

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

Inference Rules Recap



So far, we used inference rules to specify:

Inference Rules Recap



So far, we used inference rules to specify:

- the syntax of terms

$$\frac{}{0 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T}}{\text{succ } t_1 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T} \quad t_2 \in \mathcal{T}}{t_1 + t_2 \in \mathcal{T}}$$



So far, we used inference rules to specify:

- the syntax of terms

$$\frac{}{0 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T}}{\text{succ } t_1 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T} \quad t_2 \in \mathcal{T}}{t_1 + t_2 \in \mathcal{T}}$$

- the “preservation” of a predicate

$$\frac{}{P(\text{Zero})} \quad \frac{t_1 \in \text{Term} \quad P(t_1)}{P(\text{Succ } t_1)} \quad \frac{t_1 \in \text{Term} \quad P(t_1) \quad t_2 \in \text{Term} \quad P(t_2)}{P(\text{Add } t_1 \ t_2)}$$

Inference Rules Recap



- That is, we used inference rules to specify relations:



Inference Rules Recap

- That is, we used inference rules to specify relations:
- the relation \in of terms(/words) t and the set of all terms \mathcal{T}

$$t \in \mathcal{T}$$

$$\frac{}{0 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T}}{\text{succ } t_1 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T} \quad t_2 \in \mathcal{T}}{t_1 + t_2 \in \mathcal{T}}$$

Inference Rules Recap



- That is, we used inference rules to specify relations:
- the relation \in of terms(/words) t and the set of all terms \mathcal{T}

$$t \in \mathcal{T}$$

$$\frac{}{0 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T}}{\text{succ } t_1 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T} \quad t_2 \in \mathcal{T}}{t_1 + t_2 \in \mathcal{T}}$$

- the property P of terms t

$$P(t)$$

$$\frac{}{P(\text{Zero})} \quad \frac{t_1 \in \text{term} \quad P(t_1)}{P(\text{Succ } t_1)} \quad \frac{t_1 \in \text{term} \quad P(t_1) \quad t_2 \in \text{term} \quad P(t_2)}{P(\text{Add } t_1 \ t_2)}$$

Inference Rules Recap

- That is, we used inference rules to specify relations:
- the relation \in of terms(/words) t and the set of all terms \mathcal{T}

$$t \in \mathcal{T}$$

$$\frac{}{0 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T}}{\text{succ } t_1 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T} \quad t_2 \in \mathcal{T}}{t_1 + t_2 \in \mathcal{T}}$$

- the property P of terms t

$$P(t)$$

$$\frac{}{P(\text{Zero})} \quad \frac{t_1 \in \text{term} \quad P(t_1)}{P(\text{Succ } t_1)} \quad \frac{t_1 \in \text{term} \quad P(t_1) \quad t_2 \in \text{term} \quad P(t_2)}{P(\text{Add } t_1 \ t_2)}$$

- A function is a type of relation that associates one input with exactly one output.



- That is, we used inference rules to specify relations:
- the relation \in of terms(/words) t and the set of all terms \mathcal{T}

$$t \in \mathcal{T}$$

$$\frac{}{0 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T}}{\text{succ } t_1 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T} \quad t_2 \in \mathcal{T}}{t_1 + t_2 \in \mathcal{T}}$$

- the property P of terms t

$$P(t)$$

$$\frac{}{P(\text{Zero})} \quad \frac{t_1 \in \text{term} \quad P(t_1)}{P(\text{Succ } t_1)} \quad \frac{t_1 \in \text{term} \quad P(t_1) \quad t_2 \in \text{term} \quad P(t_2)}{P(\text{Add } t_1 \ t_2)}$$

- A function is a type of relation that associates one input with exactly one output.
- A binary relation is a set of (ordered) pairs where one input maybe related to more than one output.

Inference Rules Recap

- That is, we used inference rules to specify relations:
- the relation \in of terms(/words) t and the set of all terms \mathcal{T}

$$t \in \mathcal{T}$$

$$\frac{}{0 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T}}{\text{succ } t_1 \in \mathcal{T}} \quad \frac{t_1 \in \mathcal{T} \quad t_2 \in \mathcal{T}}{t_1 + t_2 \in \mathcal{T}}$$

- the property P of terms t

$$P(t)$$

$$\frac{}{P(\text{Zero})} \quad \frac{t_1 \in \text{term} \quad P(t_1)}{P(\text{Succ } t_1)} \quad \frac{t_1 \in \text{term} \quad P(t_1) \quad t_2 \in \text{term} \quad P(t_2)}{P(\text{Add } t_1 \ t_2)}$$

- A function is a type of relation that associates one input with exactly one output.
- A binary relation is a set of (ordered) pairs where one input maybe related to more than one output.
- Here is one important relation for the definition of programming languages:



Mathematically: For programming languages such a relation is the evaluation of a term:

Definition (One-Step Evaluation Relation)

The *one-step* evaluation relation \longrightarrow is the smallest binary relation that relates a term t of a language to another term t' .



Mathematically: For programming languages such a relation is the evaluation of a term:

Definition (One-Step Evaluation Relation)

The *one-step* evaluation relation \longrightarrow is the smallest binary relation that relates a term t of a language to another term t' .

Logically: This relation defines a term rewriting system.

Semantically: We talk about evaluating a term, i.e., we reduce it to another term of a smaller size.

Example: Booleans



Syntax

t	$::=$		terms:
		true	constant true
		false	constant false
		if t then t else t	conditional

v	$::=$		values:
		true	true value
		false	false value

Example: Booleans



Syntax

$t ::=$
| true
| false
| if t then t else t

terms:
constant true
constant false
conditional

$v ::=$
| true true value
| false false value

Evaluation

$\frac{}{\text{if true then } t_2 \text{ else } t_3 \longrightarrow t_2} \text{E-IFTRUE}$

$\frac{}{\text{if false then } t_2 \text{ else } t_3 \longrightarrow t_3} \text{E-IFFALSE}$

$\frac{t_1 \longrightarrow t'_1}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \longrightarrow \text{if } t'_1 \text{ then } t_2 \text{ else } t_3} \text{E-IF}$

Computations with Booleans

Derivation Trees



$s \stackrel{\text{def}}{=} \text{if true then false else false}$

Computations with Booleans

Derivation Trees



$s \stackrel{\text{def}}{=} \text{if true then false else false}$
 $t \stackrel{\text{def}}{=} \text{if } s \text{ then true else true}$

Computations with Booleans

Derivation Trees



$s \stackrel{\text{def}}{=} \text{if true then false else false}$
 $t \stackrel{\text{def}}{=} \text{if } s \text{ then true else true}$
 $u \stackrel{\text{def}}{=} \text{if false then true else true}$

Computations with Booleans

Derivation Trees



$s \stackrel{\text{def}}{=} \text{if true then false else false}$
 $t \stackrel{\text{def}}{=} \text{if } s \text{ then true else true}$
 $u \stackrel{\text{def}}{=} \text{if false then true else true}$

$\frac{}{\text{if } t \text{ then false else false} \longrightarrow \text{if } u \text{ then false else false}} \text{E-If}$

Computations with Booleans

Derivation Trees



$s \stackrel{\text{def}}{=} \text{if true then false else false}$
 $t \stackrel{\text{def}}{=} \text{if } s \text{ then true else true}$
 $u \stackrel{\text{def}}{=} \text{if false then true else true}$

$\text{if } s \text{ then true else true} \longrightarrow \text{if false then true else true}$	E-If
$\text{if } t \text{ then false else false} \longrightarrow \text{if } u \text{ then false else false}$	E-If

Computations with Booleans

Derivation Trees



$s \stackrel{\text{def}}{=} \text{if true then false else false}$
 $t \stackrel{\text{def}}{=} \text{if } s \text{ then true else true}$
 $u \stackrel{\text{def}}{=} \text{if false then true else true}$

$$\frac{\frac{\frac{}{\text{if true then false else false} \rightarrow \text{false}}{\text{if } s \text{ then true else true} \rightarrow \text{if false then true else true}}{\text{if } t \text{ then false else false} \rightarrow \text{if } u \text{ then false else false}} \quad \text{E-IfTrue} \quad \text{E-If} \quad \text{E-If}$$



Theorem (Determinacy of One-Step Evaluation)

If $t \longrightarrow t'$ and $t \longrightarrow t''$ then $t' = t''$.



Theorem (Determinacy of One-Step Evaluation)

If $t \longrightarrow t'$ and $t \longrightarrow t''$ then $t' = t''$.

Proof.

The proof is by induction **on the derivation** $t \longrightarrow t'$, i.e., on the structure of t .



Theorem (Determinacy of One-Step Evaluation)

If $t \longrightarrow t'$ and $t \longrightarrow t''$ then $t' = t''$.

Proof.

The proof is by induction **on the derivation** $t \longrightarrow t'$, i.e., on the structure of t .

Case

$$t \longrightarrow t'$$

$$t \longrightarrow t''$$



Theorem (Determinacy of One-Step Evaluation)

If $t \longrightarrow t'$ and $t \longrightarrow t''$ then $t' = t''$.

Proof.

The proof is by induction **on the derivation** $t \longrightarrow t'$, i.e., on the structure of t .

Case

E-IfTrue

$$\frac{\begin{array}{c} t \longrightarrow t' \\ \vdots \end{array}}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \longrightarrow t_2} \text{ E-IfTrue}$$

where: $t_1 = \text{true}$

$t \longrightarrow t''$

E-IfFalse:

E-If:

$t_1 = \text{false} \neq t_1 = \text{true}$

$t_1 \longrightarrow t'_1$ but $t_1 = \text{true}$ such that $\text{true} \longrightarrow ???$



Theorem (Determinacy of One-Step Evaluation)

If $t \longrightarrow t'$ and $t \longrightarrow t''$ then $t' = t''$.

Proof.

The proof is by induction **on the derivation** $t \longrightarrow t'$, i.e., on the structure of t .

Case

$t \longrightarrow t'$

$t \longrightarrow t''$

E-IfTRUE

$$\frac{\vdots}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \longrightarrow t_2} \text{ E-IfTRUE}$$

where: $t_1 = \text{true}$

E-IfFALSE:

$t_1 = \text{false} \neq t_1 = \text{true}$

E-If:

$t_1 \longrightarrow t'_1$ but $t_1 = \text{true}$ such that $\text{true} \longrightarrow ???$

E-IfFALSE

analogous with $t_1 = \text{false}$



Theorem (Determinacy of One-Step Evaluation)

If $t \longrightarrow t'$ and $t \longrightarrow t''$ then $t' = t''$.

Proof.

The proof is by induction **on the derivation** $t \longrightarrow t'$, i.e., on the structure of t .

Case

$t \longrightarrow t'$

$t \longrightarrow t''$

E-IfTRUE

$$\frac{\vdots}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \longrightarrow t_2} \text{ E-IfTRUE}$$

where: $t_1 = \text{true}$

E-IfFALSE:

$t_1 = \text{false} \neq t_1 = \text{true}$

E-If:

$t_1 \longrightarrow t'_1$ but $t_1 = \text{true}$ such that $\text{true} \longrightarrow ???$

E-IfFALSE

analogous with $t_1 = \text{false}$

E-If

E-If :

By induction hypothesis $t'_1 = t''_1$
for $t_1 \longrightarrow t'_1$ and $t_1 \longrightarrow t''_1$



Definition (Normal Form)

A term t is in *normal form* if no evaluation rule applies to it, i.e., there exists no t' such that $t \longrightarrow t'$.



Definition (Normal Form)

A term t is in *normal form* if no evaluation rule applies to it, i.e., there exists no t' such that $t \longrightarrow t'$.

- Interesting properties for our Booleans language would be:

Theorem

Every value is in normal form.



Definition (Normal Form)

A term t is in *normal form* if no evaluation rule applies to it, i.e., there exists no t' such that $t \longrightarrow t'$.

- Interesting properties for our Booleans language would be:

Theorem

Every value is in normal form.

Theorem

If t is in normal form, then t is a value.



Definition (Multi-step Evaluation Relation)

The multi-step evaluation relation \longrightarrow^* is the reflexive, transitive closure of the one-step evaluation (\longrightarrow). That is, it is the smallest relation such that



Definition (Multi-step Evaluation Relation)

The multi-step evaluation relation \longrightarrow^* is the reflexive, transitive closure of the one-step evaluation (\longrightarrow). That is, it is the smallest relation such that

$$\frac{t \longrightarrow t'}{t \longrightarrow^* t'}$$



Definition (Multi-step Evaluation Relation)

The multi-step evaluation relation \longrightarrow^* is the reflexive, transitive closure of the one-step evaluation (\longrightarrow). That is, it is the smallest relation such that

$$\frac{t \longrightarrow t'}{t \longrightarrow^* t'} \qquad \frac{}{t \longrightarrow^* t}$$



Definition (Multi-step Evaluation Relation)

The multi-step evaluation relation \longrightarrow^* is the reflexive, transitive closure of the one-step evaluation (\longrightarrow). That is, it is the smallest relation such that

$$\frac{t \longrightarrow t'}{t \longrightarrow^* t'} \qquad \frac{}{t \longrightarrow^* t} \qquad \frac{t \longrightarrow^* t' \quad t' \longrightarrow^* t''}{t \longrightarrow^* t''}$$



Definition (Multi-step Evaluation Relation)

The multi-step evaluation relation \longrightarrow^* is the reflexive, transitive closure of the one-step evaluation (\longrightarrow). That is, it is the smallest relation such that

$$\frac{t \longrightarrow t'}{t \longrightarrow^* t'} \qquad \frac{}{t \longrightarrow^* t} \qquad \frac{t \longrightarrow^* t' \quad t' \longrightarrow^* t''}{t \longrightarrow^* t''}$$

Theorem (Uniqueness of Normal Forms)

If $t \longrightarrow^ u$ and $t \longrightarrow^* u'$, where u and u' are normal forms, then $u = u'$*



Definition (Multi-step Evaluation Relation)

The multi-step evaluation relation \longrightarrow^* is the reflexive, transitive closure of the one-step evaluation (\longrightarrow). That is, it is the smallest relation such that

$$\frac{t \longrightarrow t'}{t \longrightarrow^* t'} \qquad \frac{}{t \longrightarrow^* t} \qquad \frac{t \longrightarrow^* t' \quad t' \longrightarrow^* t''}{t \longrightarrow^* t''}$$

Theorem (Uniqueness of Normal Forms)

If $t \longrightarrow^ u$ and $t \longrightarrow^* u'$, where u and u' are normal forms, then $u = u'$*

Proof.

Corollary of the determinacy of one-step evaluation. □



Theorem (Termination of Evaluation)

For every term t there is some normal form t' such that $t \longrightarrow^ t'$.*



Theorem (Termination of Evaluation)

For every term t there is some normal form t' such that $t \longrightarrow^ t'$.*

Proof.

We define a *termination measure* f for the states in our system, i.e., terms in our language. We observe that for every reduction step the size (is a natural number that) decreases and the usual order on the natural numbers is well founded. That is, if there exists a reduction $t \longrightarrow t'$ then $f(t') < f(t)$. □



Theorem (Termination of Evaluation)

For every term t there is some normal form t' such that $t \longrightarrow^ t'$.*

Proof.

We define a *termination measure* f for the states in our system, i.e., terms in our language. We observe that for every reduction step the size (is a natural number that) decreases and the usual order on the natural numbers is well founded. That is, if there exists a reduction $t \longrightarrow t'$ then $f(t') < f(t)$. □

```
Equations size (t:term) : nat :=
size TTrue := 0;
size TFalse := 0;
size (If t1 t2 t3) := 1 + (size t1) + (size t2) + (size t3).
```



Definition (Stuckness)

A closed term is *stuck* if it is in normal form but not a value.



Definition (Stuckness)

A closed term is *stuck* if it is in normal form but not a value.

Consider the extension of our Booleans language to natural numbers:

Additional Syntax

t	$::=$	\dots	terms:
		0	constant zero
		$\text{succ } t$	successor

Additional Evaluation Rules



Definition (Stuckness)

A closed term is *stuck* if it is in normal form but not a value.

Consider the extension of our Booleans language to natural numbers:

Additional Syntax

t	$::=$	\dots	terms:	v	$::=$	\dots	values:
		0	constant zero			nv	numeric value
		$\text{succ } t$	successor				

Additional Evaluation Rules



Definition (Stuckness)

A closed term is *stuck* if it is in normal form but not a value.

Consider the extension of our Booleans language to natural numbers:

Additional Syntax

t	$::=$	\dots	terms:		v	$::=$	\dots	values:		nv	$::=$	\dots	numeric values
		0	constant zero				0	zero value				0	zero value
		$\text{succ } t$	successor				nv	numeric value				$\text{succ } nv$	successor value

Additional Evaluation Rules



Definition (Stuckness)

A closed term is *stuck* if it is in normal form but not a value.

Consider the extension of our Booleans language to natural numbers:

Additional Syntax

$t ::= \dots$	terms:	$v ::= \dots$	values:	$nv ::= \dots$	numeric values
$ 0$	constant zero	$ 0$	zero value	$ 0$	zero value
$ \text{succ } t$	successor	$ nv$	numeric value	$ \text{succ } nv$	successor value

Additional Evaluation Rules

$$\frac{t \longrightarrow t'}{\text{succ } t \longrightarrow \text{succ } t'} \text{ E-Succ}$$



Definition (Stuckness)

A closed term is *stuck* if it is in normal form but not a value.

Consider the extension of our Booleans language to natural numbers:

Additional Syntax

$t ::= \dots$	terms:	$v ::= \dots$	values:	$nv ::= \dots$	numeric values
$ 0$	constant zero	$ 0$	zero value	$ 0$	zero value
$ \text{succ } t$	successor	$ nv$	numeric value	$ \text{succ } nv$	successor value

Additional Evaluation Rules

$$\frac{t \longrightarrow t'}{\text{succ } t \longrightarrow \text{succ } t'} \text{ E-Succ}$$

Can you spot the stuck terms?

Defining the Semantics of Programming Languages



Operational semantics Two flavors exist:

Small-step What we have done so far when defining \longrightarrow .

Defining the Semantics of Programming Languages



Operational semantics Two flavors exist:

Small-step What we have done so far when defining \longrightarrow .

Big-step Essentially, what we have defined so far as \longrightarrow^* but directly defined.

$$\boxed{t \Downarrow v}$$

$$\frac{}{v \Downarrow v} \text{ B-VALUE}$$

$$\frac{t_1 \Downarrow \text{true} \quad t_2 \Downarrow v_2}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \Downarrow v_2} \text{ B-IFTRUE}$$

$$\frac{t_1 \Downarrow \text{false} \quad t_3 \Downarrow v_3}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \Downarrow v_3} \text{ B-IFFALSE}$$

$$\frac{t \Downarrow nv}{\text{succ } t \Downarrow \text{succ } nv} \text{ B-SUCC}$$



Defining the Semantics of Programming Languages

Operational semantics Two flavors exist:

Small-step What we have done so far when defining \longrightarrow .

Big-step Essentially, what we have defined so far as \longrightarrow^* but directly defined.

Denotational Semantics Maps the terms of the language to elements of another domain.

Additional Syntax

$t ::= \dots$ terms:
| $\text{add } t \ t$ addition

Additional Evaluation

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

$$\frac{t_2 \longrightarrow t'_2}{\text{add } nv_1 \ t_2 \longrightarrow \text{add } nv_1 \ t'_2} \text{E-AddRIGHT}$$

$$\frac{nv_1 + nv_2 = nv_3}{\text{add } nv_1 \ nv_2 \longrightarrow nv_3} \text{E-AddRIGHT}$$

Mapping add onto the addition of natural numbers.

Defining the Semantics of Programming Languages



Operational semantics Two flavors exist:

Small-step What we have done so far when defining \longrightarrow .

Big-step Essentially, what we have defined so far as \longrightarrow^* but directly defined.

Denotational Semantics Maps the terms of the language to elements of another domain.

We will come back to this in the final part of this lecture when we want to prove properties such as semantic preservation of compiler transformations etc.

Defining the Semantics of Programming Languages



Operational semantics Two flavors exist:

Small-step What we have done so far when defining \longrightarrow .

Big-step Essentially, what we have defined so far as \longrightarrow^* but directly defined.

Denotational Semantics Maps the terms of the language to elements of another domain.

Axiomatic Semantics Instead of proving properties additionally, the properties are defined right in the system. The most prominent example is *Hoare Logic*.

Defining Relations



Haskell

Gallina (Coq)

Defining Relations



Haskell

recursive

```
data Nat = 0 | S Nat
```

Gallina (Coq)

```
Inductive nat : Type := 0 | S (n : nat).
```

Defining Relations



Haskell

recursive

```
data Nat = 0 | S Nat
```

(GADT)

```
data Nat where
```

```
  0 :: Nat
```

```
  S :: Nat -> Nat
```

Gallina (Coq)

```
Inductive nat : Type := 0 | S (n : nat).
```

```
Inductive nat : Type :=
```

```
  | 0 : nat
```

```
  | S : nat -> nat.
```

Defining Relations



Haskell

recursive

```
data Nat = 0 | S Nat
```

(GADT)

```
data Nat where
```

```
  0 :: Nat
```

```
  S :: Nat -> Nat
```

parameterized

```
data List a :: Type where
```

```
  Nil :: List a
```

```
  Cons :: a -> List a -> List a
```

Gallina (Coq)

```
Inductive nat : Type := 0 | S (n : nat).
```

```
Inductive nat : Type :=
```

```
  | 0 : nat
```

```
  | S : nat -> nat.
```

```
Inductive list (A:Type) : Set :=
```

```
  | nil : list A
```

```
  | cons : A -> list A -> list A.
```

Defining Relations



Haskell

recursive

```
data Nat = 0 | S Nat
```

(GADT)

```
data Nat where
  0 :: Nat
  S :: Nat -> Nat
```

parameterized

```
data List a :: Type where
  Nil :: List a
  Cons :: a -> List a -> List a
```

annotated/indexed

```
data Ev :: Nat -> Type where
  Ev0 :: Ev '0
  EvSS :: Ev n -> Ev ('S ('S n))
```

Gallina (Coq)

```
Inductive nat : Type := 0 | S (n : nat).
```

```
Inductive nat : Type :=
| 0 : nat
| S : nat -> nat.
```

```
Inductive list (A:Type) : Set :=
| nil : list A
| cons : A -> list A -> list A.
```

```
Inductive ev : nat -> Prop :=
| ev_0 : ev 0
| ev_SS : (n : nat), ev n -> ev (S (S n)).
```

What we have learned



- The definition of programming languages is built upon relations.
- We know the evaluation relation and
- know how to implement it?! (See the next exercise.)