CIS 500

Software Foundations

Fall 2002

30 October

Administrivia

- Prof. Pierce out of town Nov. 5 - 14
  - No office hours Nov 5, 7, 12, or 14
  - Next Wednesday: guest lecturer (on Chapter 16)
  - Following Monday: review session (led by Anne and Jim)
  - 3PM recitation cancelled on Nov 11 - go to Max's in Towne 307 instead
  - Following Wednesday: Midterm II
- There will be class on the Wednesday before Thanksgiving (Nov. 27)

Review

Subtyping

Intuitions: `S <: T` means...

- "An element of `S` may safely be used wherever an element of `T` is expected." (Official.)
- `S` is "better than" `T`.
- `S` is a subset of `T`.
- `S` is more informative / richer than `T`.

Subtype relation

$$\texttt{S} <: \texttt{S} \qquad \text{(S-Refl)}$$

$$\frac{\texttt{S} <: \texttt{U} \qquad \texttt{U} <: \texttt{T}}{\texttt{S} <: \texttt{T}} \qquad \text{(S-Trans)}$$

$$\{\texttt{l}_i\texttt{:T}_i\ ^{i\in 1..n+k}\} <: \{\texttt{l}_i\texttt{:T}_i\ ^{i\in 1..n}\} \qquad \text{(S-RcdWidth)}$$

$$\frac{\text{for each } i \qquad \texttt{S}_i <: \texttt{T}_i}{\{\texttt{l}_i\texttt{:S}_i\ ^{i\in 1..n}\} <: \{\texttt{l}_i\texttt{:T}_i\ ^{i\in 1..n}\}} \qquad \text{(S-RcdDepth)}$$

$$\frac{\{\texttt{k}_j\texttt{:S}_j\ ^{j\in 1..n}\} \text{ is a permutation of } \{\texttt{l}_i\texttt{:T}_i\ ^{i\in 1..n}\}}{\{\texttt{k}_j\texttt{:S}_j\ ^{j\in 1..n}\} <: \{\texttt{l}_i\texttt{:T}_i\ ^{i\in 1..n}\}} \qquad \text{(S-RcdPerm)}$$

$$\frac{T_1 \mathrel{<:} S_1 \qquad S_2 \mathrel{<:} T_2}{S_1 \rightarrow S_2 \mathrel{<:} T_1 \rightarrow T_2} \qquad \text{(S-Arrow)}$$

$$S \mathrel{<:} \mathtt{Top} \qquad \text{(S-Top)}$$

Subsumption Rule

$$\frac{\Gamma \vdash \mathtt{t} : \mathtt{S} \qquad \mathtt{S} <: \mathtt{T}}{\Gamma \vdash \mathtt{t} : \mathtt{T}} \qquad \text{(T-SUB)}$$

Properties

Safety

Statements of progress and preservation theorems are unchanged.

Proofs become a bit more involved, because the typing relation is *no* longer syntax directed.

Preservation

Theorem: If $\Gamma \vdash \mathtt{t} : \mathtt{T}$ and $\mathtt{t} \longrightarrow \mathtt{t}'$, then $\Gamma \vdash \mathtt{t}' : \mathtt{T}$.

Proof: By induction on typing derivations.

(Which cases are hard?)

Subsumption case

Case T-Sub: $\quad \texttt{t : S} \qquad \texttt{S <: T}$

By the induction hypothesis, $\Gamma \vdash \texttt{t}' : \texttt{S}$. By T-Sub, $\Gamma \vdash \texttt{t} : \texttt{T}$.

Subsumption case

Case T-Sub: $\quad \mathtt{t : S} \qquad \mathtt{S <: T}$

By the induction hypothesis, $\Gamma \vdash \mathtt{t}' : \mathtt{S}$. By T-Sub, $\Gamma \vdash \mathtt{t} : \mathtt{T}$.

Not hard!

Application case

Case T-App:

$\mathtt{t} = \mathtt{t}_1\ \mathtt{t}_2 \qquad \Gamma \vdash \mathtt{t}_1 : \mathtt{T}_{11}{\to}\mathtt{T}_{12} \qquad \Gamma \vdash \mathtt{t}_2 : \mathtt{T}_{11} \qquad \mathtt{T} = \mathtt{T}_{12}$

From the evaluation rules (i.e., strictly speaking, from the inversion lemma for evaluation), there are three rules by which $\mathtt{t} \longrightarrow \mathtt{t}'$ can be derived: E-App1, E-App2, and E-AppAbs. Proceed by cases.

$$\frac{\Gamma \vdash \mathtt{t}_1 : \mathtt{T}_{11}{\to}\mathtt{T}_{12} \qquad \Gamma \vdash \mathtt{t}_2 : \mathtt{T}_{11}}{\Gamma \vdash \mathtt{t}_1\ \mathtt{t}_2 : \mathtt{T}_{12}} \qquad \text{(T-App)}$$

Application case

Case T-App:

$t = t_1\ t_2 \qquad \Gamma \vdash t_1 : T_{11}{\to}T_{12} \qquad \Gamma \vdash t_2 : T_{11} \qquad T = T_{12}$

From the evaluation rules (i.e., strictly speaking, from the inversion lemma for evaluation), there are three rules by which $t \longrightarrow t'$ can be derived: E-App1, E-App2, and E-AppAbs. Proceed by cases.

Subcase E-App1: $\quad t_1 \longrightarrow t_1' \qquad t' = t_1'\ t_2$

The result follows from the induction hypothesis and T-App.

$$\frac{\Gamma \vdash t_1 : T_{11}{\to}T_{12} \qquad \Gamma \vdash t_2 : T_{11}}{\Gamma \vdash t_1\ t_2 : T_{12}} \qquad \text{(T-App)}$$

$$\frac{t_1 \longrightarrow t_1'}{t_1\ t_2 \longrightarrow t_1'\ t_2} \qquad \text{(E-App1)}$$

Case T-App (continued):

$$t = t_1\ t_2 \qquad \Gamma \vdash t_1 : T_{11} \to T_{12} \qquad \Gamma \vdash t_2 : T_{11} \qquad T = T_{12}$$

Subcase E-App2: $t_1 = v_1 \qquad t_2 \longrightarrow t_2' \qquad t' = v_1\ t_2'$

Similar.

$$\frac{\Gamma \vdash t_1 : T_{11} \to T_{12} \qquad \Gamma \vdash t_2 : T_{11}}{\Gamma \vdash t_1\ t_2 : T_{12}} \qquad \text{(T-App)}$$

$$\frac{t_2 \longrightarrow t_2'}{v_1\ t_2 \longrightarrow v_1\ t_2'} \qquad \text{(E-App2)}$$

Case T-App (continued):

$t = t_1\ t_2 \qquad \Gamma \vdash t_1 : T_{11} \to T_{12} \qquad \Gamma \vdash t_2 : T_{11} \qquad T = T_{12}$

Subcase E-AppAbs: $t_1 = \lambda x{:}S_{11}.\ t_{12} \qquad t_2 = v_2 \qquad t' = [x \mapsto v_2]t_{12}$

By the inversion lemma for the typing relation...

$$\frac{\Gamma \vdash t_1 : T_{11} \to T_{12} \qquad \Gamma \vdash t_2 : T_{11}}{\Gamma \vdash t_1\ t_2 : T_{12}} \qquad \text{(T-App)}$$

$$(\lambda x{:}T_{11}.t_{12})\ v_2 \longrightarrow [x \mapsto v_2]t_{12} \qquad \text{(E-AppAbs)}$$

Case T-App (continued):

$t = t_1\ t_2 \qquad \Gamma \vdash t_1 : T_{11} \rightarrow T_{12} \qquad \Gamma \vdash t_2 : T_{11} \qquad T = T_{12}$

Subcase E-AppAbs: $t_1 = \lambda x{:}S_{11}.\ t_{12} \qquad t_2 = v_2 \qquad t' = [x \mapsto v_2]t_{12}$

By the inversion lemma for the typing relation... T_{11} <: S_{11} and $\Gamma, x{:}S_{11} \vdash t_{12} : T_{12}$.

$$\frac{\Gamma \vdash t_1 : T_{11} \rightarrow T_{12} \qquad \Gamma \vdash t_2 : T_{11}}{\Gamma \vdash t_1\ t_2 : T_{12}} \qquad \text{(T-App)}$$

$$(\lambda x{:}T_{11}.t_{12})\ v_2 \longrightarrow [x \mapsto v_2]t_{12} \qquad \text{(E-AppAbs)}$$

Case T-App (continued):

$t = t_1\ t_2 \qquad \Gamma \vdash t_1 : T_{11} \to T_{12} \qquad \Gamma \vdash t_2 : T_{11} \qquad T = T_{12}$

Subcase E-AppAbs: $\quad t_1 = \lambda x{:}S_{11}.\ t_{12} \qquad t_2 = v_2 \qquad t' = [x \mapsto v_2]t_{12}$

By the inversion lemma for the typing relation... $T_{11} <: S_{11}$ and $\Gamma, x{:}S_{11} \vdash t_{12} : T_{12}$.

By T-Sub, $\Gamma \vdash t_2 : S_{11}$.

$$\frac{\Gamma \vdash t_1 : T_{11} \to T_{12} \qquad \Gamma \vdash t_2 : T_{11}}{\Gamma \vdash t_1\ t_2 : T_{12}} \qquad \text{(T-App)}$$

$$(\lambda x{:}T_{11}.t_{12})\ v_2 \longrightarrow [x \mapsto v_2]t_{12} \qquad \text{(E-AppAbs)}$$

Case T-App (continued):

$t = t_1\ t_2 \qquad \Gamma \vdash t_1 : T_{11} \to T_{12} \qquad \Gamma \vdash t_2 : T_{11} \qquad T = T_{12}$

Subcase E-AppAbs: $\quad t_1 = \lambda x{:}S_{11}.\ t_{12} \qquad t_2 = v_2 \qquad t' = [x \mapsto v_2]t_{12}$

By the inversion lemma for the typing relation... $T_{11} <: S_{11}$ and $\Gamma, x{:}S_{11} \vdash t_{12} : T_{12}$.

By T-Sub, $\Gamma \vdash t_2 : S_{11}$.

By the substitution lemma, $\Gamma \vdash t' : T_{12}$, and we are done.

$$\frac{\Gamma \vdash t_1 : T_{11} \to T_{12} \qquad \Gamma \vdash t_2 : T_{11}}{\Gamma \vdash t_1\ t_2 : T_{12}} \qquad \text{(T-App)}$$

$$(\lambda x{:}T_{11}.t_{12})\ v_2 \longrightarrow [x \mapsto v_2]t_{12} \qquad \text{(E-AppAbs)}$$

Inversion Lemma

Lemma: If $\Gamma \vdash \lambda x{:}S_1.s_2 : T_1 \to T_2$, then $T_1 <: S_1$ and $\Gamma, x{:}S_1 \vdash s_2 : T_2$.

Proof: Induction on typing derivations.

Inversion Lemma

Lemma: If $\Gamma \vdash \lambda x{:}S_1.s_2 : T_1{\to}T_2$, then $T_1 <: S_1$ and $\Gamma, x{:}S_1 \vdash s_2 : T_2$.

Proof: Induction on typing derivations.

Case T-Sub: $\quad \lambda x{:}S_1.s_2 : U \qquad U <: T_1{\to}T_2$

We want to say "By the induction hypothesis...", but the IH does not apply (we do not know that S is an arrow type).

Inversion Lemma

Lemma: If $\Gamma \vdash \lambda x{:}S_1.s_2 : T_1 \to T_2$, then $T_1 <: S_1$ and $\Gamma, x{:}S_1 \vdash s_2 : T_2$.

Proof: Induction on typing derivations.

Case T-Sub: $\lambda x{:}S_1.s_2 : U \qquad U <: T_1 \to T_2$

We want to say "By the induction hypothesis...", but the IH does not apply (we do not know that S is an arrow type). Need another lemma...

> **Lemma:** If $U <: T_1 \to T_2$, then U has the form $U_1 \to U_2$, with $T_1 <: U_1$ and $U_2 <: T_2$. (Proof: by induction on subtyping derivations.)

Inversion Lemma

Lemma: If $\Gamma \vdash \lambda x{:}S_1.s_2 : T_1 \to T_2$, then $T_1 <: S_1$ and $\Gamma, x{:}S_1 \vdash s_2 : T_2$.

Proof: Induction on typing derivations.

Case T-Sub: $\quad \lambda x{:}S_1.s_2 : U \qquad U <: T_1 \to T_2$

We want to say "By the induction hypothesis...", but the IH does not apply (we do not know that S is an arrow type). Need another lemma...

> **Lemma:** If $U <: T_1 \to T_2$, then U has the form $U_1 \to U_2$, with $T_1 <: U_1$ and $U_2 <: T_2$. (Proof: by induction on subtyping derivations.)

By this lemma, we know $U = U_1 \to U_2$, with $T_1 <: U_1$ and $U_2 <: T_2$.

Inversion Lemma

Lemma: If $\Gamma \vdash \lambda x{:}S_1.s_2 : T_1 \to T_2$, then $T_1 <: S_1$ and $\Gamma, x{:}S_1 \vdash s_2 : T_2$.

Proof: Induction on typing derivations.

Case T-Sub: $\lambda x{:}S_1.s_2 : U \qquad U <: T_1 \to T_2$

We want to say "By the induction hypothesis...", but the IH does not apply (we do not know that S is an arrow type). Need another lemma...

> **Lemma:** If $U <: T_1 \to T_2$, then U has the form $U_1 \to U_2$, with $T_1 <: U_1$ and $U_2 <: T_2$. (Proof: by induction on subtyping derivations.)

By this lemma, we know $U = U_1 \to U_2$, with $T_1 <: U_1$ and $U_2 <: T_2$.

The IH now applies, yielding $U_1 <: S_1$ and $\Gamma, x{:}S_1 \vdash s_2 : U_2$.

Inversion Lemma

Lemma: If $\Gamma \vdash \lambda x{:}S_1.s_2 : T_1 \to T_2$, then $T_1 <: S_1$ and $\Gamma, x{:}S_1 \vdash s_2 : T_2$.

Proof: Induction on typing derivations.

Case T-Sub: $\quad \lambda x{:}S_1.s_2 : U \qquad U <: T_1 \to T_2$

We want to say "By the induction hypothesis...", but the IH does not apply (we do not know that S is an arrow type). Need another lemma...

> **Lemma:** If $U <: T_1 \to T_2$, then U has the form $U_1 \to U_2$, with $T_1 <: U_1$ and $U_2 <: T_2$. (Proof: by induction on subtyping derivations.)

By this lemma, we know $U = U_1 \to U_2$, with $T_1 <: U_1$ and $U_2 <: T_2$.

The IH now applies, yielding $U_1 <: S_1$ and $\Gamma, x{:}S_1 \vdash s_2 : U_2$.

From $U_1 <: S_1$ and $T_1 <: U_1$, rule S-Trans gives $T_1 <: S_1$.

Inversion Lemma

Lemma: If $\Gamma \vdash \lambda x{:}S_1.s_2 : T_1 \to T_2$, then $T_1 <: S_1$ and $\Gamma, x{:}S_1 \vdash s_2 : T_2$.

Proof: Induction on typing derivations.

Case T-Sub: $\lambda x{:}S_1.s_2 : U \qquad U <: T_1 \to T_2$

We want to say "By the induction hypothesis...", but the IH does not apply (we do not know that S is an arrow type). Need another lemma...

> **Lemma:** If $U <: T_1 \to T_2$, then U has the form $U_1 \to U_2$, with $T_1 <: U_1$ and $U_2 <: T_2$. (Proof: by induction on subtyping derivations.)

By this lemma, we know $U = U_1 \to U_2$, with $T_1 <: U_1$ and $U_2 <: T_2$.

The IH now applies, yielding $U_1 <: S_1$ and $\Gamma, x{:}S_1 \vdash s_2 : U_2$.

From $U_1 <: S_1$ and $T_1 <: U_1$, rule S-Trans gives $T_1 <: S_1$.

From $\Gamma, x{:}S_1 \vdash s_2 : U_2$ and $U_2 <: T_2$, rule T-Sub gives $\Gamma, x{:}S_1 \vdash s_2 : T_2$, and we are done.

Subtyping and Other Features

Ascription and Casting

Ordinary ascription:

$$\frac{\Gamma \vdash \mathtt{t}_1 : \mathtt{T}}{\Gamma \vdash \mathtt{t}_1 \ \mathtt{as}\ \mathtt{T} : \mathtt{T}} \qquad \text{(T-Ascribe)}$$

$$\mathtt{v}_1 \ \mathtt{as}\ \mathtt{T} \longrightarrow \mathtt{v}_1 \qquad \text{(E-Ascribe)}$$

Ascription and Casting

Ordinary ascription:

$$\frac{\Gamma \vdash \mathtt{t_1} : \mathtt{T}}{\Gamma \vdash \mathtt{t_1}\ \mathtt{as}\ \mathtt{T} : \mathtt{T}} \qquad \text{(T-Ascribe)}$$

$$\mathtt{v_1}\ \mathtt{as}\ \mathtt{T} \longrightarrow \mathtt{v_1} \qquad \text{(E-Ascribe)}$$

Casting (cf. Java):

$$\frac{\Gamma \vdash \mathtt{t_1} : \mathtt{S}}{\Gamma \vdash \mathtt{t_1}\ \mathtt{as}\ \mathtt{T} : \mathtt{T}} \qquad \text{(T-Cast)}$$

$$\frac{\vdash \mathtt{v_1} : \mathtt{T}}{\mathtt{v_1}\ \mathtt{as}\ \mathtt{T} \longrightarrow \mathtt{v_1}} \qquad \text{(E-Cast)}$$

Subtyping and Variants

$$\langle l_i : T_i \;^{i \in 1..n} \rangle \quad <: \quad \langle l_i : T_i \;^{i \in 1..n+k} \rangle \qquad \text{(S-VariantWidth)}$$

$$\frac{\text{for each } i \qquad S_i <: T_i}{\langle l_i : S_i \;^{i \in 1..n} \rangle \quad <: \quad \langle l_i : T_i \;^{i \in 1..n} \rangle} \qquad \text{(S-VariantDepth)}$$

$$\frac{\langle k_j : S_j \;^{j \in 1..n} \rangle \text{ is a permutation of } \langle l_i : T_i \;^{i \in 1..n} \rangle}{\langle k_j : S_j \;^{j \in 1..n} \rangle \quad <: \quad \langle l_i : T_i \;^{i \in 1..n} \rangle} \qquad \text{(S-VariantPerm)}$$

$$\frac{\Gamma \vdash t_1 : T_1}{\Gamma \vdash \langle l_1 = t_1 \rangle : \langle l_1 : T_1 \rangle} \qquad \text{(T-Variant)}$$

Subtyping and Lists

$$\frac{\texttt{S}_1 \mathrel{<:} \texttt{T}_1}{\texttt{List }\texttt{S}_1 \mathrel{<:} \texttt{List }\texttt{T}_1} \qquad \text{(S-List)}$$

I.e., `List` is a covariant type constructor.

Subtyping and References

$$\frac{\texttt{S}_1 \ \texttt{<:}\ \texttt{T}_1 \qquad \texttt{T}_1 \ \texttt{<:}\ \texttt{S}_1}{\texttt{Ref}\ \texttt{S}_1 \ \texttt{<:}\ \texttt{Ref}\ \texttt{T}_1} \qquad \text{(S-Ref)}$$

I.e., `Ref` is *not* a covariant (nor a contravariant) type constructor.

Subtyping and Arrays

Similarly...

$$\frac{\mathtt{S_1 <: T_1} \qquad \mathtt{T_1 <: S_1}}{\mathtt{Array\ S_1 <: Array\ T_1}} \qquad \text{(S-Array)}$$

Subtyping and Arrays

Similarly...

$$\frac{\mathtt{S_1} <: \mathtt{T_1} \qquad \mathtt{T_1} <: \mathtt{S_1}}{\mathtt{Array}\ \mathtt{S_1} <: \mathtt{Array}\ \mathtt{T_1}} \qquad \text{(S-Array)}$$

$$\frac{\mathtt{S_1} <: \mathtt{T_1}}{\mathtt{Array}\ \mathtt{S_1} <: \mathtt{Array}\ \mathtt{T_1}} \qquad \text{(S-ArrayJava)}$$

This is regarded (even by the Java designers) as a mistake in the design.

References again

Observation: a value of type `Ref T` can be used in two different ways: as a **source** for values of type `T` and as a **sink** for values of type `T`.

References again

Observation: a value of type `Ref T` can be used in two different ways: as a *source* for values of type `T` and as a *sink* for values of type `T`.

Idea: Split `Ref T` into three parts:

- `Source T`: reference cell with “read cabability”
- `Sink T`: reference cell with “write cabability”
- `Ref T`: cell with both capabilities

Modified Typing Rules

$$\frac{\Gamma \mid \Sigma \vdash \mathtt{t}_1 : \mathtt{Source}\ \mathtt{T}_{11}}{\Gamma \mid \Sigma \vdash \mathtt{!t}_1 : \mathtt{T}_{11}} \qquad \text{(T-Deref)}$$

$$\frac{\Gamma \mid \Sigma \vdash \mathtt{t}_1 : \mathtt{Sink}\ \mathtt{T}_{11} \qquad \Gamma \mid \Sigma \vdash \mathtt{t}_2 : \mathtt{T}_{11}}{\Gamma \mid \Sigma \vdash \mathtt{t}_1\mathtt{:=t}_2 : \mathtt{Unit}} \qquad \text{(T-Assign)}$$

Subtyping rules

$$\frac{S_1 \mathrel{<:} T_1}{\texttt{Source}\ S_1 \mathrel{<:} \texttt{Source}\ T_1} \qquad \text{(S-Source)}$$

$$\frac{T_1 \mathrel{<:} S_1}{\texttt{Sink}\ S_1 \mathrel{<:} \texttt{Sink}\ T_1} \qquad \text{(S-Sink)}$$

$$\texttt{Ref}\ T_1 \mathrel{<:} \texttt{Source}\ T_1 \qquad \text{(S-RefSource)}$$

$$\texttt{Ref}\ T_1 \mathrel{<:} \texttt{Sink}\ T_1 \qquad \text{(S-RefSink)}$$

Capabilities

Other kinds of capabilities (e.g., send and receive capabilities on communication channels, encrypt/decrypt capabilities of cryptographic keys, ...) can be treated similarly.

Coercion semantics

[skip]

Intersection Types

The inhabitants of $T_1 \wedge T_2$ are terms belonging to both S and T—i.e., $T_1 \wedge T_2$ is an order-theoretic meet (greatest lower bound) of T_1 and T_2.

$$T_1 \wedge T_2 <: T_1 \qquad \text{(S-Inter1)}$$

$$T_1 \wedge T_2 <: T_2 \qquad \text{(S-Inter2)}$$

$$\frac{S <: T_1 \qquad S <: T_2}{S <: T_1 \wedge T_2} \qquad \text{(S-Inter3)}$$

$$S \to T_1 \wedge S \to T_2 <: S \to (T_1 \wedge T_2) \qquad \text{(S-Inter4)}$$

Intersection Types

Intersection types permit a very flexible form of finitary overloading.

$$+ : (\mathtt{Nat}{\to}\mathtt{Nat}{\to}\mathtt{Nat}) \wedge (\mathtt{Float}{\to}\mathtt{Float}{\to}\mathtt{Float})$$

This form of overloading is extremely powerful.

> Every strongly normalizing untyped lambda-term can be typed in the simply typed lambda-calculus with intersection types.
>
> $\longrightarrow$ type reconstruction problem is undecidable

Intersection types have not been used much in language designs (too powerful!), but are being intensively investigated as type systems for intermediate languages in highly optimizing compilers (cf. Church project).

Union types

Union types are also useful.

$T_1 \vee T_2$ is an untagged (non-disjoint) union of T_1 and T_2

$\longrightarrow$ no `case` construct. The only operations we can safely perform on elements of $T_1 \backslash/ T2$ are ones that make sense for both T_1 and T_2.

N.b.: untagged `union` types in C are a source of type safety violations precisely because they ignores this restriction, allowing any operation on an element of $T_1 \vee T_2$ that makes sense for either T_1 or T_2.

Union types are being used recently in type systems for XML processing languages (cf. XDuce, Xtatic).