Universiteit Utrecht

[Faculty of **Science**
Information and Computing Sciences]

Annotated Type Systems

Stolen from Stefan Holdermans

Dept. of Information and Computing Sciences, Utrecht University
P.O. Box 80.089, 3508 TB Utrecht, The Netherlands
E-mail: jur@cs.uu.nl

Type and effect systems - Introduction

Type-based approaches to static program analysis

- Static program analysis: compile-time techniques for approximating the set of values or behaviours that arise at run-time when a program is executed.
- Applications: verification, optimization.
- Different approaches: data-flow analysis, constraint-based analysis, abstract interpretation, type-based analysis.
- Type-based analysis: equipping a programming language with a nonstandard type system that keeps track of some properties of interest.
- Advantages: reuse of tools, techniques, and infrastructure (polymorphism, subtyping, type inference, . . .).
- Focus: accuracy vs. modularity.

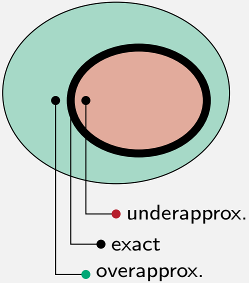
Examples

...
Side-effect analysis.
Callability analysis.
Reachability analysis.
Uniqueness analysis.
Totality analysis.
Security analysis.
Strictness analysis.
Sharing analysis.
Trust analysis.
...

...
Sign analysis.
Flow analysis.
Control-flow analysis.
Class-hierarchy analysis.
Region analysis.
Binding-time analysis.
Alias analysis.
Communication analysis.
Escape analysis.
...

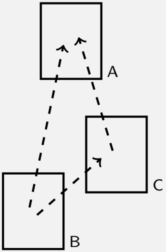
Accuracy

- Establishing nontrivial properties of programs is in general undecidable (halting problem, Rice's theorem).
- In static analysis we have to settle for "useful" approximations of properties.
- "Useful" means: sound ("erring at the safe side") and accurate (as precise as possible).



Modularity

- Breaking up a (large) program in smaller units or modules is generally considered good programming style.
- Separate compilation: compile each module in isolation.
- Advantage: only modules that have been edited need to be recompiled.
- To facilitate seperate compilation, each unit of compilation needs to be analysed in isolation, i.e., without knowledge of how it's used from within the rest of the program.



☞ Tension between accuracy and modularity: whole-program analysis typically yields more precise results.

Hindley-Milner and Algorithm W

A simple functional language

f, x	$\in$	$\mathbf{Var}$	variables
t	$\in$	$\mathbf{Tm}$	terms

$$
\begin{array}{lll}
t & ::= & \mid x \mid \lambda\ x.\,t_1 \\
 & \mid & t_1\ t_2 \\
 & \mid &
\end{array}
$$

Universiteit Utrecht

[Faculty of **Science**
Information and Computing Sciences]

A simple functional language

f, x	$\in$	$\mathbf{Var}$	variables
π	$\in$	$\mathbf{Pnt}$	program points
t	$\in$	$\mathbf{Tm}$	terms

$$
\begin{array}{lll}
t & ::= & \quad\quad\quad \mid x \mid \lambda_{\pi} x.\, t_1 \\
 & \mid & t_1\ t_2 \\
 & \mid &
\end{array}
$$

A simple functional language

f, x	$\in$	$\mathbf{Var}$	variables
π	$\in$	$\mathbf{Pnt}$	program points
t	$\in$	$\mathbf{Tm}$	terms

$$
\begin{array}{lll}
t & ::= & \mid x \mid \lambda_{\pi} x.\, t_1 \\
 & \mid & t_1\ t_2 \qquad \mid \mathbf{let}\ x = t_1\ \mathbf{in}\ t_2 \\
 & \mid &
\end{array}
$$

A simple functional language

$$
\begin{array}{lcll}
f, x & \in & \mathbf{Var} & \text{variables} \\
\pi & \in & \mathbf{Pnt} & \text{program points} \\
t & \in & \mathbf{Tm} & \text{terms}
\end{array}
$$

$$
\begin{array}{lll}
t & ::= & \mid x \mid \lambda_\pi x.\, t_1 \mid \mu f.\, \lambda_\pi x.\, t_1 \\
 & \mid & t_1\ t_2 \qquad \mid \mathbf{let}\ x = t_1\ \mathbf{in}\ t_2 \\
 & \mid &
\end{array}
$$

A simple functional language

$$
\begin{array}{rcll}
n & \in & \mathbf{Num} = \mathbb{N} & \text{numerals} \\
f, x & \in & \mathbf{Var} & \text{variables} \\
\\
\pi & \in & \mathbf{Pnt} & \text{program points} \\
t & \in & \mathbf{Tm} & \text{terms}
\end{array}
$$

$$
\begin{array}{rcll}
t & ::= & n & \mid x \mid \lambda_\pi x.\, t_1 \mid \mu f.\, \lambda_\pi x.\, t_1 \\
 & \mid & t_1\ t_2 & \mid \mathbf{let}\ x = t_1\ \mathbf{in}\ t_2 \\
 & \mid & &
\end{array}
$$

A simple functional language

$$\begin{array}{rcll} n & \in & \mathbf{Num} = \mathbb{N} & \text{numerals} \\ f, x & \in & \mathbf{Var} & \text{variables} \\ \\ \pi & \in & \mathbf{Pnt} & \text{program points} \\ t & \in & \mathbf{Tm} & \text{terms} \end{array}$$

$$\begin{array}{rcl} t & ::= & n \mid \texttt{false} \mid \texttt{true} \mid x \mid \lambda_\pi x.\, t_1 \mid \mu f.\, \lambda_\pi x.\, t_1 \\ & \mid & t_1\ t_2 \mid \mathbf{if}\ t_1\ \mathbf{then}\ t_2\ \mathbf{else}\ t_3 \mid \mathbf{let}\ x = t_1\ \mathbf{in}\ t_2 \\ & \mid & \end{array}$$

A simple functional language

n	$\in$	$\mathbf{Num} = \mathbb{N}$	numerals
f, x	$\in$	$\mathbf{Var}$	variables
$\oplus$	$\in$	$\mathbf{Op}$	binary operators
π	$\in$	$\mathbf{Pnt}$	program points
t	$\in$	$\mathbf{Tm}$	terms

$$
\begin{array}{lcl}
t & ::= & n \mid \texttt{false} \mid \texttt{true} \mid x \mid \lambda_\pi x.\, t_1 \mid \mu f.\lambda_\pi x.\, t_1 \\
 & \mid & t_1\ t_2 \mid \textbf{if}\ t_1\ \textbf{then}\ t_2\ \textbf{else}\ t_3 \mid \textbf{let}\ x = t_1\ \textbf{in}\ t_2 \\
 & \mid & t_1 \oplus t_2
\end{array}
$$

A simple functional language

n	$\in$	$\mathbf{Num} = \mathbb{N}$	numerals
f, x	$\in$	$\mathbf{Var}$	variables
$\oplus$	$\in$	$\mathbf{Op}$	binary operators
π	$\in$	$\mathbf{Pnt}$	program points
t	$\in$	$\mathbf{Tm}$	terms

$$
\begin{array}{lcl}
t & ::= & n \mid \mathtt{false} \mid \mathtt{true} \mid x \mid \lambda_\pi x.\, t_1 \mid \mu f.\, \lambda_\pi x.\, t_1 \\
 & \mid & t_1\ t_2 \mid \mathbf{if}\ t_1\ \mathbf{then}\ t_2\ \mathbf{else}\ t_3 \mid \mathbf{let}\ x = t_1\ \mathbf{in}\ t_2 \\
 & \mid & t_1 \oplus t_2
\end{array}
$$

Example:

$$
\begin{array}{l}
\mathbf{let}\ \mathit{fac} = \mu f.\, \lambda_{\mathrm{F}} x.\, \mathbf{if}\ x \equiv 0\ \mathbf{then}\ 1\ \mathbf{else}\ x * f\ (x - 1) \\
\mathbf{in}\ \ \mathit{fac}\ 6
\end{array}
$$

Monomorphic types

$$\tau \quad \in \quad \mathbf{Ty} \qquad \text{types}$$

$$\tau \quad ::= \quad \mathit{Nat} \mid \mathit{Bool} \mid \tau_1 \to \tau_2$$

Monomorphic types

$$\begin{array}{lll} \tau & \in & \mathbf{Ty} \\ \Gamma & \in & \mathbf{TyEnv} \end{array}$$

τ	$\in$	$\mathbf{Ty}$	types
Γ	$\in$	$\mathbf{TyEnv}$	type environments

$$\begin{array}{lll} \tau & ::= & \mathit{Nat} \mid \mathit{Bool} \mid \tau_1 \to \tau_2 \\ \Gamma & ::= & [\,] \mid \Gamma_1[x \mapsto \tau] \end{array}$$

Monomorphic types

τ	$\in$	$\mathbf{Ty}$	types
Γ	$\in$	$\mathbf{TyEnv}$	type environments

$$
\begin{array}{lcl}
\tau & ::= & \mathit{Nat} \mid \mathit{Bool} \mid \tau_1 \to \tau_2 \\
\Gamma & ::= & [\,] \mid \Gamma_1[x \mapsto \tau]
\end{array}
$$

Typing judgements:

$\Gamma \vdash_{\mathrm{UL}} t : \tau$	typing

"Term t has type τ assuming that any of its free variables has the type given by Γ."

Monomorphic type system: constants

$$\frac{}{\Gamma \vdash_{\mathrm{UL}} n : Nat} \; [\textit{t-num}]$$

Monomorphic type system: constants

$$\frac{}{\Gamma \vdash_{\mathrm{UL}} n : \mathit{Nat}}\ [\textit{t-num}]$$

$$\frac{}{\Gamma \vdash_{\mathrm{UL}} \texttt{false} : \mathit{Bool}}\ [\textit{t-false}]$$

$$\frac{}{\Gamma \vdash_{\mathrm{UL}} \texttt{true} : \mathit{Bool}}\ [\textit{t-true}]$$

Monomorphic type system: variables

$$\frac{\Gamma(x) = \tau}{\Gamma \vdash_{\mathrm{UL}} x : \tau} \; [\textit{t-var}]$$

Monomorphic type system: functions

$$\frac{\Gamma[x \mapsto \tau_1] \vdash_{\mathrm{UL}} t_1 : \tau_2}{\Gamma \vdash_{\mathrm{UL}} \lambda_\pi x.\, t_1 : \tau_1 \to \tau_2}\ [\textit{t-lam}]$$

Monomorphic type system: functions

$$\frac{\Gamma[x \mapsto \tau_1] \vdash_{\mathrm{UL}} t_1 : \tau_2}{\Gamma \vdash_{\mathrm{UL}} \lambda_\pi x.\, t_1 : \tau_1 \to \tau_2} \; [\textit{t-lam}]$$

$$\frac{\Gamma[f \mapsto (\tau_1 \to \tau_2)][x \mapsto \tau_1] \vdash_{\mathrm{UL}} t_1 : \tau_2}{\Gamma \vdash_{\mathrm{UL}} \mu f.\, \lambda_\pi x.\, t_1 : \tau_1 \to \tau_2} \; [\textit{t-mu}]$$

Monomorphic type system: functions

$$\frac{\Gamma[x \mapsto \tau_1] \vdash_{\mathrm{UL}} t_1 : \tau_2}{\Gamma \vdash_{\mathrm{UL}} \lambda_\pi x.\, t_1 : \tau_1 \to \tau_2}\ [\textit{t-lam}]$$

$$\frac{\Gamma[f \mapsto (\tau_1 \to \tau_2)][x \mapsto \tau_1] \vdash_{\mathrm{UL}} t_1 : \tau_2}{\Gamma \vdash_{\mathrm{UL}} \mu f.\, \lambda_\pi x.\, t_1 : \tau_1 \to \tau_2}\ [\textit{t-mu}]$$

$$\frac{\Gamma \vdash_{\mathrm{UL}} t_1 : \tau_2 \to \tau \quad \Gamma \vdash_{\mathrm{UL}} t_2 : \tau_2}{\Gamma \vdash_{\mathrm{UL}} t_1\ t_2 : \tau}\ [\textit{t-app}]$$

Monomorphic type system: conditionals

$$\frac{\Gamma \vdash_{\mathrm{UL}} t_1 : \mathit{Bool} \quad \Gamma \vdash_{\mathrm{UL}} t_2 : \tau \quad \Gamma \vdash_{\mathrm{UL}} t_3 : \tau}{\Gamma \vdash_{\mathrm{UL}} \mathbf{if}\ t_1\ \mathbf{then}\ t_2\ \mathbf{else}\ t_3 : \tau}\ [\textit{t-if}]$$

Monomorphic type system: local definitions

$$\frac{\Gamma \vdash_{\text{UL}} t_1 : \tau_1 \quad \Gamma[x \mapsto \tau_1] \vdash_{\text{UL}} t_2 : \tau}{\Gamma \vdash_{\text{UL}} \textbf{let } x = t_1 \textbf{ in } t_2 : \tau} \ [\textit{t-let}]$$

Monomorphic type system: binary operators

$$\frac{\Gamma \vdash_{\mathrm{UL}} t_1 : \tau_{\oplus}^{1} \quad \Gamma \vdash_{\mathrm{UL}} t_2 : \tau_{\oplus}^{2}}{\Gamma \vdash_{\mathrm{UL}} t_1 \oplus t_2 : \tau_{\oplus}} \; [\textit{t-op}]$$

Monomorphic type system: example

$$\frac{}{\Gamma \vdash_{\text{UL}} \mu f.\, \lambda_{\text{F}} x.\, \textbf{if } x \equiv 0 \textbf{ then } 1 \textbf{ else } x * f\ (x - 1) : \mathit{Nat} \to \mathit{Nat}}$$

Monomorphic type system: example

$$\dfrac{\dfrac{\vdots}{\Gamma_{\mathrm{F}} \vdash_{\mathrm{UL}} x \equiv 0 : \mathit{Bool}} \quad \dfrac{}{\Gamma_{\mathrm{F}} \vdash_{\mathrm{UL}} 1 : \mathit{Nat}} \quad \dfrac{\vdots}{\Gamma_{\mathrm{F}} \vdash_{\mathrm{UL}} x * f\ (x-1) : \mathit{Nat}}}{\dfrac{\Gamma_{\mathrm{F}} \vdash_{\mathrm{UL}} \mathbf{if}\ x \equiv 0\ \mathbf{then}\ 1\ \mathbf{else}\ x * f\ (x-1) : \mathit{Nat}}{\Gamma \vdash_{\mathrm{UL}} \mu f.\lambda_{\mathrm{F}} x.\,\mathbf{if}\ x \equiv 0\ \mathbf{then}\ 1\ \mathbf{else}\ x * f\ (x-1) : \mathit{Nat} \to \mathit{Nat}}}$$

$$\Gamma_{\mathrm{F}} = \Gamma[f \mapsto (\mathit{Nat} \to \mathit{Nat})][x \mapsto \mathit{Nat}]$$

Polymorphic functions

Polymorphic functions

$\lambda_{\mathrm{F}} x.\, x$

Polymorphic functions

$\lambda_{\mathrm{F}} x.\, x$

$\lambda_{\mathrm{F}} x.\, \lambda_{\mathrm{G}} y.\, x$

Polymorphic functions

$$\lambda_{\mathrm{F}} x.\, x$$

$$\lambda_{\mathrm{F}} x.\, \lambda_{\mathrm{G}} y.\, x$$

$$\lambda_{\mathrm{F}} f.\, \lambda_{\mathrm{G}} x.\, f\ x$$

Polymorphic functions

$$\lambda_{\mathrm{F}} x.\, x$$

$$\lambda_{\mathrm{F}} x.\, \lambda_{\mathrm{G}} y.\, x$$

$$\lambda_{\mathrm{F}} f.\, \lambda_{\mathrm{G}} x.\, f\ x$$

$$\mu f.\, \lambda_{\mathrm{F}} g.\, \lambda_{\mathrm{G}} x.\, \lambda_{\mathrm{H}} y.\, \mathbf{if}\ x \equiv 0\ \mathbf{then}\ y\ \mathbf{else}\ f\ g\ (x-1)\ (g\ y)$$

Polymorphic types

$\tau \in \mathbf{Ty}$ types

$\Gamma \in \mathbf{TyEnv}$ type environments

$$\tau ::= \quad | \; Nat \; | \; Bool \; | \; \tau_1 \to \tau_2$$

$$\Gamma ::= [\,] \; | \; \Gamma_1[x \mapsto \tau]$$

$\Gamma \vdash_{\mathrm{UL}} t : \tau$ typing

Polymorphic types

α	$\in$	**TyVar**	type variables
τ	$\in$	**Ty**	types
Γ	$\in$	**TyEnv**	type environments

$$\tau \quad ::= \quad \alpha \mid \mathit{Nat} \mid \mathit{Bool} \mid \tau_1 \rightarrow \tau_2$$

$$\Gamma \quad ::= \quad [\,] \mid \Gamma_1[x \mapsto \tau]$$

$\Gamma \vdash_{\mathrm{UL}} t : \tau$ typing

Polymorphic types

α	$\in$	$\mathbf{TyVar}$	type variables
τ	$\in$	$\mathbf{Ty}$	types
σ	$\in$	$\mathbf{TyScheme}$	type schemes
Γ	$\in$	$\mathbf{TyEnv}$	type environments

$$\tau ::= \alpha \mid \mathit{Nat} \mid \mathit{Bool} \mid \tau_1 \to \tau_2$$

$$\sigma ::= \tau \mid \forall \alpha.\, \sigma_1$$

$$\Gamma ::= [\,] \mid \Gamma_1[x \mapsto \tau]$$

$\Gamma \vdash_{\mathrm{UL}} t : \tau$ typing

Universiteit Utrecht

[Faculty of **Science**
Information and Computing Sciences]

Polymorphic types

α	$\in$	**TyVar**	type variables
τ	$\in$	**Ty**	types
σ	$\in$	**TyScheme**	type schemes
Γ	$\in$	**TyEnv**	type environments

$$\begin{array}{lcl}
\tau & ::= & \alpha \mid \mathit{Nat} \mid \mathit{Bool} \mid \tau_1 \rightarrow \tau_2 \\
\sigma & ::= & \tau \mid \forall\alpha.\,\sigma_1 \\
\Gamma & ::= & [\,] \mid \Gamma_1[x \mapsto \sigma]
\end{array}$$

$\Gamma \vdash_{\mathrm{UL}} t : \tau$ typing

Polymorphic types

$$\begin{array}{llll} \alpha & \in & \mathbf{TyVar} & \text{type variables} \\ \tau & \in & \mathbf{Ty} & \text{types} \\ \sigma & \in & \mathbf{TyScheme} & \text{type schemes} \\ \Gamma & \in & \mathbf{TyEnv} & \text{type environments} \end{array}$$

$$\begin{array}{lll} \tau & ::= & \alpha \mid \mathit{Nat} \mid \mathit{Bool} \mid \tau_1 \to \tau_2 \\ \sigma & ::= & \tau \mid \forall \alpha .\, \sigma_1 \\ \Gamma & ::= & [\,] \mid \Gamma_1[x \mapsto \sigma] \end{array}$$

$\Gamma \vdash_{\mathrm{UL}} t : \sigma$ typing

Polymorphic types

α	$\in$	**TyVar**	type variables
τ	$\in$	**Ty**	types
σ	$\in$	**TyScheme**	type schemes
Γ	$\in$	**TyEnv**	type environments

$$
\begin{array}{lll}
\tau & ::= & \alpha \mid \mathit{Nat} \mid \mathit{Bool} \mid \tau_1 \to \tau_2 \\
\sigma & ::= & \tau \mid \forall\alpha.\,\sigma_1 \\
\Gamma & ::= & [\,] \mid \Gamma_1[x \mapsto \sigma]
\end{array}
$$

$\Gamma \vdash_{\mathrm{UL}} t : \sigma$ typing

☞ $\mathbf{Ty} \subseteq \mathbf{TyScheme}$

Polymorphic type system: generalisation and instantiation

Introduction:

$$\frac{\Gamma \vdash_{\mathrm{UL}} t : \sigma_1 \quad \alpha \notin \mathit{ftv}(\Gamma)}{\Gamma \vdash_{\mathrm{UL}} t : \forall \alpha.\, \sigma_1} \; [\textit{t-gen}]$$

Polymorphic type system: generalisation and instantiation

Introduction:

$$\frac{\Gamma \vdash_{\text{UL}} t : \sigma_1 \qquad \alpha \notin \mathit{ftv}(\Gamma)}{\Gamma \vdash_{\text{UL}} t : \forall \alpha.\, \sigma_1} \; [\textit{t-gen}]$$

Elimination:

$$\frac{\Gamma \vdash_{\text{UL}} t : \forall \alpha.\, \sigma_1}{\Gamma \vdash_{\text{UL}} t : [\alpha \mapsto \tau_0]\sigma_1} \; [\textit{t-inst}]$$

Polymorphic type system: variables and local definitions

$$\frac{\Gamma(x) = \sigma}{\Gamma \vdash_{\mathrm{UL}} x : \sigma}\ [\mathit{t\text{-}var}]$$

Polymorphic type system: variables and local definitions

$$\frac{\Gamma(x) = \sigma}{\Gamma \vdash_{\mathrm{UL}} x : \sigma}\ [\textit{t-var}]$$

$$\frac{\Gamma \vdash_{\mathrm{UL}} t_1 : \sigma_1 \quad \Gamma[x \mapsto \sigma_1] \vdash_{\mathrm{UL}} t_2 : \tau}{\Gamma \vdash_{\mathrm{UL}} \mathbf{let}\ x = t_1\ \mathbf{in}\ t_2 : \tau}\ [\textit{t-let}]$$

Polymorphic types: example

$\lambda_{F} x.\, x : \forall \alpha.\, \alpha \to \alpha$

$\lambda_{F} x.\, \lambda_{G} y.\, x : \forall \alpha_1.\, \forall \alpha_2.\, \alpha_1 \to \alpha_2 \to \alpha_1$

$\lambda_{F} f.\, \lambda_{G} x.\, f\ x : \forall \alpha_1.\, \forall \alpha_2.\, (\alpha_1 \to \alpha_2) \to \alpha_1 \to \alpha_2$

$$\begin{aligned} &\mu f.\, \lambda_{F} g.\, \lambda_{G} x.\, \lambda_{H} y.\, \textbf{if } x \equiv 0 \textbf{ then } y \textbf{ else } f\ g\ (x-1)\ (g\ y) \\ &\quad : \forall \alpha.\, (\alpha \to \alpha) \to \mathit{Nat} \to \alpha \to \alpha \end{aligned}$$

Inference algorithm

$$\theta \quad \in \quad \mathbf{TySubst} = \mathbf{TyVar} \rightarrow_{\mathsf{fin}} \mathbf{Ty} \qquad \text{type substitution}$$

$$\begin{array}{lcll}
\mathit{generalise}_{\mathrm{UL}} & : & \mathbf{TyEnv} \times \mathbf{Ty} & \rightarrow \mathbf{TyScheme} \\
\mathit{instantiate}_{\mathrm{UL}} & : & \mathbf{TyScheme} & \rightarrow \mathbf{Ty} \\
\mathcal{U}_{\mathrm{UL}} & : & \mathbf{Ty} \times \mathbf{Ty} & \rightarrow \mathbf{TySubst} \\
\mathcal{W}_{\mathrm{UL}} & : & \mathbf{TyEnv} \times \mathbf{Tm} & \rightarrow \mathbf{Ty} \times \mathbf{TySubst}
\end{array}$$

Inference algorithm: constants

$$\mathcal{W}_{\mathrm{UL}}(\Gamma, n) = (Nat, \quad \mathit{id})$$

Inference algorithm: constants

$$\mathcal{W}_{\mathrm{UL}}(\Gamma, n) = (Nat, \quad \mathit{id})$$

$$\mathcal{W}_{\mathrm{UL}}(\Gamma, \mathtt{false}) = (Bool, \quad \mathit{id})$$

$$\mathcal{W}_{\mathrm{UL}}(\Gamma, \mathtt{true}) = (Bool, \quad \mathit{id})$$

Inference algorithm: variables

$$\mathcal{W}_{\mathrm{UL}}\ (\Gamma, x) = (\mathit{instantiate}_{\mathrm{UL}}(\Gamma(x)), \quad \mathit{id})$$

- The instantiation rule is built into the case for variables.
- By choosing fresh type variables, we commit to nothing,
- and let the actual types be determined by future unifications.

Inference algorithm: functions

$$
\begin{aligned}
\mathcal{W}_{\mathrm{UL}}\ (\Gamma, \lambda_{\pi} x.\, t_1) = {}& \text{let } \alpha_1 \text{ be fresh} \\
& \qquad (\tau_2, \theta) = \mathcal{W}_{\mathrm{UL}}(\Gamma[x \mapsto \alpha_1], t_1) \\
& \text{in } ((\theta\ \alpha_1) \to \tau_2, \quad \theta)
\end{aligned}
$$

Inference algorithm: functions

$$\begin{aligned}
\mathcal{W}_{\text{UL}}\ (\Gamma, \lambda_{\pi} x.\, t_1) = \ & \text{let } \alpha_1 \text{ be fresh} \\
& \qquad (\tau_2, \theta) = \mathcal{W}_{\text{UL}}(\Gamma[x \mapsto \alpha_1], t_1) \\
& \text{in } ((\theta\ \alpha_1) \to \tau_2, \quad \theta)
\end{aligned}$$

$$\begin{aligned}
& \mathcal{W}_{\text{UL}}\ (\Gamma, \mu f.\, \lambda_{\pi} x.\, t_1) = \\
& \quad \text{let } \alpha_1, \alpha_2 \text{ be fresh} \\
& \qquad (\tau_2, \theta_1) = \mathcal{W}_{\text{UL}}(\Gamma[f \mapsto (\alpha_1 \to \alpha_2)][x \mapsto \alpha_1], t_1) \\
& \qquad \theta_2 = \mathcal{U}_{\text{UL}}(\tau_2, \theta_1\ \alpha_2) \\
& \quad \text{in } (\theta_2\ (\theta_1\ \alpha_1) \to \theta_2\ \tau_2, \quad \theta_2 \circ \theta_1)
\end{aligned}$$

Inference algorithm: functions

$$\begin{aligned}\mathcal{W}_{\mathrm{UL}}\ (\Gamma, \lambda_\pi x.\ t_1) = \text{let } & \alpha_1 \text{ be fresh} \\ & (\tau_2, \theta) = \mathcal{W}_{\mathrm{UL}}(\Gamma[x \mapsto \alpha_1], t_1) \\ \text{in } & ((\theta\ \alpha_1) \to \tau_2, \quad \theta)\end{aligned}$$

$$\begin{aligned}&\mathcal{W}_{\mathrm{UL}}\ (\Gamma, \mu f.\lambda_\pi x.\ t_1) = \\ &\quad \text{let } \alpha_1, \alpha_2 \text{ be fresh} \\ &\qquad (\tau_2, \theta_1) = \mathcal{W}_{\mathrm{UL}}(\Gamma[f \mapsto (\alpha_1 \to \alpha_2)][x \mapsto \alpha_1], t_1) \\ &\qquad \theta_2 = \mathcal{U}_{\mathrm{UL}}(\tau_2, \theta_1\ \alpha_2) \\ &\quad \text{in } (\theta_2\ (\theta_1\ \alpha_1) \to \theta_2\ \tau_2, \quad \theta_2 \circ \theta_1)\end{aligned}$$

$$\begin{aligned}\mathcal{W}_{\mathrm{UL}}\ (\Gamma, t_1\ t_2) = \text{let } & (\tau_1, \theta_1) = \mathcal{W}_{\mathrm{UL}}(\Gamma, t_1) \\ & (\tau_2, \theta_2) = \mathcal{W}_{\mathrm{UL}}(\theta_1\ \Gamma, t_2) \\ & \alpha \text{ be fresh} \\ & \theta_3 = \mathcal{U}_{\mathrm{UL}}(\theta_2\ \tau_1, \tau_2 \to \alpha) \\ \text{in } & (\theta_3\ \alpha, \quad \theta_3 \circ \theta_2 \circ \theta_1)\end{aligned}$$

Universiteit Utrecht

[Faculty of Science
Information and Computing Sciences]

Unification

- To combine (join) two given types we apply unification
- I.e., in case rule for applications, $\mathcal{U}_{\mathrm{UL}}(\theta_2\ \tau_1, \tau_2 \to \alpha)$
- Unification computes a substitution from two types: $\mathcal{U}_{\mathrm{UL}} : \mathbf{Ty} \times \mathbf{Ty} \to \mathbf{TySubst}$
- If $\mathcal{U}_{\mathrm{UL}}(t_1, t_2) = \theta$ then $\theta\ t_1 = \theta\ t_2$
  - And θ is the least such substitution
- Ex. $\mathcal{U}_{\mathrm{UL}}(\alpha_1 \to \mathit{Nat} \to \mathit{Bool}, \mathit{Nat} \to \mathit{Nat} \to \alpha_2)$ equals θ with $\theta(\alpha_1) = \mathit{Nat}$ and $\theta(\alpha_2) = \mathit{Bool}$
- Note: unification is basically the $\sqcup$ in the lattice of monotypes

Unification Algorithm

$$
\begin{aligned}
&\mathcal{U}_{\mathrm{UL}}\ (\mathit{Nat},\ \mathit{Nat}) \ = \mathit{id}\\
&\mathcal{U}_{\mathrm{UL}}\ (\mathit{Bool}, \mathit{Bool}) = \mathit{id}\\
&\mathcal{U}_{\mathrm{UL}}\ (\tau_1 \to \tau_2, \tau_3 \to \tau_4) = \theta_2 \circ \theta_1\\
&\quad \textbf{where}\\
&\qquad \theta_1 = \mathcal{U}_{\mathrm{UL}}\ (\tau_1, \tau_3)\\
&\qquad \theta_2 = \mathcal{U}_{\mathrm{UL}}\ (\theta_1\ \tau_2, \theta_1\ \tau_4)\\
&\mathcal{U}_{\mathrm{UL}}\ (\alpha, \tau) = [\alpha \mapsto \tau]\ \textbf{if}\ \mathit{chk}\ (\alpha, \tau)\\
&\mathcal{U}_{\mathrm{UL}}\ (\tau, \alpha) = [\alpha \mapsto \tau]\ \textbf{if}\ \mathit{chk}\ (\alpha, \tau)\\
&\mathcal{U}_{\mathrm{UL}}\ (\_, \_) \ = \mathsf{fail}
\end{aligned}
$$

Here, $\mathit{chk}\ (\alpha, \tau)$ returns true if $\tau = \alpha$ or α is not a free variable in τ.

Universiteit Utrecht

[Faculty of **Science**
Information and Computing Sciences]

Inference algorithm: conditionals

$$
\begin{array}{l}
\mathcal{W}_{\mathrm{UL}}(\Gamma, \textbf{if } t_1 \textbf{ then } t_2 \textbf{ else } t_3) = \\
\quad \text{let } (\tau_1, \theta_1) = \mathcal{W}_{\mathrm{UL}}(\Gamma, t_1) \\
\qquad (\tau_2, \theta_2) = \mathcal{W}_{\mathrm{UL}}(\theta_1\ \Gamma, t_2) \\
\qquad (\tau_3, \theta_3) = \mathcal{W}_{\mathrm{UL}}(\theta_2\ (\theta_1\ \Gamma), t_3) \\
\qquad \theta_4 = \mathcal{U}_{\mathrm{UL}}(\theta_3\ (\theta_2\ \tau_1), \mathit{Bool}) \\
\qquad \theta_5 = \mathcal{U}_{\mathrm{UL}}(\theta_4\ (\theta_3\ \tau_2), \theta_4\ \tau_3) \\
\quad \text{in } (\theta_5\ (\theta_4\ \tau_3), \quad \theta_5 \circ \theta_4 \circ \theta_3 \circ \theta_2 \circ \theta_1)
\end{array}
$$

- Subsitutions are applied as soon as possible.
- Error prone process of putting the right composition of substitutions everywhere.
- Substitutions are idempotent: blindly applying all of them all the time can only influence efficiency.

Inference algorithm: local definitions

$$
\begin{array}{l}
\mathcal{W}_{\mathrm{UL}}(\Gamma, \textbf{let}\ x = t_1\ \textbf{in}\ t_2) = \\
\quad \mathsf{let}\ (\tau_1, \theta_1) = \mathcal{W}_{\mathrm{UL}}(\Gamma, t_1) \\
\qquad\quad (\tau, \theta_2) \ = \mathcal{W}_{\mathrm{UL}}((\theta_1\ \Gamma)[x \mapsto \textit{generalise}_{\mathrm{UL}}(\theta_1\ \Gamma, \tau_1)], t_2) \\
\quad \mathsf{in}\ \ (\tau, \quad \theta_2 \circ \theta_1)
\end{array}
$$

Inference algorithm: binary operators

$$
\begin{aligned}
&\mathcal{W}_{\mathrm{UL}}(\Gamma, t_1 \oplus t_2) = \\
&\quad \text{let } (\tau_1, \theta_1) = \mathcal{W}_{\mathrm{UL}}(\Gamma, t_1) \\
&\qquad\;\; (\tau_2, \theta_2) = \mathcal{W}_{\mathrm{UL}}(\theta_1\ \Gamma, t_2) \\
&\qquad\;\; \theta_3 = \mathcal{U}_{\mathrm{UL}}(\theta_2\ \tau_1, \tau_{\oplus}^{1}) \\
&\qquad\;\; \theta_4 = \mathcal{U}_{\mathrm{UL}}(\theta_3\ \tau_2, \tau_{\oplus}^{2}) \\
&\quad \text{in } \;(\tau_{\oplus}, \quad \theta_4 \circ \theta_3 \circ \theta_2 \circ \theta_1)
\end{aligned}
$$

Control-flow analysis

Control-flow analysis (or closure analysis) determines:

For each function application, which functions may be applied.

Control-flow Analysis with Annotated Types

Annotated types

$\varphi \quad \in \quad \mathbf{Ann}$ annotations

$\varphi \quad ::= \quad \emptyset \mid \{\pi\} \mid \varphi_1 \cup \varphi_2$

Annotated types

φ	$\in$	$\mathbf{Ann}$	annotations
$\widehat{\tau}$	$\in$	$\widehat{\mathbf{Ty}}$	annotated types

$$\varphi ::= \emptyset \mid \{\pi\} \mid \varphi_1 \cup \varphi_2$$

$$\widehat{\tau} ::= \alpha \mid Nat \mid Bool \mid \widehat{\tau}_1 \xrightarrow{\varphi} \widehat{\tau}_2$$

Annotated types

φ	$\in$	$\mathbf{Ann}$	annotations
$\widehat{\tau}$	$\in$	$\widehat{\mathbf{Ty}}$	annotated types
$\widehat{\sigma}$	$\in$	$\widehat{\mathbf{TyScheme}}$	annotated type schemes

$$
\begin{array}{lcl}
\varphi & ::= & \emptyset \mid \{\pi\} \mid \varphi_1 \cup \varphi_2 \\
\widehat{\tau} & ::= & \alpha \mid \mathit{Nat} \mid \mathit{Bool} \mid \widehat{\tau}_1 \xrightarrow{\varphi} \widehat{\tau}_2 \\
\widehat{\sigma} & ::= & \widehat{\tau} \mid \forall \alpha .\, \widehat{\sigma}_1
\end{array}
$$

Annotated types

φ	$\in$	$\mathbf{Ann}$	annotations
$\widehat{\tau}$	$\in$	$\widehat{\mathbf{Ty}}$	annotated types
$\widehat{\sigma}$	$\in$	$\widehat{\mathbf{TyScheme}}$	annotated type schemes
$\widehat{\Gamma}$	$\in$	$\widehat{\mathbf{TyEnv}}$	annotated type environments

$$
\begin{array}{lcl}
\varphi & ::= & \emptyset \mid \{\pi\} \mid \varphi_1 \cup \varphi_2 \\
\widehat{\tau} & ::= & \alpha \mid \mathit{Nat} \mid \mathit{Bool} \mid \widehat{\tau}_1 \xrightarrow{\varphi} \widehat{\tau}_2 \\
\widehat{\sigma} & ::= & \widehat{\tau} \mid \forall \alpha.\, \widehat{\sigma}_1 \\
\widehat{\Gamma} & ::= & [\,] \mid \widehat{\Gamma}_1[x \mapsto \widehat{\sigma}]
\end{array}
$$

Annotated types

φ	$\in$	$\mathbf{Ann}$	annotations
$\widehat{\tau}$	$\in$	$\widehat{\mathbf{Ty}}$	annotated types
$\widehat{\sigma}$	$\in$	$\widehat{\mathbf{TyScheme}}$	annotated type schemes
$\widehat{\Gamma}$	$\in$	$\widehat{\mathbf{TyEnv}}$	annotated type environments

$$
\begin{array}{lcl}
\varphi & ::= & \emptyset \mid \{\pi\} \mid \varphi_1 \cup \varphi_2 \\
\widehat{\tau} & ::= & \alpha \mid \mathit{Nat} \mid \mathit{Bool} \mid \widehat{\tau}_1 \xrightarrow{\varphi} \widehat{\tau}_2 \\
\widehat{\sigma} & ::= & \widehat{\tau} \mid \forall \alpha . \, \widehat{\sigma}_1 \\
\widehat{\Gamma} & ::= & [\,] \mid \widehat{\Gamma}_1[x \mapsto \widehat{\sigma}]
\end{array}
$$

$\widehat{\Gamma} \vdash_{\mathrm{CFA}} t : \widehat{\sigma}$ control-flow analysis

Control-flow analysis: constants

$$\frac{}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} n : \mathit{Nat}} \; [\mathit{cfa\text{-}num}]$$

Control-flow analysis: constants

$$\frac{}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} n : \mathit{Nat}} \; [\mathit{cfa\text{-}num}]$$

$$\frac{}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} \mathtt{false} : \mathit{Bool}} \; [\mathit{cfa\text{-}false}]$$

$$\frac{}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} \mathtt{true} : \mathit{Bool}} \; [\mathit{cfa\text{-}true}]$$

Control-flow analysis: variables

$$\frac{\widehat{\Gamma}(x) = \widehat{\sigma}}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} x : \widehat{\sigma}} \; [\textit{cfa-var}]$$

Control-flow analysis: functions

$$\frac{\widehat{\Gamma}[x \mapsto \widehat{\tau}_1] \vdash_{\mathrm{CFA}} t_1 : \widehat{\tau}_2}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} \lambda_\pi x.\, t_1 : \widehat{\tau}_1 \xrightarrow{\{\pi\}} \widehat{\tau}_2} \; [\textit{cfa-lam}]$$

Control-flow analysis: functions

$$\frac{\widehat{\Gamma}[x \mapsto \widehat{\tau}_1] \vdash_{\mathrm{CFA}} t_1 : \widehat{\tau}_2}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} \lambda_\pi x.\, t_1 : \widehat{\tau}_1 \xrightarrow{\{\pi\}} \widehat{\tau}_2} \; [\textit{cfa-lam}]$$

$$\frac{\widehat{\Gamma}[f \mapsto (\widehat{\tau}_1 \xrightarrow{\{\pi\}} \widehat{\tau}_2)][x \mapsto \widehat{\tau}_1] \vdash_{\mathrm{CFA}} t_1 : \widehat{\tau}_2}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} \mu f.\, \lambda_\pi x.\, t_1 : \widehat{\tau}_1 \xrightarrow{\{\pi\}} \widehat{\tau}_2} \; [\textit{cfa-mu}]$$

Control-flow analysis: functions

$$\frac{\widehat{\Gamma}[x \mapsto \widehat{\tau}_1] \vdash_{\text{CFA}} t_1 : \widehat{\tau}_2}{\widehat{\Gamma} \vdash_{\text{CFA}} \lambda_\pi x.\, t_1 : \widehat{\tau}_1 \xrightarrow{\{\pi\}} \widehat{\tau}_2} \; [\textit{cfa-lam}]$$

$$\frac{\widehat{\Gamma}[f \mapsto (\widehat{\tau}_1 \xrightarrow{\{\pi\}} \widehat{\tau}_2)][x \mapsto \widehat{\tau}_1] \vdash_{\text{CFA}} t_1 : \widehat{\tau}_2}{\widehat{\Gamma} \vdash_{\text{CFA}} \mu f.\, \lambda_\pi x.\, t_1 : \widehat{\tau}_1 \xrightarrow{\{\pi\}} \widehat{\tau}_2} \; [\textit{cfa-mu}]$$

$$\frac{\widehat{\Gamma} \vdash_{\text{CFA}} t_1 : \widehat{\tau}_2 \xrightarrow{\varphi} \widehat{\tau} \quad \widehat{\Gamma} \vdash_{\text{CFA}} t_2 : \widehat{\tau}_2}{\widehat{\Gamma} \vdash_{\text{CFA}} t_1 \; t_2 : \widehat{\tau}} \; [\textit{cfa-app}]$$

- φ describes what may be applied!

Control-flow analysis: conditionals

$$\frac{\widehat{\Gamma} \vdash_{\mathrm{CFA}} t_1 : \mathit{Bool} \quad \widehat{\Gamma} \vdash_{\mathrm{CFA}} t_2 : \widehat{\tau} \quad \widehat{\Gamma} \vdash_{\mathrm{CFA}} t_3 : \widehat{\tau}}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} \mathbf{if}\ t_1\ \mathbf{then}\ t_2\ \mathbf{else}\ t_3 : \widehat{\tau}} \ [\textit{cfa-if}]$$

Control-flow analysis: local definitions

$$\frac{\widehat{\Gamma} \vdash_{\mathrm{CFA}} t_1 : \widehat{\sigma}_1 \quad \widehat{\Gamma}[x \mapsto \widehat{\sigma}_1] \vdash_{\mathrm{CFA}} t_2 : \widehat{\tau}}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} \mathbf{let}\ x = t_1\ \mathbf{in}\ t_2 : \widehat{\tau}}\ [\textit{cfa-let}]$$

Control-flow analysis: binary operators

$$\frac{\widehat{\Gamma} \vdash_{\mathrm{CFA}} t_1 : \tau_{\oplus}^{1} \quad \widehat{\Gamma} \vdash_{\mathrm{CFA}} t_2 : \tau_{\oplus}^{2}}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} t_1 \oplus t_2 : \tau_{\oplus}} \ [\textit{cfa-op}]$$

Control-flow analysis: example

$(\lambda_{\mathrm{F}} x.\, x)\ (\lambda_{\mathrm{G}} y.\, y)$

Control-flow analysis: example

$(\lambda_{F}\, x.\, x)\ (\lambda_{G}\, y.\, y)$

$$\frac{}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} (\lambda_{F}\, x.\, x)\ (\lambda_{G}\, y.\, y) : \forall \alpha.\, \alpha \xrightarrow{\{G\}} \alpha}$$

Control-flow analysis: example

$$(\lambda_{F} x.\, x)\ (\lambda_{G} y.\, y)$$

$$\dfrac{\dfrac{\dfrac{\vdots}{\widehat{\Gamma}[x \mapsto \widehat{\tau}_{G}] \vdash_{\text{CFA}} x : \widehat{\tau}_{G}}}{\widehat{\Gamma} \vdash_{\text{CFA}} \lambda_{F} x.\, x : \widehat{\tau}_{G} \xrightarrow{\{F\}} \widehat{\tau}_{G}} \qquad \dfrac{\dfrac{\vdots}{\widehat{\Gamma}[y \mapsto \alpha] \vdash_{\text{CFA}} y : \alpha}}{\widehat{\Gamma} \vdash_{\text{CFA}} \lambda_{G} y.\, y : \widehat{\tau}_{G}}}{\dfrac{\widehat{\Gamma} \vdash_{\text{CFA}} (\lambda_{F} x.\, x)\ (\lambda_{G} y.\, y) : \widehat{\tau}_{G}}{\widehat{\Gamma} \vdash_{\text{CFA}} (\lambda_{F} x.\, x)\ (\lambda_{G} y.\, y) : \forall \alpha.\, \alpha \xrightarrow{\{G\}} \alpha}}$$

$$\widehat{\tau}_{G} = \alpha \xrightarrow{\{G\}} \alpha$$

Higher-order functions

$$
\begin{array}{l}
\textbf{let } f = \lambda_{\color{blue}{\mathrm{F}}} x.\, x + 1 \textbf{ in} \\
\textbf{let } g = \lambda_{\color{blue}{\mathrm{G}}} y.\, y * 2 \textbf{ in} \\
\textbf{let } h = \lambda_{\color{blue}{\mathrm{H}}} z.\, z\ 3 \quad \textbf{in} \\
h\ g + h\ f
\end{array}
$$

Higher-order functions

$$
\begin{aligned}
&\textbf{let } f = \lambda_{\mathrm{F}} x.\ x + 1 \ \textbf{ in}\\
&\textbf{let } g = \lambda_{\mathrm{G}} y.\ y * 2 \ \textbf{ in}\\
&\textbf{let } h = \lambda_{\mathrm{H}} z.\ z\ 3 \quad \textbf{ in}\\
&h\ g + h\ f
\end{aligned}
$$

$$
\begin{aligned}
f &: \quad Nat \xrightarrow{\{\mathrm{F}\}} Nat\\
g &: \quad Nat \xrightarrow{\{\mathrm{G}\}} Nat
\end{aligned}
$$

Higher-order functions

$$\begin{array}{l}
\textbf{let } f = \lambda_{\mathrm{F}} x.\, x + 1 \textbf{ in} \\
\textbf{let } g = \lambda_{\mathrm{G}} y.\, y * 2 \textbf{ in} \\
\textbf{let } h = \lambda_{\mathrm{H}} z.\, z\ 3 \textbf{ in} \\
h\ g + h\ f
\end{array}$$

$$\begin{array}{lcl}
f & : & Nat \xrightarrow{\{\mathrm{F}\}} Nat \\
g & : & Nat \xrightarrow{\{\mathrm{G}\}} Nat \\
h & : & (Nat \xrightarrow{??} Nat) \xrightarrow{\{\mathrm{H}\}} Nat
\end{array}$$

Higher-order functions

$$\begin{array}{l}
\textbf{let } f = \lambda_{\text{F}} x.\ x + 1 \ \textbf{in} \\
\textbf{let } g = \lambda_{\text{G}} y.\ y * 2 \ \ \textbf{in} \\
\textbf{let } h = \lambda_{\text{H}} z.\ z\ 3 \quad \textbf{in} \\
h\ g + h\ f
\end{array}$$

$$\begin{array}{lcl}
f & : & Nat \xrightarrow{\{\text{F}\}} Nat \\
g & : & Nat \xrightarrow{\{\text{G}\}} Nat \\
h & : & (Nat \xrightarrow{??} Nat) \xrightarrow{\{\text{H}\}} Nat
\end{array}$$

Should we have $h : (Nat \xrightarrow{\{\text{F}\}} Nat) \xrightarrow{\{\text{H}\}} Nat$ or $h : (Nat \xrightarrow{\{\text{G}\}} Nat) \xrightarrow{\{\text{H}\}} Nat$?

Conditionals

$$
\begin{aligned}
\lambda_{\mathrm{H}} z.\ &\mathbf{if}\ && z \equiv 0 \\
&\mathbf{then}\ && \lambda_{\mathrm{F}} x.\ x + 1 \\
&\mathbf{else}\ && \lambda_{\mathrm{G}} y.\ y * 2
\end{aligned}
$$

Conditionals

$$\begin{aligned}\lambda_{\mathrm{H}} z.\ \mathbf{if}\ & \ z \equiv 0 \\ \mathbf{then}\ & \ \lambda_{\mathrm{F}} x.\ x + 1 \\ \mathbf{else}\ & \ \lambda_{\mathrm{G}} y.\ y * 2\end{aligned}$$

Should we have $Nat \xrightarrow{\{\mathrm{H}\}} (Nat \xrightarrow{\{\mathrm{F}\}} Nat)$ or $Nat \xrightarrow{\{\mathrm{H}\}} (Nat \xrightarrow{\{\mathrm{G}\}} Nat)$?

Subeffecting

$$\frac{\widehat{\Gamma}[x \mapsto \widehat{\tau}_1] \vdash_{\mathrm{CFA}} t_1 : \widehat{\tau}_2}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} \lambda_\pi x.\, t_1 : \widehat{\tau}_1 \xrightarrow{\{\pi\}\cup\varphi} \widehat{\tau}_2}\ [\mathit{cfa\text{-}lam}]$$

Subeffecting

$$\frac{\widehat{\Gamma}[x \mapsto \widehat{\tau}_1] \vdash_{\mathrm{CFA}} t_1 : \widehat{\tau}_2}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} \lambda_\pi x.\, t_1 : \widehat{\tau}_1 \xrightarrow{\{\pi\} \cup \varphi} \widehat{\tau}_2} \ [\textit{cfa-lam}]$$

$$\frac{\widehat{\Gamma}[f \mapsto (\widehat{\tau}_1 \xrightarrow{\{\pi\} \cup \varphi} \widehat{\tau}_2)][x \mapsto \widehat{\tau}_1] \vdash_{\mathrm{CFA}} t_1 : \widehat{\tau}_2}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} \mu f.\, \lambda_\pi x.\, t_1 : \widehat{\tau}_1 \xrightarrow{\{\pi\} \cup \varphi} \widehat{\tau}_2} \ [\textit{cfa-mu}]$$

Subeffecting: example

$$\begin{array}{l}
\mathbf{let}\ f = \lambda_{\mathrm{F}} x.\ x + 1\ \mathbf{in} \\
\mathbf{let}\ g = \lambda_{\mathrm{G}} y.\ y * 2\ \mathbf{in} \\
\mathbf{let}\ h = \lambda_{\mathrm{H}} z.\ z\ 3\ \ \mathbf{in} \\
h\ g + h\ f
\end{array}$$

$$\begin{array}{lll}
f & : & \mathit{Nat} \xrightarrow{\{\mathrm{F},\mathrm{G}\}} \mathit{Nat} \\
g & : & \mathit{Nat} \xrightarrow{\{\mathrm{F},\mathrm{G}\}} \mathit{Nat} \\
h & : & (\mathit{Nat} \xrightarrow{\{\mathrm{F},\mathrm{G}\}} \mathit{Nat}) \xrightarrow{\{\mathrm{H}\}} \mathit{Nat}
\end{array}$$

Subeffecting: example

$$\begin{aligned}
\lambda_{\mathrm{H}} z.\ \mathbf{if}\quad & z \equiv 0 \\
\mathbf{then}\ & \lambda_{\mathrm{F}} x.\ x + 1 \\
\mathbf{else}\ & \lambda_{\mathrm{G}} y.\ y * 2
\end{aligned}$$

$$\mathit{Nat} \xrightarrow{\{\mathrm{H}\}} (\mathit{Nat} \xrightarrow{\{\mathrm{F},\mathrm{G}\}} \mathit{Nat})$$

Inference algorithm: simple types

β	$\in$	$\mathbf{AnnVar}$	annotation variables
$\widehat{\tau}$	$\in$	$\widehat{\mathbf{SimpleTy}}$	simple types
$\widehat{\sigma}$	$\in$	$\widehat{\mathbf{SimpleTyScheme}}$	simple type schemes
$\widehat{\Gamma}$	$\in$	$\widehat{\mathbf{SimpleTyEnv}}$	simple type environments
$\widehat{\theta}$	$\in$	$\widehat{\mathbf{TySubst}}$	hybrid type substitution
C	$\in$	$\mathbf{Constr}$	constraint

$$
\begin{array}{lcl}
\widehat{\tau} & ::= & \alpha \mid Nat \mid Bool \mid \widehat{\tau}_1 \xrightarrow{\beta} \widehat{\tau}_2 \\
\widehat{\sigma} & ::= & \widehat{\tau} \mid \forall \alpha . \widehat{\sigma}_1 \\
\widehat{\Gamma} & ::= & [\,] \mid \widehat{\Gamma}_1[x \mapsto \widehat{\sigma}] \\
C & ::= & \emptyset \mid \{\beta \supseteq \varphi\} \mid C_1 \cup C_2
\end{array}
$$

Inference algorithm

$$
\begin{array}{lcl}
\mathit{generalise}_{\mathrm{CFA}} & : & \widehat{\mathbf{SimpleTyEnv}} \times \widehat{\mathbf{SimpleTy}} \rightarrow \\
 & & \widehat{\mathbf{SimpleTyScheme}} \\
\mathit{instantiate}_{\mathrm{CFA}} & : & \widehat{\mathbf{SimpleTyScheme}} \rightarrow \widehat{\mathbf{Ty}} \\
\mathcal{U}_{\mathrm{CFA}} & : & \widehat{\mathbf{SimpleTy}} \times \widehat{\mathbf{SimpleTy}} \rightarrow \\
 & & \widehat{\mathbf{TySubst}} \\
\mathcal{W}_{\mathrm{CFA}} & : & \widehat{\mathbf{SimpleTyEnv}} \times \mathbf{Tm} \rightarrow \\
 & & \widehat{\mathbf{SimpleTy}} \times \widehat{\mathbf{TySubst}} \times \mathbf{Constr}
\end{array}
$$

Inference algorithm: constants

$$\mathcal{W}_{\mathrm{CFA}}(\widehat{\Gamma}, n) = (Nat, \quad \mathit{id}, \quad \emptyset)$$

$$\mathcal{W}_{\mathrm{CFA}}(\widehat{\Gamma}, \texttt{false}) = (Bool, \quad \mathit{id}, \quad \emptyset)$$

$$\mathcal{W}_{\mathrm{CFA}}(\widehat{\Gamma}, \texttt{true}) = (Bool, \quad \mathit{id}, \quad \emptyset)$$

Inference algorithm: variables

$$\mathcal{W}_{\mathrm{CFA}}\ (\widehat{\Gamma}, x) = (\mathit{instantiate}_{\mathrm{CFA}}(\widehat{\Gamma}(x)), \quad \mathit{id}, \quad \emptyset)$$

Inference algorithm: functions

$$
\begin{aligned}
\mathcal{W}_{\text{CFA}}\ (\widehat{\Gamma}, \lambda_{\pi} x.\, t_1) = \text{let } & \alpha_1 \text{ be fresh} \\
& (\widehat{\tau}_2, \widehat{\theta}, C_1) = \mathcal{W}_{\text{CFA}}(\widehat{\Gamma}[x \mapsto \alpha_1], t_1) \\
& \beta \text{ be fresh} \\
\text{in } ((\widehat{\theta}\ \alpha_1) \xrightarrow{\beta} \widehat{\tau}_2, \quad & \widehat{\theta}, C_1 \cup \{\beta \supseteq \{\pi\}\})
\end{aligned}
$$

- Introduce fresh variables for annotations.
- Invariant: only variables as annotations in types.
- Put concrete information about the variable into C.
- Solve constraints later to obtain actual sets.
- Simplifies unification substantially.

Changes to unification

Only the case for function changes:

$$\ldots$$

$$\mathcal{U}_{\mathrm{UL}}\ (\tau_1 \xrightarrow{\beta_1} \tau_2, \tau_3 \xrightarrow{\beta_2} \tau_4) = \theta_2 \circ \theta_1 \circ \theta_0$$

where

$$\theta_0 = [\beta_1 \mapsto \beta_2]$$
$$\theta_1 = \mathcal{U}_{\mathrm{UL}}\ (\theta_0\ \tau_1, \theta_0\ \tau_3)$$
$$\theta_2 = \mathcal{U}_{\mathrm{UL}}\ (\theta_1\ (\theta_0\ \tau_2), \theta_1\ (\theta_0\ \tau_4))$$

$$\ldots$$

No need to recurse on annotations: just map one variable to the other.

Inference algorithm: recursive functions

$$
\begin{aligned}
&\mathcal{W}_{\mathrm{CFA}}\ (\widehat{\Gamma}, \mu f\,.\,\lambda_{\pi} x\,.\,t_1) = \\
&\quad \text{let } \alpha_1, \alpha_2, \beta \text{ be fresh} \\
&\qquad (\widehat{\tau}_2, \widehat{\theta}_1, C_1) = \mathcal{W}_{\mathrm{CFA}}(\widehat{\Gamma}[f \mapsto (\alpha_1 \xrightarrow{\beta} \alpha_2)][x \mapsto \alpha_1], t_1) \\
&\qquad \widehat{\theta}_2 = \mathcal{U}_{\mathrm{CFA}}(\widehat{\tau}_2, \widehat{\theta}_1\ \alpha_2) \\
&\quad \text{in } \ (\widehat{\theta}_2\ (\widehat{\theta}_1\ \alpha_1) \xrightarrow{\widehat{\theta}_2\ (\widehat{\theta}_1\ \beta)} \widehat{\theta}_2\ \widehat{\tau}_2, \quad \widehat{\theta}_2 \circ \widehat{\theta}_1, \\
&\qquad\quad (\widehat{\theta}_2\ C_1) \cup \{\widehat{\theta}_2\ (\widehat{\theta}_1\ \beta) \supseteq \{\pi\}\})
\end{aligned}
$$

Constraints: example

$$
\begin{array}{l}
\textbf{let } f \;= \lambda_{\mathrm{F}} x.\, x + 1 \;\textbf{ in} \\
\textbf{let } g = \lambda_{\mathrm{G}} y.\, y * 2 \;\;\textbf{ in} \\
\textbf{let } h = \lambda_{\mathrm{H}} z.\, z\ 3 \quad\;\textbf{ in} \\
h\ g + h\ f
\end{array}
$$

Constraints: example

$$\begin{array}{l} \textbf{let } f = \lambda_{\mathrm{F}} x.\ x + 1 \textbf{ in} \\ \textbf{let } g = \lambda_{\mathrm{G}} y.\ y * 2 \textbf{ in} \\ \textbf{let } h = \lambda_{\mathrm{H}} z.\ z\ 3 \textbf{ in} \\ h\ g + h\ f \end{array}$$

$$\begin{array}{lcl} f & : & Nat \xrightarrow{\beta_1} Nat \\ g & : & Nat \xrightarrow{\beta_2} Nat \\ h & : & (Nat \xrightarrow{\beta_3} Nat) \xrightarrow{\{\mathrm{H}\}} Nat \end{array}$$

Universiteit Utrecht
[Faculty of Science
Information and Computing Sciences]

Constraints: example

$$
\begin{array}{l}
\textbf{let } f = \lambda_{F} x.\ x + 1 \textbf{ in} \\
\textbf{let } g = \lambda_{G} y.\ y * 2 \textbf{ in} \\
\textbf{let } h = \lambda_{H} z.\ z\ 3 \textbf{ in} \\
h\ g + h\ f
\end{array}
$$

$$
\begin{array}{lcl}
f & : & Nat \xrightarrow{\beta_1} Nat \\
g & : & Nat \xrightarrow{\beta_2} Nat \\
h & : & (Nat \xrightarrow{\beta_3} Nat) \xrightarrow{\{H\}} Nat
\end{array}
$$

$$
\widehat{\theta}(\beta_1) = \beta_3
$$

$$
\widehat{\theta}(\beta_2) = \beta_3
$$

Constraints: example

$$
\begin{array}{l}
\textbf{let } f = \lambda_{\text{F}} x.\ x + 1 \textbf{ in} \\
\textbf{let } g = \lambda_{\text{G}} y.\ y * 2 \textbf{ in} \\
\textbf{let } h = \lambda_{\text{H}} z.\ z\ 3 \textbf{ in} \\
h\ g + h\ f
\end{array}
$$

$$
\begin{array}{lcl}
f & : & Nat \xrightarrow{\beta_1} Nat \\
g & : & Nat \xrightarrow{\beta_2} Nat \\
h & : & (Nat \xrightarrow{\beta_3} Nat) \xrightarrow{\{\text{H}\}} Nat
\end{array}
$$

$$\widehat{\theta}(\beta_1) = \beta_3$$
$$\widehat{\theta}(\beta_2) = \beta_3$$

$$C = \{\beta_1 \supseteq \{\text{F}\}, \beta_2 \supseteq \{\text{G}\}\}$$

Constraints: example

$$
\begin{array}{l}
\textbf{let } f = \lambda_{\mathrm{F}} x.\, x + 1 \textbf{ in} \\
\textbf{let } g = \lambda_{\mathrm{G}} y.\, y * 2 \textbf{ in} \\
\textbf{let } h = \lambda_{\mathrm{H}} z.\, z\ 3 \textbf{ in} \\
h\ g + h\ f
\end{array}
$$

$$
\begin{array}{lcl}
f & : & \mathit{Nat} \xrightarrow{\beta_1} \mathit{Nat} \\
g & : & \mathit{Nat} \xrightarrow{\beta_2} \mathit{Nat} \\
h & : & (\mathit{Nat} \xrightarrow{\beta_3} \mathit{Nat}) \xrightarrow{\{\mathrm{H}\}} \mathit{Nat}
\end{array}
$$

$$
\begin{array}{l}
\widehat{\theta}(\beta_1) = \beta_3 \\
\widehat{\theta}(\beta_2) = \beta_3
\end{array}
$$

$$
\begin{array}{ll}
C & = \{\beta_1 \supseteq \{\mathrm{F}\}, \beta_2 \supseteq \{\mathrm{G}\}\} \\
\widehat{\theta}\ C & = \{\beta_3 \supseteq \{\mathrm{F}\}, \beta_3 \supseteq \{\mathrm{G}\}\}
\end{array}
$$

Constraints: example

$$\begin{array}{l} \textbf{let } f = \lambda_{\mathrm{F}} x.\, x + 1 \textbf{ in} \\ \textbf{let } g = \lambda_{\mathrm{G}} y.\, y * 2 \textbf{ in} \\ \textbf{let } h = \lambda_{\mathrm{H}} z.\, z\; 3 \textbf{ in} \\ h\; g + h\; f \end{array}$$

$$\begin{array}{lll} f & : & Nat \xrightarrow{\beta_1} Nat \\ g & : & Nat \xrightarrow{\beta_2} Nat \\ h & : & (Nat \xrightarrow{\beta_3} Nat) \xrightarrow{\{\mathrm{H}\}} Nat \end{array}$$

$$\widehat{\theta}(\beta_1) = \beta_3$$
$$\widehat{\theta}(\beta_2) = \beta_3$$

$$C = \{\beta_1 \supseteq \{\mathrm{F}\}, \beta_2 \supseteq \{\mathrm{G}\}\}$$
$$\widehat{\theta}\; C = \{\beta_3 \supseteq \{\mathrm{F}\}, \beta_3 \supseteq \{\mathrm{G}\}\}$$

Least solution: $\beta_3 = \{\mathrm{F}, \mathrm{G}\}$.

Poisoning

Naive use of subeffecting is fatal for the precision of your analysis:

$$
\begin{array}{ll}
\textbf{let}\ f\ = \lambda_{\mathrm{F}} x.\ x + 1 & \textbf{in} \\
\textbf{let}\ g\ = \lambda_{\mathrm{G}} y.\ y * 2 & \textbf{in} \\
\textbf{let}\ h\ = \lambda_{\mathrm{H}} z.\ \textbf{if}\ z \equiv 0\ \textbf{then}\ f\ \textbf{else}\ g & \textbf{in} \\
f &
\end{array}
$$

$$
\mathit{Nat} \xrightarrow{\{\mathrm{F},\mathrm{G}\}} \mathit{Nat}
$$

Universiteit Utrecht

[Faculty of **Science**
Information and Computing Sciences]

Separate rule for subeffecting

$$\frac{\widehat{\Gamma} \vdash_{\mathrm{CFA}} t : \widehat{\tau}_1 \xrightarrow{\varphi} \widehat{\tau}_2}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} t : \widehat{\tau}_1 \xrightarrow{\varphi \cup \varphi'} \widehat{\tau}_1} \; [\mathit{cfa\text{-}sub}]$$

Separate rule for subeffecting

$$\frac{\widehat{\Gamma} \vdash_{\mathrm{CFA}} t : \widehat{\tau}_1 \xrightarrow{\varphi} \widehat{\tau}_2}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} t : \widehat{\tau}_1 \xrightarrow{\varphi \cup \varphi'} \widehat{\tau}_1} \; [\textit{cfa-sub}]$$

We can remove the subeffecting from the lambda rule:

$$\frac{\widehat{\Gamma}[x \mapsto \widehat{\tau}_1] \vdash_{\mathrm{CFA}} t_1 : \widehat{\tau}_2}{\widehat{\Gamma} \vdash_{\mathrm{CFA}} \lambda_\pi x.\, t_1 : \widehat{\tau}_1 \xrightarrow{\{\pi\}} \widehat{\tau}_2} \; [\textit{cfa-lam}]$$

Separate compilation?

$$
\begin{aligned}
&\textbf{let } f = \lambda_{\mathrm{F}} x.\ x + 1 && \textbf{in} \\
&\textbf{let } g = \lambda_{\mathrm{G}} y.\ y * 2 && \textbf{in} \\
&\textbf{let } h = \lambda_{\mathrm{H}} z.\ z\ 3 && \textbf{in} \\
&h\ g + h\ f
\end{aligned}
$$

$$
\begin{aligned}
f &: \quad \mathit{Nat} \xrightarrow{\{\mathrm{F}\}} \mathit{Nat} \\
g &: \quad \mathit{Nat} \xrightarrow{\{\mathrm{G}\}} \mathit{Nat} \\
h &: \quad (\mathit{Nat} \xrightarrow{\{\mathrm{F},\mathrm{G}\}} \mathit{Nat}) \xrightarrow{\{\mathrm{H}\}} \mathit{Nat}
\end{aligned}
$$

Separate compilation?

$$\begin{aligned}
&\mathbf{let}\ f = \lambda_{\mathrm{F}} x.\ x + 1\ \mathbf{in} \\
&\mathbf{let}\ g = \lambda_{\mathrm{G}} y.\ y * 2\ \ \mathbf{in} \\
&\mathbf{let}\ h = \lambda_{\mathrm{H}} z.\ z\ 3\ \ \ \ \mathbf{in} \\
&h\ g + h\ f
\end{aligned}$$

$$\begin{aligned}
f &: \quad \mathit{Nat} \xrightarrow{\{\mathrm{F}\}} \mathit{Nat} \\
g &: \quad \mathit{Nat} \xrightarrow{\{\mathrm{G}\}} \mathit{Nat} \\
h &: \quad (\mathit{Nat} \xrightarrow{\{\mathrm{F},\mathrm{G}\}} \mathit{Nat}) \xrightarrow{\{\mathrm{H}\}} \mathit{Nat}
\end{aligned}$$

☞ We need to analyse the whole program to accurately determine the domain of h.

Subeffecting and subtyping

- We have now seen subeffecting at work.
- The main ideas of all of these are:
  - compute types and annotations independent of context,
  - allow to weaken the outcomes whenever convenient.
- Weakening provides a form of context-sensitiveness.
- In (shape conformant) subtyping we may also weaken annotations deeper in the type.