

Foundations of Software Fall 2023

Week 7

Plan

PREVIOUSLY: unit, sequencing, let, pairs, tuples

TODAY:

1. options, variants
2. recursion
3. state

NEXT: exceptions?

NEXT: polymorphic (not so simple) typing

Records

$t ::= \dots$
 $\{l_i = t_i \mid i \in 1..n\}$
 $t.l$

terms
record
projection

$v ::= \dots$
 $\{l_i = v_i \mid i \in 1..n\}$

values
record value

$T ::= \dots$
 $\{l_i : T_i \mid i \in 1..n\}$

types
type of records

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 =  $\lambda$ a:Addr.
  case a of
    inl x  $\Rightarrow$  x.firstlast
  | inr y  $\Rightarrow$  y.name;
```

New syntactic forms

$t ::=$...	terms
	$\text{inl } t$	tagging (left)
	$\text{inr } t$	tagging (right)
	$\text{case } t \text{ of } \text{inl } x \Rightarrow t \mid \text{inr } x \Rightarrow t$	case
$v ::=$...	values
	$\text{inl } v$	tagged value (left)
	$\text{inr } v$	tagged value (right)
$T ::=$...	types
	$T + T$	sum type

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

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

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

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

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

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

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

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

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

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:

$$\boxed{t \longrightarrow t'}$$

$$\begin{array}{l} \text{case (inl } v_0 \text{ as } T_0) \\ \text{of inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2 \\ \longrightarrow [x_1 \mapsto v_0]t_1 \end{array} \quad (\text{E-CASEINL})$$

$$\begin{array}{l} \text{case (inr } v_0 \text{ as } T_0) \\ \text{of inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2 \\ \longrightarrow [x_2 \mapsto v_0]t_2 \end{array} \quad (\text{E-CASEINR})$$

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

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

Variants

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

New syntactic forms

$t ::= \dots$
 $\langle l=t \rangle \text{ as } T$
 case t of $\langle l_i=x_i \rangle \Rightarrow t_i \quad i \in 1..n$

terms
 tagging
 case

$T ::= \dots$
 $\langle l_i:T_i \quad i \in 1..n \rangle$

types
 type of variants

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

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

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

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

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

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;
```