

Type Systems

Lecture 1 Oct. 20th, 2004
Sebastian Maneth

<http://lampwww.epfl.ch/teaching/typeSystems/2004>

Today

1. Organizational Matters
2. What is this course about?
3. Where do “types” come from?
4. Def. of the small language [Expr](#). Its syntax and semantics.
5. Structural Induction on [Expr](#)'s

1. Organizational Matters

Lectures:

We 13:15-15:00, INM203

Sebastian Maneth
BC360, 021-69 31226

Exercises (lab):

We 15:15-17:00, INR 331

Burak Emir
INR320, 021-69 36867

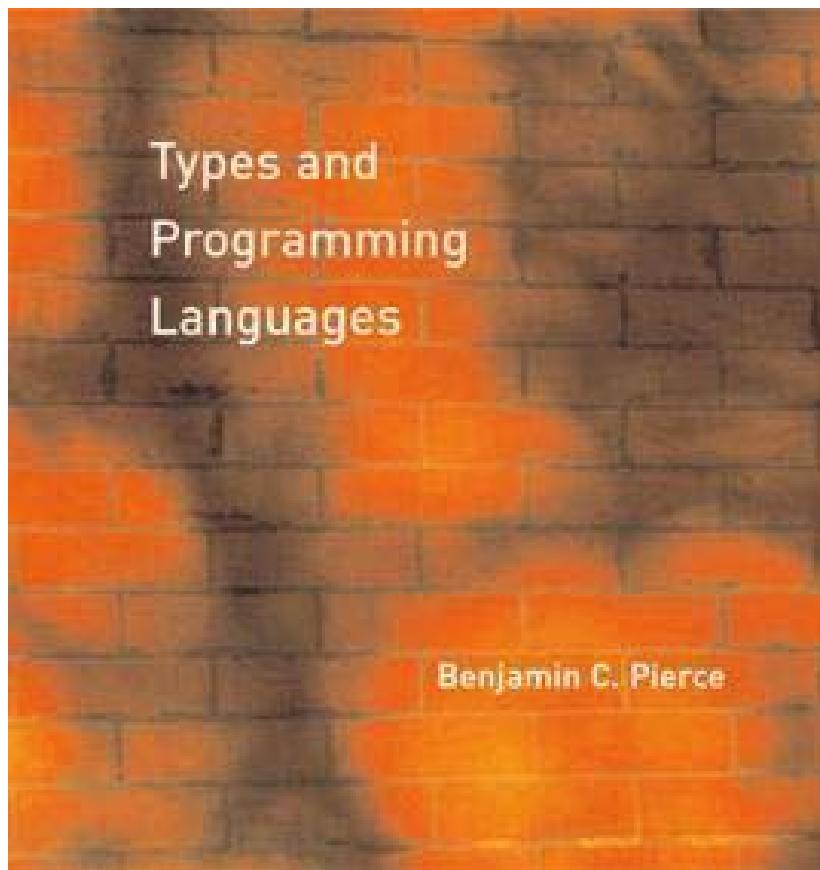
(last 3 lectures by Martin Odersky)

To get credits you have to:

- 1/3 { → 1-2 written assignments
 → one programming assignment
2/3 → oral examination

1. Organizational Matters

Course Book: Benjamin Pierce, “**Types and Programming Languages**”



MIT Press, 2002.

We will strictly follow this book!

So: Good to buy it!

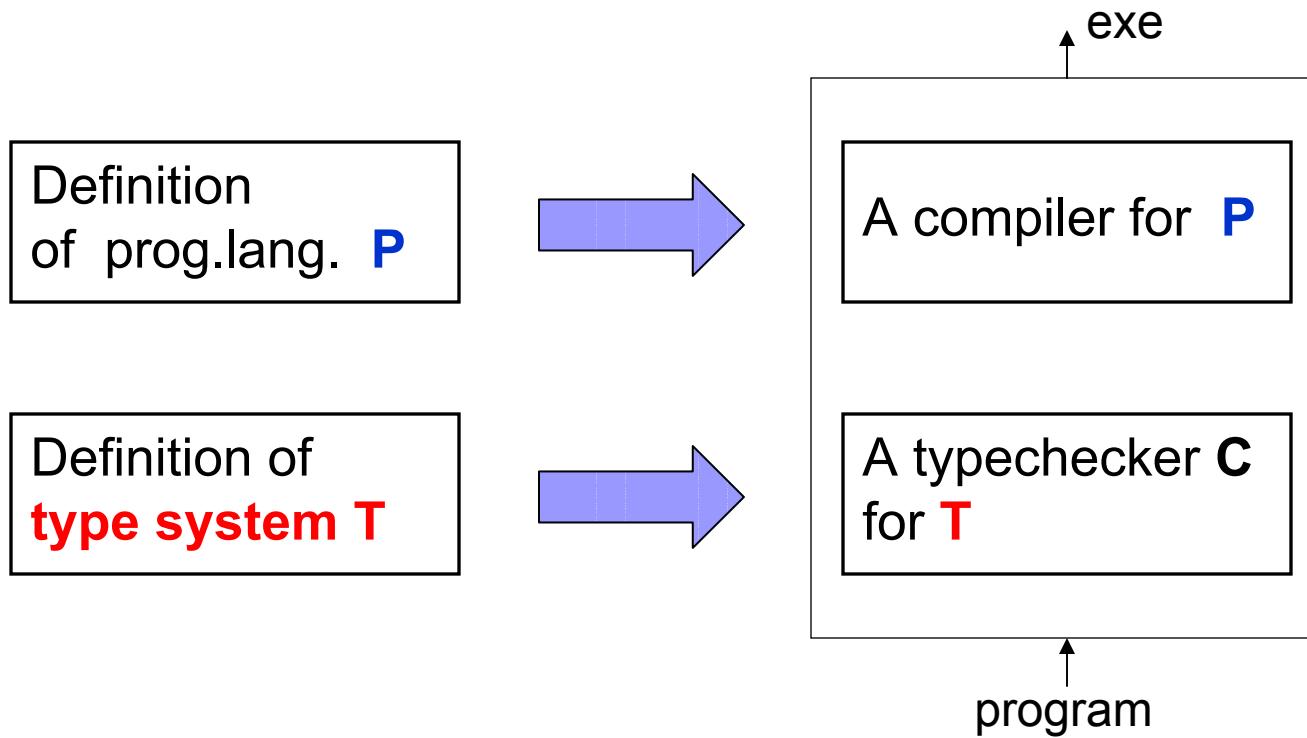
Type Systems for Programming Languages

What for ??

→ to prevent **execution errors**.

A PL in which all well-typed programs are free of execution errors
is called **type sound**.

Type Systems for Programming Languages



- is (P, T) type sound?
- is T decidable?
- does **C** correctly implement T ?

What you will learn in this course:

- how to **define** a type system **T** (to allow for unambiguous implementations)
- how to formally **prove** that (P, T) is type sound
- how to **implement** a typechecker for **T**

Type Systems in Programming Languages

What for ??

→ to prevent **execution errors**.

Execution Errors

trapped

computation stops immediately

untrapped

later causes arbitrary behavior

examples:

- division by zero
- accessing an illegal addr.
- jump to a wrong addr.
- accessing past the end of an array

A program is **SAFE** if it does not have untrapped errors.

A PL is **SAFE** if all its programs are safe.

Execution Errors

trapped

computation stops immediately

untrapped

later causes arbitrary behavior

examples:

- division by zero
- accessing an illegal addr.
- jump to a wrong addr.
- accessing past the end of an array

A program is **SAFE** if it does not have untrapped errors.

A PL is **SAFE** if all its programs are.

trapped + some “forbidden” untrapped errors := well-behaved

What is a TYPE, in our context?

A **type** is an upper bound of the **range of values** that a **program variable** can assume during execution.

e.g. if x has **type Boolean**, then in all runs it should only take one of the values **true / false**.

→ $\text{not}(x)$ has a meaning in every run

PLs in which variables can be given nontrivial types are called **TYPED languages**.

safe/unsafe and typed/untyped

	typed	untyped
safe	ML, Java	LISP
unsafe	C	Assembler

safety \Rightarrow integrity of run-time structures

\Rightarrow enables garbage collection



saves code size / develop. time

(price: performance)

safe/unsafe and typed/untyped

	typed	untyped
safe	ML, Java	LISP
unsafe	C	Assembler

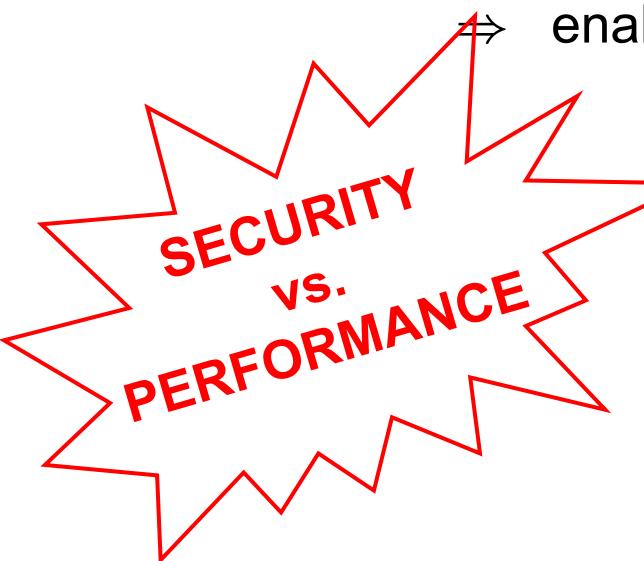
safety \Rightarrow integrity of run-time structures

enables garbage collection



saves code size / develop. time

(price: performance)



```
var x : Boolean
```

```
:
```

```
x := 10;
```



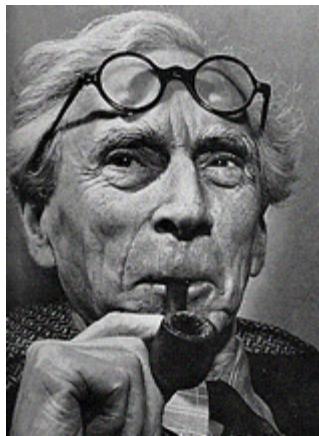
typechecker should
complain!

caveat: of course no one knows if this line will ever be executed!
... but ... it just not SAFE to have it.

should **not** be allowed to write such a program: it has **no meaning!**

TYPE SYSTEMS are there to PROTECT YOU from making
stupid (obvious) mistakes.

Type Theory is much older than PLs!



Bertrand Russell (1872-1970)

1901 Russell's Paradox Let $P = \{ Q \in \text{sets} \mid Q \notin Q \}$

then: $P \in P \Leftrightarrow P \notin P$

- ⇒ Naive set theory is inconsistent!
- ⇒ MUST eliminate self-referential defs.
to make set theory consistent

HOW?

1903 define a **hierarchy of types**: individuals, sets, sets of set, etc.

Any well defined set can only have elements from lower levels.

Course Outline

- today: Intro, Arithm. Expressions, Induction, Evaluation → LAB1
- next: (untyped) Lambda-Calculus → LAB2 untyped λ -evaluator
- 3rd: Simply-Typed Lambda-Calculus → LAB3 simply typed w. let/fix
- 4rd: Simple Extensions, Subtyping → LAB4 subtyping on records
- 5th: Subtyping, Featherweight Java → LAB5
- 6th: Recursive Types I
- 7th: Recursive Types II
- 8th: Polymorphism I
- 9th: Polymorphism II
- 10th: Bounded Quantification
- 11-13th: Scala's Type System (by Martin Odersky)

Syntax and Semantics of PLs

1960 Irons, Syntax-Directed Compiler for ALGOL 60



Syntax and Semantics of PLs

1960 Irons, Syntax-Directed Compiler for ALGOL 60



Defining | Translating

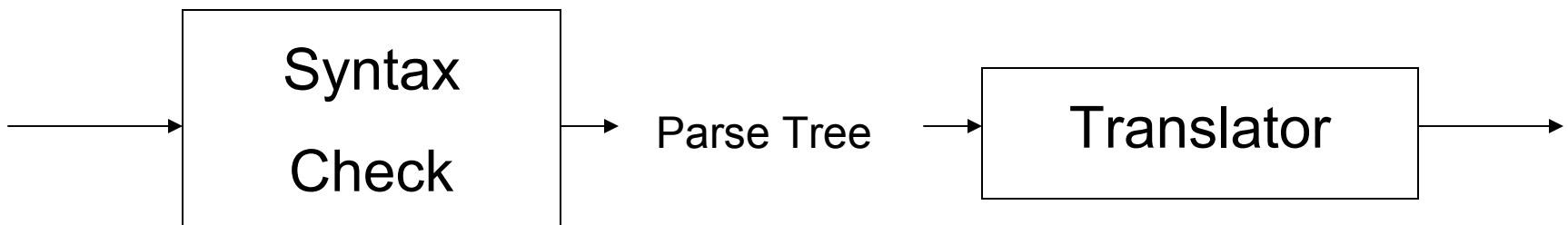
Syntax and Semantics of PLs

1960 Irons, Syntax-Directed Compiler for ALGOL 60



Defining | Translating

1966 Younger, $O(n^3)$ Parsing of **Context-Free Grammars**



Syntax and Semantics of PLs

Until today, **EBNF** (ext. cf. grammar) is used to describe the **syntax of a programming language**.

Example: Arithmetic Expressions

```
Expr ::= true | false | zero  
Expr ::= if Expr then Expr else Expr  
Expr ::= succ Expr  
Expr ::= pred Expr  
Expr ::= isZero Expr
```

Derivable Expressions:

- pred succ zero
- if isZero pred succ zero then zero else true
- if zero then true else false

Syntax and Semantics of PLs

Until today, **EBNF** (ext. cf. grammar) is used to describe the **syntax of a programming language**.

Example: Arithmetic Expressions

```
Expr ::= true | false | zero  
Expr ::= if Expr then Expr else Expr  
Expr ::= succ (Expr)  
Expr ::= pred (Expr)  
Expr ::= isZero (Expr)
```

Derivable Expressions:

- pred (succ (zero))
- if isZero (pred (succ (zero))) then zero else true
- if zero then true else false

Syntax and Semantics of PLs

Until today, **EBNF** (ext. cf. grammar) is used to describe the **syntax of a programming language**.

Example: Arithmetic Expressions

```
Expr ::= true | false | zero  
Expr ::= if Expr then Expr else Expr  
Expr ::= succ (Expr)  
Expr ::= pred (Expr)  
Expr ::= isZero (Expr)
```

Derivable Expressions:

- pred (succ (zero))
- if isZero (pred (succ (zero))) then zero else true
- if zero then true else false

 semantics??

Syntax and Semantics of PLs

Alternative Formalism: Inference Rules

The set of expressions is the smallest set E such that:

$$\text{true} \in E$$

$$\text{false} \in E$$

$$\text{zero} \in E$$

$$\frac{t_1 \in E}{\text{succ } t_1 \in E}$$

$$\frac{t_1 \in E}{\text{pred } t_1 \in E}$$

$$\frac{t_1 \in E}{\text{isZero } t_1 \in E}$$

$$\frac{t_1 \in E \quad t_2 \in E \quad t_3 \in E}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \in E}$$

Syntax and Semantics of PLs

1. **Operational Semantics:** behavior defined in terms of abstract machines
2. **Denotational Semantics:** maps programs by an interpretation function into a collection of semantic domains (such as, e.g., numbers, functions, etc.)
3. **Axiomatic Semantics:** proves properties of a program by applying laws about program behavior (e.g., given that properties P hold before a statement, what properties Q hold after executing it?)

Syntax and **Semantics** of PLs

1. **Operational Semantics:** behavior defined in terms of abstract machines
2. **Denotational Semantics:** maps programs by an interpretation function into a collection of semantic domains (such as, e.b., numbers, functions, etc)
3. **Axiomatic Semantics:** proves properties of a program by applying laws about program behavior (e.g., given that properties P hold before a statement, what properties Q hold after executing it?)

Semantics of Expr

Expr ::= true | false | zero

Expr ::= if Expr then Expr else Expr

Expr ::= succ (Expr)

Expr ::= pred (Expr)

Expr ::= isZero (Expr)

Val ::= true | false | NVal

NVal ::= zero | succ NVal

Evaluation Relation → on Expr's

if true then t_2 else $t_3 \rightarrow t_2$

if false then t_2 else $t_3 \rightarrow t_3$

$$\frac{t_1 \rightarrow t_1'}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \rightarrow \text{if } t_1' \text{ then } t_2 \text{ else } t_3}$$

Semantics of Expr

Expr ::= true | false | zero

Expr ::= if Expr then Expr else Expr

Expr ::= succ (Expr)

Expr ::= pred (Expr)

Expr ::= isZero (Expr)

Val ::= true | false | NVal

NVal ::= zero | succ NVal

Evaluation Relation → on Expr's

if true then t_2 else $t_3 \rightarrow t_2$

if false then t_2 else $t_3 \rightarrow t_3$

$$\frac{t_1 \rightarrow t_1'}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \rightarrow \text{if } t_1' \text{ then } t_2 \text{ else } t_3}$$

$$\frac{t_1 \rightarrow t_1'}{\text{succ } t_1 \rightarrow \text{succ } t_1'}$$

pred zero → zero

$$\frac{t_1 \rightarrow t_1'}{\text{pred } t_1 \rightarrow \text{pred } t_1'}$$

isZero zero → true

pred succ nv₁ → nv₁

isZero succ nv₁ → false

Semantics of Expr

Example: if isZero pred succ pred zero then zero else succ zero

$$\frac{t_1 \rightarrow t_1'}{\text{succ } t_1 \rightarrow \text{succ } t_1'}$$

pred zero \rightarrow zero

pred succ nv₁ \rightarrow nv₁

$$\frac{t_1 \rightarrow t_1'}{\text{pred } t_1 \rightarrow \text{pred } t_1'}$$

isZero zero \rightarrow true

isZero succ nv₁ \rightarrow false

$$\frac{t_1 \rightarrow t_1'}{\begin{array}{c} \text{if } t_1 \text{ then } t_2 \text{ else } t_3 \rightarrow \\ \text{if } t_1' \text{ then } t_2 \text{ else } t_3 \end{array}}$$

$$\frac{t_1 \rightarrow t_1'}{\text{isZero } t_1 \rightarrow \text{isZero } t_1'}$$

Semantics of Expr

redex

Example: if isZero pred succ pred zero then zero else succ zero

$$t_1 \rightarrow t_1'$$

$$\frac{t_1 \rightarrow t_1'}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3} \rightarrow \\ \text{if } t_1' \text{ then } t_2 \text{ else } t_3 \in E$$

$$t_1 \rightarrow t_1'$$

$$\frac{}{\text{succ } t_1 \rightarrow \text{succ } t_1'}$$

$$t_1 \rightarrow t_1'$$

$$\frac{}{\text{pred } t_1 \rightarrow \text{pred } t_1'}$$

$$t_1 \rightarrow t_1'$$

$$\frac{}{\text{isZero } t_1 \rightarrow \text{isZero } t_1'}$$

$$\text{pred zero} \rightarrow \text{zero}$$

$$\text{isZero zero} \rightarrow \text{true}$$

$$\text{pred succ nv}_1 \rightarrow \text{nv}_1$$

$$\text{isZero succ nv}_1 \rightarrow \text{false}$$

Semantics of Expr

redex

Example: if isZero pred succ pred zero then zero else succ zero
→ if isZero pred succ zero then zero else succ zero

$$\frac{t_1 \rightarrow t_1'}{\text{succ } t_1 \rightarrow \text{succ } t_1'}$$

pred zero → zero

pred succ nv₁ → nv₁

$$\frac{t_1 \rightarrow t_1'}{\text{pred } t_1 \rightarrow \text{pred } t_1'}$$

isZero zero → true

isZero succ nv₁ → false

$$\frac{t_1 \rightarrow t_1'}{\begin{array}{c} \text{if } t_1 \text{ then } t_2 \text{ else } t_3 \rightarrow \\ \text{if } t_1' \text{ then } t_2 \text{ else } t_3 \end{array}} \quad E$$

$$\frac{t_1 \rightarrow t_1'}{\text{isZero } t_1 \rightarrow \text{isZero } t_1'}$$

Semantics of Expr

Example: if isZero pred succ pred zero then zero else succ zero
 redex
→ if isZero pred succ zero then zero else succ zero

$$\frac{t_1 \rightarrow t_1'}{\text{succ } t_1 \rightarrow \text{succ } t_1'}$$

pred zero → zero

pred succ nv₁ → nv₁

$$\frac{t_1 \rightarrow t_1'}{\text{pred } t_1 \rightarrow \text{pred } t_1'}$$

isZero zero → true

isZero succ nv₁ → false

$$\frac{\begin{array}{c} t_1 \rightarrow t_1' \\ \hline \text{if } t_1 \text{ then } t_2 \text{ else } t_3 \end{array}}{\text{if } t_1' \text{ then } t_2 \text{ else } t_3 \rightarrow E}$$
$$\frac{t_1 \rightarrow t_1'}{\text{isZero } t_1 \rightarrow \text{isZero } t_1'}$$

Semantics of Expr

Example: if isZero pred succ pred zero then zero else succ zero
 redex
→ if isZero pred succ zero then zero else succ zero
→ if isZero zero then zero else succ zero

$$\frac{t_1 \rightarrow t_1'}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \rightarrow \text{if } t_1' \text{ then } t_2 \text{ else } t_3} \quad E$$
$$\frac{t_1 \rightarrow t_1'}{\text{pred } t_1 \rightarrow \text{pred } t_1'}$$
$$\frac{t_1 \rightarrow t_1'}{\text{isZero } t_1 \rightarrow \text{isZero } t_1'}$$

pred zero → zero

isZero zero → true

pred succ nv₁ → nv₁

isZero succ nv₁ → false

Semantics of Expr

Example: if isZero pred succ pred zero then zero else succ zero

→ if isZero pred succ zero then zero else succ zero
redex
→ if isZero zero then zero else succ zero

$$\frac{t_1 \rightarrow t_1'}{\text{succ } t_1 \rightarrow \text{succ } t_1'}$$

pred zero → zero

pred succ nv₁ → nv₁

$$\frac{t_1 \rightarrow t_1'}{\text{pred } t_1 \rightarrow \text{pred } t_1'}$$

isZero zero → true

isZero succ nv₁ → false

$$\frac{t_1 \rightarrow t_1'}{\text{if } t_1 \text{ then } t_2 \text{ else } t_3 \rightarrow \text{if } t_1' \text{ then } t_2 \text{ else } t_3 E}$$

$$\frac{t_1 \rightarrow t_1'}{\text{isZero } t_1 \rightarrow \text{isZero } t_1'}$$

Semantics of Expr

Example: if isZero pred succ pred zero then zero else succ zero

→ if isZero pred succ zero then zero else succ zero
redex

→ if isZero zero then zero else succ zero

→ if true then zero else succ zero

$$\frac{t_1 \rightarrow t_1'}{\text{succ } t_1 \rightarrow \text{succ } t_1'}$$

pred zero → zero

pred succ nv₁ → nv₁

$$\frac{t_1 \rightarrow t_1'}{\text{pred } t_1 \rightarrow \text{pred } t_1'}$$

isZero zero → true

isZero succ nv₁ → false

$$\frac{t_1 \rightarrow t_1'}{\text{isZero } t_1 \rightarrow \text{isZero } t_1'}$$

Semantics of Expr

Example: if isZero pred succ pred zero then zero else succ zero

→ if isZero pred succ zero then zero else succ zero

→ if isZero zero then zero else succ zero

→ if true then zero else succ zero

redex

$$\frac{t_1 \rightarrow t_1'}{\text{succ } t_1 \rightarrow \text{succ } t_1'}$$

pred zero → zero

$$\frac{t_1 \rightarrow t_1'}{\text{pred } t_1 \rightarrow \text{pred } t_1'}$$

if true then t_2 else $t_3 \rightarrow t_2$

$$\frac{t_1 \rightarrow t_1'}{\text{isZero } t_1 \rightarrow \text{isZero } t_1'}$$

pred succ $\text{nv}_1 \rightarrow \text{nv}_1$

isZero succ $\text{nv}_1 \rightarrow \text{false}$

isZero zero → true

Semantics of Expr

Example: if isZero pred succ [pred zero] then zero else succ zero

→ if isZero [pred succ zero] then zero else succ zero

→ if isZero zero [then zero else succ zero]

→ ^{redex} if true [then zero else succ zero]

→ zero

$$\frac{t_1 \rightarrow t_1'}{\text{succ } t_1 \rightarrow \text{succ } t_1'}$$

pred zero → zero

$$\frac{t_1 \rightarrow t_1'}{\text{pred } t_1 \rightarrow \text{pred } t_1'}$$

if true then t_2 else $t_3 \rightarrow t_2$

$$\frac{t_1 \rightarrow t_1'}{\text{isZero } t_1 \rightarrow \text{isZero } t_1'}$$

pred succ nv₁ → nv₁

isZero succ nv₁ → false isZero zero → true

Induction on the Structure of Expr's

The set of expressions is the smallest set E such that:

1. $\text{true}, \text{false}, \text{zero} \in E$
2. $\underline{\text{if } t_1, t_2, t_3 \in E, \text{ then succ } t_1, \text{pred } t_1, \text{isZero } t_1 \in E}$
 $\underline{\text{and if } t_1 \text{ then } t_2 \text{ else } t_3 \in E}$

inductive definition

→ we can define / proof things about Expr's by induction!

Example: for any Expr t define its **size** as

1. $\underline{\text{if } t = \text{true} \mid \text{false} \mid \text{zero} \text{ then } \text{size}(t) = 0}$
2. $\underline{\text{if } t = \text{succ } t_1 \mid \text{pred } t_1 \mid \text{isZero } t_1 \text{ then } \text{size}(t) = \text{size}(t_1) + 1}$
 $\underline{\text{if } t = \text{if } t_1 \text{ then } t_2 \text{ else } t_3 \text{ then } \text{size}(t) = \text{size}(t_1) + \text{size}(t_2) + \text{size}(t_3) + 1}$

Proof by Induction on the Structure of Expr's

Theorem. \rightarrow is deterministic: if $t \rightarrow t'$ and $t \rightarrow t''$ then $t' = t''$

Proof. by induction on the structure of t

1. if $t = \text{true} \mid \text{false} \mid \text{zero}$ then $t' = t'' = t$
2. if $t = \text{succ } t_1$ then

$$\frac{t_1 \rightarrow t_1'}{\text{succ } t_1 \rightarrow \text{succ } t_1'},$$

only rule for $\text{succ}(\dots)$

Proof by Induction on the Structure of Expr's

Theorem. \rightarrow is deterministic: if $t \rightarrow t'$ and $t \rightarrow t''$ then $t' = t''$

Proof. by induction on the structure of t

1. if $t = \text{true} \mid \text{false} \mid \text{zero}$ then $t' = t'' = t$
2. if $t = \text{succ } t_1$ then $t' = \text{succ } t'_1$ and $t'' = \text{succ } t''_1$
for t'_1, t''_1 with $t_1 \rightarrow t'_1$ and $t_1 \rightarrow t''_1$

Proof by Induction on the Structure of Expr's

Theorem. \rightarrow is deterministic: if $t \rightarrow t'$ and $t \rightarrow t''$ then $t' = t''$

Proof. by induction on the structure of t

1. if $t = \text{true} \mid \text{false} \mid \text{zero}$ then $t' = t'' = t$
2. if $t = \text{succ } t_1$ then $t' = \text{succ } t'_1$ and $t'' = \text{succ } t''_1$
for t'_1, t''_1 with $\underbrace{t_1 \rightarrow t'_1 \text{ and } t_1 \rightarrow t''_1}$
by induction $t'_1 = t''_1$

Proof by Induction on the Structure of Expr's

Theorem. \rightarrow is deterministic: if $t \rightarrow t'$ and $t \rightarrow t''$ then $t' = t''$

Proof. by induction on the structure of t

1. if $t = \text{true} \mid \text{false} \mid \text{zero}$ then $t' = t'' = t$
2. if $t = \text{succ } t_1$ then $t' = \text{succ } t'_1$ and $t'' = \text{succ } t''_1$
for t'_1, t''_1 with $\underbrace{t_1 \rightarrow t'_1 \text{ and } t_1 \rightarrow t''_1}$
by induction $t'_1 = t''_1$

Thus, also $t' = t''$.

Proof by Induction on the Structure of Expr's

Theorem. \rightarrow is deterministic: if $t \rightarrow t'$ and $t \rightarrow t''$ then $t' = t''$

Proof. by induction on the structure of t

1. if $t = \text{true} \mid \text{false} \mid \text{zero}$ then $t' = t'' = t$
2. if $t = \text{pred } t_1$ then

if $t_1 = \text{succ } t_{11}$ then $t' = t'' = t_{11}$

because $\boxed{\text{pred succ nv}_1 \rightarrow \text{nv}_1}$ is **only** rule applicable.

Proof by Induction on the Structure of Expr's

Theorem. \rightarrow is deterministic: if $t \rightarrow t'$ and $t \rightarrow t''$ then $t' = t''$

Proof. by induction on the structure of t

1. if $t = \text{true} | \text{false} | \text{zero}$ then $t' = t'' = t$

2. if $t = \text{pred } t_1$ then

if $t_1 = \text{succ } t_{11}$ then $t' = t'' = t_{11}$

because $\boxed{\text{pred succ nv}_1 \rightarrow \text{nv}_1}$ is **only** rule applicable.

otherwise $t' = \text{pred } t'_1$ and $t'' = \text{pred } t''_1$
with $t_1 \rightarrow t'_1$ and $t_1 \rightarrow t''_1$

Proof by Induction on the Structure of Expr's

Theorem. \rightarrow is deterministic: if $t \rightarrow t'$ and $t \rightarrow t''$ then $t' = t''$

Proof. by induction on the structure of t

1. if $t = \text{true} | \text{false} | \text{zero}$ then $t' = t'' = t$

2. if $t = \text{pred } t_1$ then

if $t_1 = \text{succ } t_{11}$ then $t' = t'' = t_{11}$

because $\boxed{\text{pred succ nv}_1 \rightarrow \text{nv}_1}$ is **only** rule applicable.

otherwise $t' = \text{pred } t'_1$ and $t'' = \text{pred } t''_1$

with $t_1 \xrightarrow{} t'_1$ and $t_1 \xrightarrow{} t''_1$

by induction $t'_1 = t''_1$

Thus, also $t' = t''$.

Proof by Induction on the Structure of Expr's

Theorem. \rightarrow is deterministic: if $t \rightarrow t'$ and $t \rightarrow t''$ then $t' = t''$

Proof. by induction on the structure of t

1. if $t = \text{true} \mid \text{false} \mid \text{zero}$ then $t' = t'' = t$
2. if $t = \text{if } t_1 \text{ then } t_2 \text{ else } t_3$ then
if $t_1 = \text{true}$ then $t' = t'' = t_2$
if $t_1 = \text{false}$ then $t' = t'' = t_3$

Proof by Induction on the Structure of Expr's

Theorem. \rightarrow is deterministic: if $t \rightarrow t'$ and $t \rightarrow t''$ then $t' = t''$

Proof. by induction on the structure of t

1. if $t = \text{true} \mid \text{false} \mid \text{zero}$ then $t' = t'' = t$
 2. if $t = \text{if } t_1 \text{ then } t_2 \text{ else } t_3$ then

if $t_1 = \text{true}$ **then** $t' = t'' = t_2$

if $t_1 = \text{false}$ **then** $t' = t'' = t_3$

otherwise $t' = \text{if } t_1 \text{ then } t_2 \text{ else } t_3$ and
 $t'' = \text{if } t_1'' \text{ then } t_2 \text{ else } t_3$

with $t_1 \rightarrow t_1'$ and $t_1 \rightarrow t_1''$

by induction $t_1' = t_1''$

Thus, also $t' = t''$.

Questions:

1. Is \rightarrow still deterministic if we add the new rule

succ pred nv₁ \rightarrow nv₁

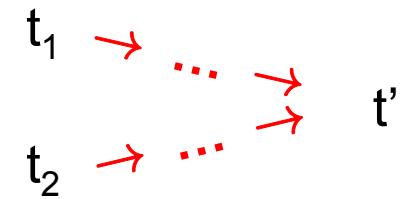
Which rule must be removed now, to keep a sane semantics?

2. What if redexes can be chosen freely? Is \rightarrow still determin.?

(i.e., rules can be applied to arbitrary sub-Expr's)

Is \rightarrow confluent? Is it terminating?

if $t \xrightarrow{} t_1$ $t \xrightarrow{} t_2$ then there is a t' such that



Summary

- we have defined the **syntax** of the small language called **Expr**.
- we have given a **semantics** to **Expr's** by means of an evaluation relation.
- we have proved by **induction** that for every **Expr** there is at most one other **Expr** that can be derived by the evaluation relation.

Next Lecture

How to define a small language for defining **functions**?

- function definition and application: **the lambda-calculus**