

# Foundations of Certified Programming Language and Compiler Design

**Dr.-Ing. Sebastian Ertel**

Composable Operating Systems Group, Barkhausen Institute

# Outline



Lecture	Logic	Formalisms	PL
1	Propositional and first-order logic		
2			Functional programming
3		Syntax and Semantics	
4			The untyped lambda calculus
5		Types	
6			The typed lambda calculus
7			Polymorphism
8		Curry-Howard	
9			Higher-order types
10			Dependent types

# The simply typed lambda calculus



## Syntax:

$t ::=$	$x$	terms:	$T ::=$		types:
$ $	$\lambda x:T.t$	variable	$ $	$T \rightarrow T$	type of functions
$ $	$t\ t$	abstraction	$\Gamma ::=$		contexts:
$ $		application	$ $	$\emptyset$	empty context
$v ::=$	$\lambda x:T.t$	values:	$ $	$\Gamma, x:T$	term variable binding
$ $		abstraction value			

## Semantics:

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

$$\frac{t_1 \longrightarrow t'_1}{t_1\ t_2 \longrightarrow t'_1\ t_2} \text{ E-APP1}$$

$$\frac{t_2 \longrightarrow t'_2}{v_1\ t_2 \longrightarrow v_1\ t'_2} \text{ E-APP2}$$

$$\frac{}{(\lambda x:T.t_{12})\ v_2 \longrightarrow [x \mapsto v_2]t_{12}} \text{ E-APPABS}$$

# The simply typed lambda calculus



## Syntax:

$t$	$::=$	$x$	terms:	$T$	$::=$	$T \rightarrow T$	types:
		$\lambda x : T. t$	variable				type of functions
		$t t$	abstraction	$\Gamma$	$::=$	$\emptyset$	contexts:
			application			$\Gamma, x : T$	empty context
$v$	$::=$	$\lambda x : T. t$	values:				term variable binding
			abstraction value				

## Typing

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

$$\frac{x : T \in \Gamma}{\Gamma \vdash x : T} \text{ T-VAR}$$

$$\frac{\Gamma, x : T_1 \vdash t_2 : T_2}{\Gamma \vdash \lambda x : T_1. t_2 : T_1 \rightarrow T_2} \text{ T-ABS}$$

$$\frac{\Gamma \vdash t_1 : T_{11} \rightarrow T_{12} \quad \Gamma \vdash t_2 : T_{11}}{\Gamma \vdash t_1 t_2 : T_{12}} \text{ T-APP}$$



## Derivation Trees

- Typing derivations:

$$\frac{\frac{\frac{x : \text{Bool} \in x : \text{Bool}}{x : \text{Bool} \vdash x : \text{Bool}} \text{ T-VAR} \quad \frac{}{\vdash \text{true} : \text{Bool}} \text{ T-TRUE}}{\lambda x : \text{Bool}. x : \text{Bool} \rightarrow \text{Bool} \quad \vdash (\lambda x : \text{Bool}. x) \text{ true} : \text{Bool}} \text{ T-APP}$$

- Check: Show the derivation tree for

$f : \text{Bool} \rightarrow \text{Bool} \vdash f (\text{if false then true else false}) : \text{Bool}$

$$\frac{\frac{f : \text{Bool} \rightarrow \text{Bool} \in \Gamma}{\Gamma \vdash f : \text{Bool} \rightarrow \text{Bool}} \text{ T-VAR} \quad \frac{\frac{\frac{\frac{}{\vdash \text{false} : \text{Bool}} \text{ T-FALSE} \quad \frac{}{\vdash \text{true} : \text{Bool}} \text{ T-TRUE}}{\vdash \text{if false then true else false} : \text{Bool}} \text{ T-IF}}{f : \text{Bool} \rightarrow \text{Bool} \vdash f (\text{if false then true else false}) : \text{Bool}} \text{ T-APP}$$

- Check: Find a context  $\Gamma$  for  $f \ x \ y : \text{Bool}$ .

$$\Gamma = \begin{array}{lll} f : \text{Bool} \rightarrow \text{Bool} \rightarrow \text{Bool}, & f : \text{Nat} \rightarrow \text{Nat} \rightarrow \text{Bool}, & f : \text{T} \rightarrow \text{T} \rightarrow \text{Bool} \\ x : \text{Bool}, y : \text{Bool} & x : \text{Nat}, y : \text{Nat} & x : \text{T}, y : \text{T} \end{array}$$

# Properties



## Lemmas:

Inversion of the typing relation.

Canonical forms.

## Theorems:

Uniqueness of types.

### Theorem (Progress)

*Suppose  $t$  is a closed, well-typed term, i.e.,  $\vdash t : T$ . Then either  $t$  is a value or there exists some  $t'$  such that  $t \longrightarrow t'$ .*

### Theorem (Normalization)

*If  $\vdash t : T$ , then  $t$  is normalizable.*

### Theorem (Preservation)

*If  $\Gamma \vdash t : T$  and  $t \longrightarrow t'$ , then  $\Gamma \vdash t' : T$*

What is the type of  $(\lambda x.x\ x)\ (\lambda x.x\ x)$ ?



- Currently, we cannot implement the STLC because we are missing the base case for our types.
- Let's study extensions:



# Extensions

## Base Types

- Base types: `Nat`, `Bool`, `String`, `Float`, ...
- Let's add some uninterpreted/unknown base types without any primitive operations.

New syntactic forms:

$$\begin{array}{ll} T ::= \dots & \text{types:} \\ | \mathbf{A} & \text{base type} \end{array}$$

- Consider:  $\lambda x : A. x : A \rightarrow A$
- We could assume that  $A = \text{Nat}$  is some number:  $(\lambda x : A. x : A \rightarrow A) 5$





- But where do these values come from and what actually is  $\mathbb{Nat}$ ?
- We do not want to add unknown values/types?!
- We would like to have a closed system to reason about it.



New syntactic forms:

$t ::= \dots$	terms:	$v ::= \dots$	values:
<code>unit</code>	constant unit	<code>unit</code>	constant unit

$T ::= \dots$	types:
<code>Unit</code>	unit type

New Typing Rules:

$\Gamma \vdash t : T$

$\frac{}{\text{unit} : \text{Unit}} \text{ T-UNIT}$

# Ascription

- Document your code ... with types!

New syntactic forms:

$t ::= \dots$  terms:  
 $| \text{ } t \text{ as } T \text{ ascription}$

New Evaluation Rules:

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

$$\frac{}{v_1 \text{ as } T_1 \longrightarrow v_1} \text{E-ASCRIBE1}$$

$$\frac{t_1 \longrightarrow t'_1}{t_1 \text{ as } T_1 \longrightarrow t'_1 \text{ as } T_1} \text{E-ASCRIBE2}$$

New Typing Rules:

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

$$\frac{\Gamma \vdash t_1 : T_1}{\Gamma \vdash t_1 \text{ as } T_1 : T_1} \text{T-ASCRIBE}$$

# Product Types a.k.a. Pairs a.k.a. Tuples (with two elements)



New syntactic forms:

$t ::= \dots$	terms:	$v ::= \dots$	values:
$\{t, t\}$	pair	$\{v, v\}$	pair
$t.1$	first projection		
$t.2$	second projection	$T ::= \dots$	types:
		$T_1 \times T_2$	product type

New Evaluation Rules:

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

$$\frac{}{\{v_1, v_2\}.1 \longrightarrow v_1} \text{E-PAIRBETA1}$$

$$\frac{}{\{v_1, v_2\}.2 \longrightarrow v_2} \text{E-PAIRBETA2}$$

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

$$\frac{t_2 \longrightarrow t'_2}{t_2.2 \longrightarrow t'_2.2} \text{E-PROJ2}$$

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

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



# Product Types a.k.a. Pairs a.k.a. Tuples (with two elements)

New syntactic forms:

$t ::= \dots$	terms:	$v ::= \dots$	values:
$\{t, t\}$	pair	$\{v, v\}$	pair
$t.1$	first projection		
$t.2$	second projection	$T ::= \dots$	types:
		$T_1 \times T_2$	product type

New Typing Rules:

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

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

$$\frac{\Gamma \vdash t : T_1 \times T_2}{\Gamma \vdash t.1 : T_1} \text{ T-PROJ1}$$

$$\frac{\Gamma \vdash t : T_1 \times T_2}{\Gamma \vdash t.2 : T_2} \text{ T-PROJ2}$$

# Sum Types



## New syntactic forms:

$t ::= \dots$

| `inl t`  
| `inr t`  
| `case t of inl x  $\Rightarrow$  t | inr x  $\Rightarrow$  t`

terms:

tagging (left)  
tagging (right)  
case

$v ::= \dots$

values:

| `inl v` tagged value (left)  
| `inr v` tagged value (right)

$T ::= \dots$

types:

| `T + T` sum type

## New Evaluation Rules:

$t \longrightarrow t'$

$$\frac{t \longrightarrow t'}{\text{inl } t \longrightarrow \text{inl } t'} \text{ E-INL}$$

$$\frac{t \longrightarrow t'}{\text{inr } t \longrightarrow \text{inr } t'} \text{ E-INR}$$

# Sum Types



New syntactic forms:

$t ::= \dots$	terms:	$v ::= \dots$	values:
<code>inl t</code>	tagging (left)	<code>inl v</code>	tagged value (left)
<code>inr t</code>	tagging (right)	<code>inr v</code>	tagged value (right)
<code>case t of inl x <math>\Rightarrow</math> t   inr x <math>\Rightarrow</math> t</code>	case		
		$T ::= \dots$	types:
		<code>T + T</code>	sum type

New Evaluation Rules:

$$t \longrightarrow t'$$

$$\frac{}{\text{case inl } v_1 \text{ of inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2 \longrightarrow [x_1 \mapsto v_1]t_1} \text{E-CASEINL}$$

$$\frac{}{\text{case inr } v_2 \text{ of inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2 \longrightarrow [x_2 \mapsto v_2]t_2} \text{E-CASEINR}$$

$$\frac{t_0 \longrightarrow t'_0}{\text{case } t_0 \text{ of inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2 \longrightarrow \text{case } t'_0 \text{ of inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2} \text{E-CASE}$$

# Sum Types



New syntactic forms:

$t ::= \dots$

| `inl t`  
| `inr t`  
| `case t of inl x  $\Rightarrow$  t | inr x  $\Rightarrow$  t`

terms:

tagging (left)  
tagging (right)  
case

$v ::= \dots$

values:

| `inl v` tagged value (left)  
| `inr v` tagged value (right)

$T ::= \dots$

types:

| `T + T` sum type

New Typing Rules:

$\Gamma \vdash t : T$

$$\frac{\Gamma \vdash t : T_1}{\Gamma \vdash \text{inl } t : T_1 + T_2} \text{ T-INL}$$

$$\frac{\Gamma \vdash t : T_2}{\Gamma \vdash \text{inl } t : T_1 + T_2} \text{ T-INR}$$

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



# Labeled Product Types a.k.a. Records



New syntactic forms:

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

New Evaluation Rules:

$$t \longrightarrow t'$$

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

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

# Labeled Product Types a.k.a. Records



New syntactic forms:

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

New Evaluation Rules:

$$t \longrightarrow t'$$

$$\frac{t_j \longrightarrow t'_j}{\{l_i = v_i \mid i \in 1..j-1, l_j = t_j, l_k = t_k \mid k \in j+1..n\} \longrightarrow \{l_i = v_i \mid i \in 1..j-1, l_j = t'_j, l_k = t_k \mid k \in j+1..n\}} \text{ E-Rcd}$$

# Labeled Product Types a.k.a. Records



New syntactic forms:

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

New Typing Rules:

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

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

$$\frac{\Gamma \vdash t_1 : \{l_i : T_i \mid i \in 1..j-1\}, l_j : T_j, l_k : T_k \mid k \in j+1..n}{\Gamma \vdash \{t_1.l_j\} : T_j} \text{ T-PROJ}$$

## Sum Types (again)

### Uniqueness of Typing



Valid types for `inl 5` are `Nat + Nat`, `Nat + Bool` etc.

New syntactic forms:

$t ::= \dots$	terms:	$v ::= \dots$	values:
<code>inl <math>t</math> as <math>T</math></code>	tagging (left)	<code>inl <math>v</math> as <math>T</math></code>	tagged value (left)
<code>inr <math>t</math> as <math>T</math></code>	tagging (right)	<code>inr <math>v</math> as <math>T</math></code>	tagged value (right)

New Evaluation Rules:

$t \longrightarrow t'$	
$\frac{}{\text{case } (\text{inl } v_1 \text{ as } T) \text{ of } \text{inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2 \longrightarrow [x_1 \mapsto v_1]t_1}$	E-CASEINL
$\frac{}{\text{case } (\text{inr } v_2 \text{ as } T) \text{ of } \text{inl } x_1 \Rightarrow t_1 \mid \text{inr } x_2 \Rightarrow t_2 \longrightarrow [x_2 \mapsto v_2]t_2}$	E-CASEINR

## Sum Types (again)

### Uniqueness of Typing



Valid types for `inl 5` are `Nat + Nat`, `Nat + Bool` etc.

New syntactic forms:

$t ::= \dots$	terms:	$v ::= \dots$	values:
<code>inl t</code> <b>as</b> <code>T</code>	tagging (left)	<code>inl v</code> <b>as</b> <code>T</code>	tagged value (left)
<code>inr t</code> <b>as</b> <code>T</code>	tagging (right)	<code>inr v</code> <b>as</b> <code>T</code>	tagged value (right)

New Typing Rules:

$$\boxed{\Gamma \vdash t : T} \qquad
 \frac{\Gamma \vdash t : T_1}{\Gamma \vdash \text{inl } t \text{ **as** } T_1 + T_2 : T_1 + T_2} \text{ T-INL} \qquad
 \frac{\Gamma \vdash t : T_2}{\Gamma \vdash \text{inl } t \text{ **as** } T_1 + T_2 : T_1 + T_2} \text{ T-INR}$$

# Labeled Sum Types a.k.a. Variants a.k.a. Algebraic Datatypes



New syntactic forms:

$t ::= \dots$

|  $\langle l=t \rangle \text{ as } T$   
|  $\text{case } t \text{ of } \langle l_i=t_i \rangle \Rightarrow t_i^{i \in 1..n}$

terms:

tagging  
case

$v ::= \dots$

|  $\langle l=v \rangle \text{ as } T$  tagged value

values:

$T ::= \dots$

types:

|  $\langle l_i : T_i^{i \in 1..n} \rangle$  type of variants

New Evaluation Rules:

$t \longrightarrow t'$

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

# Labeled Sum Types a.k.a. Variants a.k.a. Algebraic Datatypes



New syntactic forms:

$t ::= \dots$	terms:	$v ::= \dots$	values:
$\langle l=t \rangle \text{ as } T$	tagging	$\langle l=v \rangle \text{ as } T$	tagged value
$\text{case } t \text{ of } \langle l_i=t_i \rangle \Rightarrow t_i^{i \in 1..n}$	case	$T ::= \dots$	types:
		$\langle l_i:T_i^{i \in 1..n} \rangle$	type of variants

New Evaluation Rules:

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

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

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

# Labeled Sum Types a.k.a. Variants a.k.a. Algebraic Datatypes



New syntactic forms:

$t ::= \dots$ $  \quad \langle l=t \rangle \text{ as } T$ $  \quad \text{case } t \text{ of } \langle l_i=t_i \rangle \Rightarrow t_i^{i \in 1..n}$	terms: tagging case	$v ::= \dots$ $  \quad \langle l=v \rangle \text{ as } T$ $T ::= \dots$ $  \quad \langle l_i : T_i^{i \in 1..n} \rangle$	values: tagged value  types: type of variants
---	---------------------------	---	---

New Typing Rules:

$\Gamma \vdash t : T$

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

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



# Examples



## STLC++

```
Bool = <false:Unit, true:Unit>
false = <false=unit> as Bool
true = <true=unit> as Bool

Week = <weekday:Unit, weekend:Unit>
weekday = <weekday=unit> as Week
weekend = <weekend=unit> as Week
weekendYet =  $\lambda x:\text{Week}.$ 
  case x of
    <weekday=unit>  $\Rightarrow$  false
    <weekend=unit>  $\Rightarrow$  true

OptionNat = <none:Unit, some:Nat>
NatList = <nil:Unit, cons:{Nat, NatList}>
Nat = <zero:Unit, succ:Nat>
```

## Coq

```
Inductive bool := False | True.

Inductive week = Weekday | Weekend.

Inductive optionNat = None | Some (_:nat_)
Inductive natList = Nil | Cons (_:nat) (_:natList).
Inductive nat = Zero | Succ (_:nat).
```