{t_x}}_{t_1}\ \texttt{in}\ \underbrace{\texttt{if}\ \underbrace{(\underbrace{f}_{t_f}\ \underbrace{(\texttt{iszero}\ \underbrace{0}_{t_5})}_{t_4})}_{t_3}\ \texttt{then}\ \underbrace{(\underbrace{f}_{t_f}\ \underbrace{11}_{t_7})}_{t_6}\ \texttt{else}\ \underbrace{(\underbrace{f}_{t_f}\ \underbrace{22}_{t_9})}_{t_8}}_{t_2}}_{t_0}$$

Exercise 4

Equations			Substitution
t_0	$=$	t_2	
t_f	$=$	t_1	
t_1	$=$	$t_x \rightarrow t_x$	
t_3	$=$	bool	
t_6	$=$	t_2	
t_8	$=$	t_2	
t_f	$=$	$t_4 \rightarrow t_3$	
t_4	$=$	bool	
t_5	$=$	int	
t_f	$=$	$t_7 \rightarrow t_6$	
t_f	$=$	$t_9 \rightarrow t_8$	
t_7	$=$	int	
t_9	$=$	int	

Exercise 4: Unification

Apply the unification algorithm. Initially, the substitution is empty:

Equations	Substitution
$t_0 = t_2$	
$t_f = t_1$	
$t_1 = t_x \rightarrow t_x$	
$t_3 = \text{bool}$	
$t_6 = t_2$	
$t_8 = t_2$	
$t_f = t_4 \rightarrow t_3$	
$t_4 = \text{bool}$	
$t_5 = \text{int}$	
$t_f = t_7 \rightarrow t_6$	
$t_f = t_9 \rightarrow t_8$	
$t_7 = \text{int}$	
$t_9 = \text{int}$	

Exercise 4: Unification

Move each of the first six LHS-variable equations in turn. When $t_1 = t_x \to t_x$ is moved, it propagates into the entry $t_f = t_1$:

Equations			Substitution		
t_f	$=$	$t_4 \to t_3$	t_0	$=$	t_2
t_4	$=$	bool	t_f	$=$	$t_x \to t_x$
t_5	$=$	int	t_1	$=$	$t_x \to t_x$
t_f	$=$	$t_7 \to t_6$	t_3	$=$	bool
t_f	$=$	$t_9 \to t_8$	t_6	$=$	t_2
t_7	$=$	int	t_8	$=$	t_2
t_9	$=$	int			

Exercise 4: Unification

Process $t_f = t_4 \to t_3$. Apply the substitution: $t_f \mapsto t_x \to t_x$ and $t_3 \mapsto \text{bool}$, so the equation becomes $t_x \to t_x = t_4 \to \text{bool}$. Neither side is a variable; decompose:

Equations			Substitution		
t_x	$=$	t_4	t_0	$=$	t_2
t_x	$=$	bool	t_f	$=$	$t_x \to t_x$
t_4	$=$	bool	t_1	$=$	$t_x \to t_x$
t_5	$=$	int	t_3	$=$	bool
t_f	$=$	$t_7 \to t_6$	t_6	$=$	t_2
t_f	$=$	$t_9 \to t_8$	t_8	$=$	t_2
t_7	$=$	int			
t_9	$=$	int			

Exercise 4: Unification

Move $t_x = t_4$ (propagate into t_f, t_1); apply substitution to $t_x = \text{bool}$ to get $t_4 = \text{bool}$ and move (propagate); the old equation $t_4 = \text{bool}$ now becomes $\text{bool} = \text{bool}$, discard; move $t_5 = \text{int}$:

Equations			Substitution		
t_f	$=$	$t_7 \to t_6$	t_0	$=$	t_2
t_f	$=$	$t_9 \to t_8$	t_f	$=$	$\text{bool} \to \text{bool}$
t_7	$=$	int	t_1	$=$	$\text{bool} \to \text{bool}$
t_9	$=$	int	t_3	$=$	bool
			t_6	$=$	t_2
			t_8	$=$	t_2
			t_x	$=$	bool
			t_4	$=$	bool
			t_5	$=$	int

Exercise 4: Unification

Process $t_f = t_7 \to t_6$. Apply substitution: bool $\to$ bool $= t_7 \to t_2$. Decompose, switch sides, and move $t_7 =$ bool then $t_2 =$ bool (propagate $t_2 \mapsto$ bool into t_0, t_6, t_8):

Equations			Substitution		
t_f	$=$	$t_9 \to t_8$	t_0	$=$	bool
t_7	$=$	int	t_f	$=$	bool $\to$ bool
t_9	$=$	int	t_1	$=$	bool $\to$ bool
			t_3	$=$	bool
			t_6	$=$	bool
			t_8	$=$	bool
			t_x	$=$	bool
			t_4	$=$	bool
			t_5	$=$	int
			t_7	$=$	bool
			t_2	$=$	bool

Exercise 4: Unification

Process $t_f = t_9 \to t_8$. Apply substitution: bool $\to$ bool $= t_9 \to$ bool. Decompose: bool $= t_9$ (switch & move as $t_9 =$ bool), and bool $=$ bool (discard):

Equations	Substitution
$t_7 = \text{int}$	$t_0 = \text{bool}$
$t_9 = \text{int}$	$t_f = \text{bool} \to \text{bool}$
	$t_1 = \text{bool} \to \text{bool}$
	$t_3 = \text{bool}$
	$t_6 = \text{bool}$
	$t_8 = \text{bool}$
	$t_x = \text{bool}$
	$t_4 = \text{bool}$
	$t_5 = \text{int}$
	$t_7 = \text{bool}$
	$t_2 = \text{bool}$
	$t_9 = \text{bool}$

Exercise 4: Unification (Failure)

Process $t_7 = \text{int}$. Apply the substitution: $t_7 \mapsto \text{bool}$, so the equation becomes $\text{bool} = \text{int}$. Neither side is a variable, and the two sides are contradictory base types.

Unification fails. The expression
$\texttt{let}\ f = \texttt{proc}\ (x)\ x\ \texttt{in}\ \texttt{if}\ (f\ (\texttt{iszero}\ 0))\ \texttt{then}\ (f\ 11)\ \texttt{else}\ (f\ 22)$ is *not typable*.

Why? The identity function f is used both on a Boolean (in the condition) and on integers (in the branches). The monomorphic type system can only assign f a single type, but no single type can be both $\text{bool} \to \text{bool}$ and $\text{int} \to \text{int}$. This is the classic motivating example for *let-polymorphism*.

Substitution

Solutions of type equations are represented by substitution:

$$\boldsymbol{S} \in \mathit{Subst} = \mathit{TyVar} \to \boldsymbol{T}$$

Applying a substitution to a type:

$$\begin{array}{rcl}
\boldsymbol{S}(\mathrm{int}) & = & \mathrm{int} \\
\boldsymbol{S}(\mathrm{bool}) & = & \mathrm{bool} \\
\boldsymbol{S}(\boldsymbol{\alpha}) & = & \begin{cases} \boldsymbol{t} & \text{if } \boldsymbol{\alpha} \mapsto \boldsymbol{t} \in \boldsymbol{S} \\ \boldsymbol{\alpha} & \text{otherwise} \end{cases} \\
\boldsymbol{S}(\boldsymbol{T}_1 \to \boldsymbol{T}_2) & = & \boldsymbol{S}(\boldsymbol{T}_1) \to \boldsymbol{S}(\boldsymbol{T}_2)
\end{array}$$

Example

Applying the substitution

$$S = \{t_1 \mapsto \text{int}, t_2 \mapsto \text{int} \to \text{int}\}$$

to to the type $(t_1 \to t_2) \to (t_3 \to \text{int})$:

$$\begin{aligned}
&S((t_1 \to t_2) \to (t_3 \to \text{int})) \\
&= S(t_1 \to t_2) \to S(t_3 \to \text{int}) \\
&= (S(t_1) \to S(t_2)) \to (S(t_3) \to S(\text{int})) \\
&= (\text{int} \to (\text{int} \to \text{int})) \to (t_3 \to \text{int})
\end{aligned}$$

Unification

Update the current substitution with equality $t_1 \doteq t_2$.

$$\textbf{unify} : T \times T \times \mathit{Subst} \rightarrow \mathit{Subst}$$

$$
\begin{array}{rcl}
\textbf{unify}(\text{int}, \text{int}, S) & = & S \\
\textbf{unify}(\text{bool}, \text{bool}, S) & = & S \\
\textbf{unify}(\alpha, \alpha, S) & = & S \\
\textbf{unify}(\alpha, t, S) & = & \begin{cases} \text{fail} & \alpha \text{ occurs in } t \\ \text{extend } S \text{ with } \alpha \doteq t & \text{otherwise} \end{cases} \\
\textbf{unify}(t, \alpha, S) & = & \textbf{unify}(\alpha, t, S) \\
\textbf{unify}(t_1 \rightarrow t_2, t_1' \rightarrow t_2', S) & = & \text{let } S' = \textbf{unify}(t_1, t_1', S) \text{ in} \\
 & & \text{let } S'' = \textbf{unify}(S'(t_2), S'(t_2'), S') \text{ in} \\
 & & \quad S'' \\
\textbf{unify}(\_, \_, \_) & = & \text{fail}
\end{array}
$$

Exercises

- $\textbf{unify}(\boldsymbol{\alpha}, \text{int} \to \text{int}, \emptyset) = \{\boldsymbol{\alpha} \mapsto \text{int} \to \text{int}\}$
- $\textbf{unify}(\boldsymbol{\alpha}, \text{int} \to \boldsymbol{\alpha}, \emptyset) = \text{fail} \quad (\text{occurs check: } \boldsymbol{\alpha} \text{ occurs in int} \to \boldsymbol{\alpha})$
- $\textbf{unify}(\boldsymbol{\alpha} \to \boldsymbol{\beta}, \text{int} \to \text{int}, \emptyset) = \{\boldsymbol{\alpha} \mapsto \text{int}, \boldsymbol{\beta} \mapsto \text{int}\}$
- $\textbf{unify}(\boldsymbol{\alpha} \to \boldsymbol{\beta}, \text{int} \to \boldsymbol{\alpha}, \emptyset) = \{\boldsymbol{\alpha} \mapsto \text{int}, \boldsymbol{\beta} \mapsto \text{int}\}$

Solving Equations

$$\textbf{unifyall} : \mathit{TyEqn} \rightarrow \mathit{Subst} \rightarrow \boldsymbol{Subst}$$

$$\begin{array}{rcl} \textbf{unifyall}(\emptyset, \boldsymbol{S}) & = & \boldsymbol{S} \\ \textbf{unifyall}((t_1 \doteq t_2) \wedge \boldsymbol{u}, \boldsymbol{S}) & = & \text{let } \boldsymbol{S}' = \textbf{unify}(\boldsymbol{S}(t_1), \boldsymbol{S}(t_2), \boldsymbol{S}) \\ & & \text{in } \textbf{unifyall}(\boldsymbol{u}, \boldsymbol{S}') \end{array}$$

Let $\mathcal{U}$ be the final unification algorithm:

$$\mathcal{U}(\boldsymbol{u}) = \textbf{unifyall}(\boldsymbol{u}, \emptyset)$$

typeof : $E \to T$

The final type inference algorithm that composes equation derivation ($\mathcal{V}$) and equation solving ($\mathcal{U}$):

$$
\begin{aligned}
&\textbf{typeof}(E) = \\
&\quad \textbf{let } S = \mathcal{U}(\mathcal{V}(\emptyset, E, \boldsymbol{\alpha})) \quad (\text{new } \boldsymbol{\alpha}) \\
&\quad \textbf{in } S(\boldsymbol{\alpha})
\end{aligned}
$$

Examples

- **typeof**$((\texttt{proc}\ (x)\ x)\ 1) =$
- **typeof**$(\texttt{let}\ x = 1\ \texttt{in}\ \texttt{proc}(y)\ (x + y)) =$

Summary

Automatic type inference:

- derive type equations from the program text, and
- solve the equations by unification algorithm.