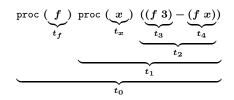
## COSE212: Programming Languages

Lecture 11 — Automatic Type Inference (2)

Hakjoo Oh 2016 Fall

## Finding a Solution of Type Equations

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



Equations			Solution		
$t_0$	=	$t_f  ightarrow t_1$	$t_0$	=	$(int \to int) \to (int \to int)$
$t_1$	=	$t_x  ightarrow t_2$	$t_1$	=	int  o int
$t_3$	=	int	$t_2$	=	int
$t_4$	=	int	$t_3$	=	int
$t_2$	=	int		=	
$t_f$	=	$int \to t_3$	$ t_f $	=	int  o int
$t_f$	=	$t_x  ightarrow t_4$		=	

Such a solution can be found by the unification algorithm.

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

Equations	Substitution
$t_0 = t_f  o t_1$	
$t_1 = t_x  ightarrow t_2$	
$t_3 = int$	
$t_4 = int$	
$t_2 = int$	
$t_f \; = \; int  o t_3$	
$t_f = t_x  ightarrow t_4$	

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

Equations	Substitution		
$t_1 = t_x  ightarrow t_2$	$t_0 = t_f \rightarrow t_1$		
$t_3 \; = \; int$			
$t_4 = int$			
$t_2 = int$			
$t_f \; = \; int  o t_3$			
$t_f = t_x  o t_4$			

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			
$t_3 = \text{int}$	$t_0 = t_f  ightarrow (t_x  ightarrow t_2)$			
$t_4 \;=\; int$	$t_1 = t_x  ightarrow t_2$			
$t_{2}$ = int				
$t_f$ = int $ o t_3$				
$t_f = t_x \rightarrow t_4$				

Same for the next three equations:

Equations	Substitution
$t_4$ = int	$t_0 = t_f  ightarrow (t_x  ightarrow t_2)$
$t_{2}$ = int	$egin{array}{cccc} t_1 & = & t_x  ightarrow t_2 \end{array}$
$t_f$ $=$ int $ o t_3$	$t_3$ = int
$t_f = t_x  ightarrow t_4$	
Equations	Substitution
$t_2 = int$	$t_0 = t_f  ightarrow (t_x  ightarrow t_2)$
$t_f$ $=$ int $ o t_3$	$t_1 = t_x \rightarrow t_2$
$t_f = t_x  ightarrow t_4$	$t_3$ = int
•	$\mid t_4 \mid = int$
Equations	Substitution
$t_f = {\sf int}  o t_3$	$t_0 = t_f  ightarrow (t_x  ightarrow { m int})$
$t_f^- = t_x  ightarrow t_4$	$t_1 \; = \; t_x  o int$
	$t_3 = int$
	$egin{array}{lll} t_4 &=& { m int} \ t_2 &=& { m int} \end{array}$
	$\mid t_{2} \mid = \mid$ int

Consider the next equation  $t_f={\rm int}\to t_3$ . 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		
$t_f = int  o int$	$t_0 = t_f  ightarrow (t_x  ightarrow { ext{int}})$		
$t_f = t_x  ightarrow t_4$	$egin{array}{lll} t_1 &=& t_x  ightarrow  ext{int} \ t_3 &=&  ext{int} \ t_4 &=&  ext{int} \ \end{array}$		
	$t_3$ = int		
	$t_4 = int$		
	$t_2$ = int		

Move the resulting equation to the substitution and update it.

		Equations			Substitution
$\overline{t_f}$	=	$t_x  ightarrow t_4$	$t_0$	=	$(int  o int)  o (t_x  o int)$
			$t_1$	=	$t_x  ightarrow$ int
			$t_3$	=	int
			$t_4$	=	int
			$t_2$	=	int
			$\mid t_f \mid$	=	$(\operatorname{int}  o \operatorname{int})  o (t_x  o \operatorname{int})$ $t_x  o \operatorname{int}$ $\operatorname{int}$ $\operatorname{int}$ $\operatorname{int}$ $\operatorname{int}$ $\operatorname{int}$ $\operatorname{int}$ $\operatorname{int}  o \operatorname{int}$

Apply the substitution to the equation:

Equations	Substitution
$int  o int \ = \ t_x  o int$	$t_0 = (int  o int)  o (t_x  o int)$
	$t_1 = t_x  o int$
	$t_3$ = int
	$t_4$ = int
	$t_2$ = int
	$egin{array}{lll} t_0 &=& (\operatorname{int}  ightarrow \operatorname{int})  ightarrow (t_x  ightarrow \operatorname{int}) \ t_1 &=& t_x  ightarrow \operatorname{int} \ t_3 &=& \operatorname{int} \ t_4 &=& \operatorname{int} \ t_2 &=& \operatorname{int} \ t_f &=& \operatorname{int}  ightarrow \operatorname{int}  i$

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

Equations			Substitution		
	=	$t_x$	$t_0$	=	$(int  o int)  o (t_x  o int)$
int	=	int	$t_1$	=	$t_{m{x}}  ightarrow int$
			$t_3$	=	int
			$t_4$	=	int
			$t_2$	=	int
			$t_f$	=	$(\operatorname{int}  o \operatorname{int})  o (t_x  o \operatorname{int})$ $t_x  o \operatorname{int}$ $\operatorname{int}$ $\operatorname{int}$ $\operatorname{int}$ $\operatorname{int}$ $\operatorname{int}$ $\operatorname{int}  o \operatorname{int}$

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

Equations	Substitution			
int = int	$t_0$	=	$(int \to int) \to (int \to int)$	
	$t_1$	=	int  o int	
	$t_3$	=	int	
	$t_4$	=	int	
	$t_2$	=	int	
	$t_f$	=	int  o int	
	$t_{x}$	=	$\begin{array}{l} \operatorname{int} \to \operatorname{int} \\ \operatorname{int} \\ \operatorname{int} \\ \operatorname{int} \\ \operatorname{int} \\ \operatorname{int} \\ \operatorname{int} \to \operatorname{int} \\ \operatorname{int} \end{array}$	

The final substitution is the solution of the original equations.

## Unification Algorithm

For each equation in turn,

- Apply the current substitution to the equation.
- If the left-hand side is a variable, move it to the substitution and substitute the right-hand side for each occurrence of the variable in the substitution. (If the right-hand side is a variable, switch the sides and do the same thing).

Basically, the algorithm follows the two steps. Two execptions: If neither side is a variable, simplify the equation, which eventually generates an equation whose left- or right-hand side is a variable. If the equation is always true, discard it.

$$\begin{array}{cccc} \operatorname{proc} & \underbrace{\begin{pmatrix} f \\ t_f \end{pmatrix}}_{t_f} & \underbrace{\begin{pmatrix} f & 11 \\ t_1 \end{pmatrix}}_{t_0} \\ \\ t_0 & = & t_f \to t_1 \\ t_f & = & \operatorname{int} \to t_1 \end{array}$$

 $\begin{array}{c|cccc} & & & & & & \\ \hline t_0 & = & t_f \rightarrow t_1 & & & \\ t_f & = & \mathsf{int} \rightarrow t_1 & & & \\ \end{array}$ 

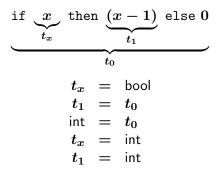
Equations	Substitution		
$t_f = int  o t_1$	$egin{array}{cccc} t_0 &=& t_f  ightarrow t_1 \end{array}$		

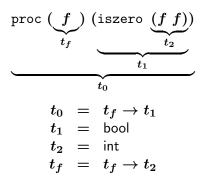
Equations Substitution 
$$\begin{array}{c|cccc} \mathsf{Equations} & \mathsf{Substitution} \\ & t_0 & = & (\mathsf{int} \to t_1) \to t_1 \\ & t_f & = & \mathsf{int} \to t_1 \end{array}$$

The type is *polymorphic* in  $t_1$ .

•

2





#### Exercises

For each following expression, perform the type inference and find its type, or determine that no such type exists.

- **1** let x = 4 in  $(x \ 3)$
- 2 let f = proc(z) z in proc(x) ((f x) 1)
- 3 let p = iszero 1 in if p then 88 else 99
- $\textcircled{\scriptsize 1} \texttt{let} \ f \ = \texttt{proc} \ (x) \ x \ \texttt{in} \ \texttt{if} \ (f \ (\texttt{iszero0})) \ \texttt{then} \ (f \ 11) \ \texttt{else} \ (f \ 22)$

## Summary

#### Automatic type inference:

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