

CIS 500
Software Foundations
Fall 2006

October 16

Any Questions?

Plan

“We have the technology...”

- ▶ In this lecture and the next, we're going to cover some simple extensions of the typed-lambda calculus (TAPL Chapter 11).
 1. Products, records
 2. Sums, variants
 3. Recursion
- ▶ We're skipping Chapters 10 and 12.

Erasure and Typability

Eraseure

We can transform terms in λ_{\rightarrow} to terms of the untyped lambda-calculus simply by erasing type annotations on lambda-abstractions.

$$\begin{aligned} \text{erase}(x) &= x \\ \text{erase}(\lambda x:T_1. t_2) &= \lambda x. \text{erase}(t_2) \\ \text{erase}(t_1 t_2) &= \text{erase}(t_1) \text{erase}(t_2) \end{aligned}$$

Typability

Conversely, an untyped λ -term m is said to be *typable* if there is some term t in the simply typed lambda-calculus, some type T , and some context Γ such that $\text{erase}(t) = m$ and $\Gamma \vdash t : T$.

This process is called *type reconstruction* or *type inference*.

Typability

Conversely, an untyped λ -term m is said to be *typable* if there is some term t in the simply typed lambda-calculus, some type T , and some context Γ such that $\text{erase}(t) = m$ and $\Gamma \vdash t : T$.

This process is called *type reconstruction* or *type inference*.

Example: Is the term

$\lambda x. x x$

typable?

The Curry-Howard Correspondence

Intro vs. elim forms

An *introduction form* for a given type gives us a way of *constructing* elements of this type.

An *elimination form* for a type gives us a way of *using* elements of this type.

The Curry-Howard Correspondence

In *constructive logics*, a proof of P must provide *evidence* for P .

- ▶ “law of the excluded middle” — $P \vee \neg P$ — not recognized.

A proof of $P \wedge Q$ is a *pair* of evidence for P and evidence for Q .

A proof of $P \supset Q$ is a *procedure* for transforming evidence for P into evidence for Q .

Propositions as Types

LOGIC

propositions

proposition $P \supset Q$

proposition $P \wedge Q$

proof of proposition P

proposition P is provable

PROGRAMMING LANGUAGES

types

type $P \rightarrow Q$

type $P \times Q$

term t of type P

type P is inhabited (by some term)

evaluation

Propositions as Types

LOGIC

propositions

proposition $P \supset Q$

proposition $P \wedge Q$

proof of proposition P

proposition P is provable

proof simplification

(a.k.a. “cut elimination”)

PROGRAMMING LANGUAGES

types

type $P \rightarrow Q$

type $P \times Q$

term t of type P

type P is inhabited (by some term)

On to real programming
languages...

Base types

Up to now, we've formulated “base types” (e.g. `Nat`) by adding them to the syntax of types, extending the syntax of terms with associated constants (`zero`) and operators (`succ`, etc.) and adding appropriate typing and evaluation rules. We can do this for as many base types as we like.

For more theoretical discussions (as opposed to programming) we can often ignore the term-level inhabitants of base types, and just treat these types as uninterpreted constants.

E.g., suppose `B` and `C` are some base types. Then we can ask (without knowing anything more about `B` or `C`) whether there are any types `S` and `T` such that the term

$$(\lambda f:S. \lambda g:T. f\ g) (\lambda x:B. x)$$

is well typed.

The Unit type

$t ::= \dots$
 unit

terms
constant unit

$v ::= \dots$
 unit

values
constant unit

$T ::= \dots$
 Unit

types
unit type

New typing rules

$\boxed{\Gamma \vdash t : T}$

$\Gamma \vdash \text{unit} : \text{Unit}$

(T-UNIT)

Sequencing

$t ::= \dots$
 $t_1; t_2$

terms

Sequencing

$t ::= \dots$
 $t_1; t_2$

terms

$$\frac{t_1 \longrightarrow t'_1}{t_1; t_2 \longrightarrow t'_1; t_2} \quad (\text{E-SEQ})$$

$$\text{unit}; t_2 \longrightarrow t_2 \quad (\text{E-SEQNEXT})$$

$$\frac{\Gamma \vdash t_1 : \text{Unit} \quad \Gamma \vdash t_2 : T_2}{\Gamma \vdash t_1; t_2 : T_2} \quad (\text{T-SEQ})$$

Derived forms

- ▶ Syntactic sugar
- ▶ Internal language vs. external (surface) language

Sequencing as a derived form

$$t_1; t_2 \stackrel{\text{def}}{=} (\lambda x:\text{Unit}. t_2) t_1$$

where $x \notin FV(t_2)$

Equivalence of the two definitions

[board]

Ascription

New syntactic forms

$t ::= \dots$
 $t \text{ as } T$

New evaluation rules

$v_1 \text{ as } T \longrightarrow v_1$

(E-ASCRIIBE)

$$\frac{t_1 \longrightarrow t'_1}{t_1 \text{ as } T \longrightarrow t'_1 \text{ as } T}$$

(E-ASCRIIBE1)

New typing rules

$$\frac{\Gamma \vdash t_1 : T}{\Gamma \vdash t_1 \text{ as } T : T}$$

(T-ASCRIIBE)

terms

ascription

$t \longrightarrow t'$

$\Gamma \vdash t : T$

Ascription as a derived form

$t \text{ as } T \stackrel{\text{def}}{=} (\lambda x:T. x) t$

Let-bindings

New syntactic forms

$t ::= \dots$

$\text{let } x=t \text{ in } t$

terms

let binding

New evaluation rules

$t \longrightarrow t'$

$\text{let } x=v_1 \text{ in } t_2 \longrightarrow [x \mapsto v_1]t_2$ (E-LETV)

$$\frac{t_1 \longrightarrow t'_1}{\text{let } x=t_1 \text{ in } t_2 \longrightarrow \text{let } x=t'_1 \text{ in } t_2}$$
 (E-LET)

New typing rules

$\Gamma \vdash t : T$

$$\frac{\Gamma \vdash t_1 : T_1 \quad \Gamma, x:T_1 \vdash t_2 : T_2}{\Gamma \vdash \text{let } x=t_1 \text{ in } t_2 : T_2}$$
 (T-LET)

Pairs, tuples, and records

Pairs

$t ::= \dots$
 $\{t, t\}$
 $t.1$
 $t.2$

terms
pair
first projection
second projection

$v ::= \dots$
 $\{v, v\}$

values
pair value

$T ::= \dots$
 $T_1 \times T_2$

types
product type

Evaluation rules for pairs

$$\{v_1, v_2\}.1 \longrightarrow v_1 \quad (\text{E-PAIRBETA1})$$

$$\{v_1, v_2\}.2 \longrightarrow v_2 \quad (\text{E-PAIRBETA2})$$

$$\frac{t_1 \longrightarrow t'_1}{t_1.1 \longrightarrow t'_1.1} \quad (\text{E-PROJ1})$$

$$\frac{t_1 \longrightarrow t'_1}{t_1.2 \longrightarrow t'_1.2} \quad (\text{E-PROJ2})$$

$$\frac{t_1 \longrightarrow t'_1}{\{t_1, t_2\} \longrightarrow \{t'_1, t_2\}} \quad (\text{E-PAIR1})$$

$$\frac{t_2 \longrightarrow t'_2}{\{v_1, t_2\} \longrightarrow \{v_1, t'_2\}} \quad (\text{E-PAIR2})$$

Typing rules for pairs

$$\frac{\Gamma \vdash t_1 : T_1 \quad \Gamma \vdash t_2 : T_2}{\Gamma \vdash \{t_1, t_2\} : T_1 \times T_2} \quad (\text{T-PAIR})$$

$$\frac{\Gamma \vdash t_1 : T_{11} \times T_{12}}{\Gamma \vdash t_{1.1} : T_{11}} \quad (\text{T-PROJ1})$$

$$\frac{\Gamma \vdash t_1 : T_{11} \times T_{12}}{\Gamma \vdash t_{1.2} : T_{12}} \quad (\text{T-PROJ2})$$

Tuples

$t ::= \dots$ *terms*
 $\{t_i^{i \in 1..n}\}$ *tuple*
 $t.i$ *projection*

$v ::= \dots$ *values*
 $\{v_i^{i \in 1..n}\}$ *tuple value*

$T ::= \dots$ *types*
 $\{T_i^{i \in 1..n}\}$ *tuple type*

Evaluation rules for tuples

$$\{v_i \mid i \in 1..n\}.j \longrightarrow v_j \quad (\text{E-PROJTUPLE})$$

$$\frac{t_1 \longrightarrow t'_1}{t_1.i \longrightarrow t'_1.i} \quad (\text{E-PROJ})$$

$$\frac{t_j \longrightarrow t'_j}{\begin{array}{l} \{v_i \mid i \in 1..j-1, t_j, t_k \mid k \in j+1..n\} \\ \longrightarrow \{v_i \mid i \in 1..j-1, t'_j, t_k \mid k \in j+1..n\} \end{array}} \quad (\text{E-TUPLE})$$

Typing rules for tuples

$$\frac{\text{for each } i \quad \Gamma \vdash t_i : T_i}{\Gamma \vdash \{t_i\}_{i \in 1..n} : \{T_i\}_{i \in 1..n}} \quad (\text{T-TUPLE})$$

$$\frac{\Gamma \vdash t_1 : \{T_i\}_{i \in 1..n}}{\Gamma \vdash t_1.j : T_j} \quad (\text{T-PROJ})$$

Records

$t ::= \dots$	<i>terms</i>
$\{l_i = t_i \mid i \in 1..n\}$	<i>record</i>
$t.l$	<i>projection</i>
$v ::= \dots$	<i>values</i>
$\{l_i = v_i \mid i \in 1..n\}$	<i>record value</i>
$T ::= \dots$	<i>types</i>
$\{l_i : T_i \mid i \in 1..n\}$	<i>type of records</i>

Evaluation rules for records

$$\{l_i = v_i \mid i \in 1..n\} . l_j \longrightarrow v_j \quad (\text{E-PROJRCd})$$

$$\frac{t_1 \longrightarrow t'_1}{t_1 . l \longrightarrow t'_1 . l} \quad (\text{E-PROJ})$$

$$\frac{t_j \longrightarrow t'_j}{\begin{array}{l} \{l_i = v_i \mid i \in 1..j-1, l_j = t_j, l_k = t_k \mid k \in j+1..n\} \\ \longrightarrow \{l_i = v_i \mid i \in 1..j-1, l_j = t'_j, l_k = t_k \mid k \in j+1..n\} \end{array}} \quad (\text{E-RCd})$$

Typing rules for records

$$\frac{\text{for each } i \quad \Gamma \vdash t_i : T_i}{\Gamma \vdash \{l_i = t_i \mid i \in 1..n\} : \{l_i : T_i \mid i \in 1..n\}} \quad (\text{T-RCD})$$

$$\frac{\Gamma \vdash t_1 : \{l_i : T_i \mid i \in 1..n\}}{\Gamma \vdash t_1.l_j : T_j} \quad (\text{T-PROJ})$$

Sums and variants

Sums – motivating example

```
PhysicalAddr = {firstlast:String, addr:String}
VirtualAddr  = {name:String, email:String}
Addr         = PhysicalAddr + VirtualAddr
inl  : "PhysicalAddr → PhysicalAddr+VirtualAddr"
inr  : "VirtualAddr  → PhysicalAddr+VirtualAddr"
```

```
getName = λa:Addr.
  case a of
    inl x ⇒ x.firstlast
  | inr y ⇒ y.name;
```

New syntactic forms

<code>t ::= ...</code>	<i>terms</i>
<code>inl t</code>	<i>tagging (left)</i>
<code>inr t</code>	<i>tagging (right)</i>
<code>case t of inl x⇒t inr x⇒t</code>	<i>case</i>
<code>v ::= ...</code>	<i>values</i>
<code>inl v</code>	<i>tagged value (left)</i>
<code>inr v</code>	<i>tagged value (right)</i>
<code>T ::= ...</code>	<i>types</i>
<code>T+T</code>	<i>sum type</i>

T_1+T_2 is a *disjoint union* of T_1 and T_2 (the tags `inl` and `inr` ensure disjointness)

New evaluation rules

$$t \longrightarrow t'$$

$$\begin{array}{l} \text{case (inl } v_0) \\ \text{of inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2 \end{array} \longrightarrow [x_1 \mapsto v_0]t_1 \quad (\text{E-CASEINL})$$

$$\begin{array}{l} \text{case (inr } v_0) \\ \text{of inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2 \end{array} \longrightarrow [x_2 \mapsto v_0]t_2 \quad (\text{E-CASEINR})$$

$$\frac{t_0 \longrightarrow t'_0}{\begin{array}{l} \text{case } t_0 \text{ of inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2 \\ \longrightarrow \text{case } t'_0 \text{ of inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2 \end{array}} \quad (\text{E-CASE})$$

$$\frac{t_1 \longrightarrow t'_1}{\text{inl } t_1 \longrightarrow \text{inl } t'_1} \quad (\text{E-INL})$$

$$\frac{t_1 \longrightarrow t'_1}{\text{inr } t_1 \longrightarrow \text{inr } t'_1} \quad (\text{E-INR})$$

New typing rules

$\Gamma \vdash t : T$

$$\frac{\Gamma \vdash t_1 : T_1}{\Gamma \vdash \text{inl } t_1 : T_1 + T_2} \quad (\text{T-INL})$$

$$\frac{\Gamma \vdash t_1 : T_2}{\Gamma \vdash \text{inr } t_1 : T_1 + T_2} \quad (\text{T-INR})$$

$$\frac{\Gamma \vdash t_0 : T_1 + T_2 \quad \Gamma, x_1 : T_1 \vdash t_1 : T \quad \Gamma, x_2 : T_2 \vdash t_2 : T}{\Gamma \vdash \text{case } t_0 \text{ of inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2 : T} \quad (\text{T-CASE})$$

Sums and Uniqueness of Types

Problem:

If t has type T , then $\text{inl } t$ has type $T+U$ for every U .

I.e., we've lost uniqueness of types.

Possible solutions:

- ▶ “Infer” U as needed during typechecking
- ▶ Give constructors different names and only allow each name to appear in one sum type (requires generalization to “variants,” which we'll see next) — OCaml's solution
- ▶ Annotate each inl and inr with the intended sum type.

For simplicity, let's choose the third.

New syntactic forms

`t ::= ...`
`inl t as T`
`inr t as T`

terms
tagging (left)
tagging (right)

`v ::= ...`
`inl v as T`
`inr v as T`

values
tagged value (left)
tagged value (right)

Note that `as T` here is not the ascription operator that we saw before — i.e., not a separate syntactic form: in essence, there is an ascription “built into” every use of `inl` or `inr`.

New typing rules

$\Gamma \vdash t : T$

$$\frac{\Gamma \vdash t_1 : T_1}{\Gamma \vdash \text{inl } t_1 \text{ as } T_1+T_2 : T_1+T_2} \quad (\text{T-INL})$$

$$\frac{\Gamma \vdash t_1 : T_2}{\Gamma \vdash \text{inr } t_1 \text{ as } T_1+T_2 : T_1+T_2} \quad (\text{T-INR})$$

Evaluation rules ignore annotations:

$t \longrightarrow t'$

case (inl v_0 as T_0)
of inl $x_1 \Rightarrow t_1$ | inr $x_2 \Rightarrow t_2$ (E-CASEINL)
 $\longrightarrow [x_1 \mapsto v_0]t_1$

case (inr v_0 as T_0)
of inl $x_1 \Rightarrow t_1$ | inr $x_2 \Rightarrow t_2$ (E-CASEINR)
 $\longrightarrow [x_2 \mapsto v_0]t_2$

$$\frac{t_1 \longrightarrow t'_1}{\text{inl } t_1 \text{ as } T_2 \longrightarrow \text{inl } t'_1 \text{ as } T_2} \quad \text{(E-INL)}$$

$$\frac{t_1 \longrightarrow t'_1}{\text{inr } t_1 \text{ as } T_2 \longrightarrow \text{inr } t'_1 \text{ as } T_2} \quad \text{(E-INR)}$$

Variants

Just as we generalized binary products to labeled records, we can generalize binary sums to labeled *variants*.

New syntactic forms

$t ::= \dots$
 $\langle l=t \rangle$ as T
 $\text{case } t \text{ of } \langle l_j=x_j \rangle \Rightarrow t_j \quad i \in 1..n$

terms
tagging
case

$T ::= \dots$
 $\langle l_j:T_j \quad i \in 1..n \rangle$

types
type of variants

New evaluation rules

$$t \longrightarrow t'$$

$$\text{case } \langle l_j = v_j \rangle \text{ as } T \text{ of } \langle l_i = x_i \rangle \Rightarrow t_i \text{ }^{i \in 1..n} \longrightarrow [x_j \mapsto v_j] t_j \quad (\text{E-CASEVARIANT})$$

$$\frac{t_0 \longrightarrow t'_0}{\text{case } t_0 \text{ of } \langle l_i = x_i \rangle \Rightarrow t_i \text{ }^{i \in 1..n} \longrightarrow \text{case } t'_0 \text{ of } \langle l_i = x_i \rangle \Rightarrow t_i \text{ }^{i \in 1..n}} \quad (\text{E-CASE})$$

$$\frac{t_i \longrightarrow t'_i}{\langle l_i = t_i \rangle \text{ as } T \longrightarrow \langle l_i = t'_i \rangle \text{ as } T} \quad (\text{E-VARIANT})$$

New typing rules

$\Gamma \vdash t : T$

$$\frac{\Gamma \vdash t_j : T_j}{\Gamma \vdash \langle l_j = t_j \rangle \text{ as } \langle l_i : T_i \rangle_{i \in 1..n} : \langle l_i : T_i \rangle_{i \in 1..n}} \text{ (T-VARIANT)}$$

$$\frac{\begin{array}{c} \Gamma \vdash t_0 : \langle l_i : T_i \rangle_{i \in 1..n} \\ \text{for each } i \quad \Gamma, x_i : T_i \vdash t_i : T \end{array}}{\Gamma \vdash \text{case } t_0 \text{ of } \langle l_i = x_i \rangle \Rightarrow t_i \text{ }_{i \in 1..n} : T} \text{ (T-CASE)}$$

Example

```
Addr = <physical:PhysicalAddr, virtual:VirtualAddr>;
```

```
a = <physical=pa> as Addr;
```

```
getName = λa:Addr.
```

```
  case a of
```

```
    <physical=x> ⇒ x.firstlast
```

```
  | <virtual=y> ⇒ y.name;
```

Options

Just like in OCaml...

```
OptionalNat = <none:Unit, some:Nat>;
```

```
Table = Nat → OptionalNat;
```

```
emptyTable = λn:Nat. <none=unit> as OptionalNat;
```

```
extendTable =
```

```
  λt:Table. λm:Nat. λv:Nat.
```

```
    λn:Nat.
```

```
      if equal n m then <some=v> as OptionalNat
```

```
      else t n;
```

```
x = case t(5) of
```

```
  <none=u> ⇒ 999
```

```
  | <some=v> ⇒ v;
```

Enumerations

```
Weekday = <monday:Unit, tuesday:Unit, wednesday:Unit,  
          thursday:Unit, friday:Unit>;
```

```
nextBusinessDay = λw:Weekday.
```

```
  case w of <monday=x>    ⇒ <tuesday=unit> as Weekday  
           | <tuesday=x>   ⇒ <wednesday=unit> as Weekday  
           | <wednesday=x> ⇒ <thursday=unit> as Weekday  
           | <thursday=x> ⇒ <friday=unit> as Weekday  
           | <friday=x>   ⇒ <monday=unit> as Weekday;
```

Recursion

Recursion in λ_{\rightarrow}

- ▶ In λ_{\rightarrow} , all programs terminate. (Cf. Chapter 12.)
- ▶ Hence, untyped terms like `omega` and `fix` are not typable.
- ▶ But we can *extend* the system with a (typed) fixed-point operator...

Example

```
ff = λie:Nat→Bool.  
    λx:Nat.  
      if iszero x then true  
      else if iszero (pred x) then false  
      else ie (pred (pred x));  
  
iseven = fix ff;  
  
iseven 7;
```

New syntactic forms

$t ::= \dots$
 $\text{fix } t$

terms

fixed point of t

New evaluation rules

$$t \longrightarrow t'$$

$$\text{fix } (\lambda x:T_1.t_2) \longrightarrow [x \mapsto (\text{fix } (\lambda x:T_1.t_2))]t_2 \quad (\text{E-FIXBETA})$$

$$\frac{t_1 \longrightarrow t'_1}{\text{fix } t_1 \longrightarrow \text{fix } t'_1} \quad (\text{E-FIX})$$

New typing rules

$\Gamma \vdash t : T$

$$\frac{\Gamma \vdash t_1 : T_1 \rightarrow T_1}{\Gamma \vdash \text{fix } t_1 : T_1}$$

(T-FIX)

A more convenient form

`letrec x:T1=t1 in t2` $\stackrel{\text{def}}{=} \text{let } x = \text{fix } (\lambda x:T_1.t_1) \text{ in } t_2$

```
letrec iseven : Nat → Bool =
  λx:Nat.
    if iszero x then true
    else if iszero (pred x) then false
    else iseven (pred (pred x))
in
  iseven 7;
```