Arrays

Using array as an expression, on the right-hand side

$$\frac{\Gamma \vdash \mathrm{a}\colon\ \mathrm{Array(T)} \qquad \Gamma \vdash \mathrm{i}\colon\ \mathrm{Int}}{\Gamma \vdash \mathrm{a[i]:T}}$$

Assigning to an array

$$\frac{\Gamma \vdash \mathrm{a}\colon\ \mathrm{Array(T)} \qquad \Gamma \vdash \mathrm{i}\colon\ \mathrm{Int} \qquad \Gamma \vdash \mathrm{e}\colon\ \mathrm{T}}{\Gamma \vdash (\mathrm{a[i]}\ = \mathrm{e})\colon\ \mathrm{void}}$$

Example with Arrays

```
def next(a : Array[Int], k : Int) : Int = {
    a[k] = a[a[k]]
}
```

Given Γ = {(a, Array(Int)), (k, Int)}, check $\Gamma \vdash$ a[k] = a[a[k]]: void

$$\dfrac{\dfrac{\Gamma \vdash \mathrm{a}\!:\ \mathrm{Array(Int)} \qquad \dfrac{\Gamma \vdash \mathrm{a}\!:\ \mathrm{Array(Int)} \qquad \Gamma \vdash \mathrm{k}\!:\ \mathrm{Int}}{\Gamma \vdash \mathrm{a[k]}\!:\ \mathrm{Int}}}{\Gamma \vdash \mathrm{a[a[k]]}\!:\!\mathrm{Int}} \qquad \Gamma \vdash \mathrm{a}\!:\ \mathrm{Array(Int)} \qquad \Gamma \vdash \mathrm{k}\!:\!\mathrm{Int}}{\Gamma \vdash \mathrm{a[k]} = \mathrm{a[a[k]]}\!:\ \text{void}}$$

Type Rules (1)

$$\frac{(\mathrm{x}:\ \mathrm{T}) \in \Gamma}{\Gamma \vdash \mathrm{x}:\ \mathrm{T}}$$ variable

$$\frac{}{\mathrm{IntConst(k)}:\ \mathrm{Int}}$$ constant

$$\frac{\Gamma \vdash e_1 : T_1 \ \ \ldots \ \ \Gamma \vdash e_n : T_n \qquad \Gamma \vdash f : (T_1 \times \cdots \times T_n \to T)}{\Gamma \vdash f(e_1, \ldots, e_n) : T}$$ function application

$$\frac{\Gamma \vdash e_1:\ \mathrm{Int} \qquad \Gamma \vdash e_2:\ \mathrm{Int}}{\Gamma \vdash (e_1 + e_2):\ \mathrm{Int}}$$ plus $$\frac{\Gamma \vdash e_1:\ \mathrm{String} \qquad \Gamma \vdash e_2:\ \mathrm{String}}{\Gamma \vdash (e_1 + e_2):\ \mathrm{String}}$$

$$\frac{\Gamma \vdash \mathrm{b}:\ \mathrm{Boolean} \qquad \Gamma \vdash e_1 :\ \mathrm{T} \qquad \Gamma \vdash e_2 :\ \mathrm{T}}{\Gamma \vdash (\mathrm{if(b)}\ e_1\ \mathrm{else}\ e_2) :\ \mathrm{T}}$$ if

$$\frac{\Gamma \vdash \mathrm{b}:\ \mathrm{Boolean} \qquad \Gamma \vdash \mathrm{s}:\ \mathrm{void}}{\Gamma \vdash (\mathrm{while(b)\ s}):\ \mathrm{void}}$$ while

$$\frac{(\mathrm{x},\ \mathrm{T}) \in \Gamma \qquad \Gamma \vdash \mathrm{e}:\ \mathrm{T}}{\Gamma \vdash (\mathrm{x{=}e}):\ \mathrm{void}}$$ assignment

Type Rules (2)

$$\frac{\Gamma \vdash \mathrm{e}:\ \mathrm{T}}{\Gamma \vdash \{\mathrm{e}\}:\ \mathrm{T}} \qquad \frac{}{\Gamma \vdash \{\}:\ \mathrm{void}}$$

$$\frac{\Gamma \oplus \{(x, T_1)\} \vdash \{t_2;\ \ldots\ ; t_n\}:\ \mathrm{T}}{\Gamma \vdash \{\mathrm{var}\ x : T_1; t_2;\ \ldots\ ; t_n\}:\ \mathrm{T}}$$

$$\frac{\Gamma \vdash s_1:\ \mathrm{void} \qquad \Gamma \vdash \{t_2;\ \ldots\ ; t_n\}:\ \mathrm{T}}{\Gamma \vdash \{s_1; t_2;\ \ldots\ ; t_n\}:\ \mathrm{T}}$$

block

$$\frac{\Gamma \vdash \mathrm{a}:\ \mathrm{Array}(\mathrm{T}) \qquad \Gamma \vdash \mathrm{i}:\ \mathrm{Int}}{\Gamma \vdash \mathrm{a}[\mathrm{i}]:\ \mathrm{T}}$$

array use

$$\frac{\Gamma \vdash \mathrm{a}:\ \mathrm{Array}(\mathrm{T}) \qquad \Gamma \vdash \mathrm{i}:\ \mathrm{Int} \qquad \Gamma \vdash \mathrm{e}:\ \mathrm{T}}{\Gamma \vdash \mathrm{a}[\mathrm{i}] = \mathrm{e}}$$

array
assignment

Type Rules (3)

Γ^c - top-level environment of class C

```
class C {
  var x: Int;
  def m(p: Int): Boolean = {…}
}
```

$\downarrow$

Γ^c = { (x, Int), (m, C x Int $\rightarrow$ Boolean)}

$$\frac{\Gamma \vdash e : C \quad \Gamma^C \vdash m : \mathsf{C} \times T_1 \times \ \ldots \ \times T_n \rightarrow T_{n+1} \qquad \Gamma \vdash e_i : T_i \quad 1 \leq i \leq n}{\Gamma \vdash e.m(e_1, \ \ldots \ , e_n) : T_{n+1}}$$

method invocation

$$\frac{\Gamma \vdash \mathrm{e:\ C} \qquad \Gamma^C \vdash \mathrm{f:\ T}}{\Gamma \vdash \mathrm{e.f:\ T}}$$

field use

$$\frac{\Gamma \vdash \mathrm{e:\ C} \qquad \Gamma^C \vdash \mathrm{f:\ T} \qquad \Gamma \vdash \mathrm{x:\ T}}{\Gamma \vdash \mathrm{(e.f = x):\ void}}$$

field assignment

Does this program type check?

```
class Rectangle {
  var width: Int
  var height: Int
  var xPos: Int
  var yPos: Int
  def area(): Int = {
    if (width > 0 && height > 0)
     width * height
   else 0
  }
  def resize(maxSize: Int) {
    while (area > maxSize) {
      width = width / 2
      height = height / 2
    }
  }
}
```

$$\Gamma_0 = \left\{ \begin{array}{c} \text{w: Int, h: Int,} \\ \text{x: Int, y: Int,} \\ area : \text{Unit} \rightarrow \text{Int,} \\ resize : \text{Int} \rightarrow \text{Unit} \end{array} \right\}$$

Type check: area

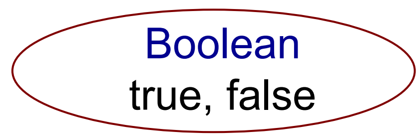
Type check: resize

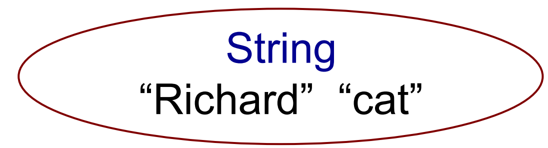
Semantics of Types

- Operational view: Types are named entities
  - such as the primitive types (Int, Bool etc.) and explicitly declared classes, traits …
  - their meaning is given by methods they have
  - constructs such as inheritance establish relationships between classes
- Mathematically, Types are sets of values
  - Int = { ..., -2, -1, 0, 1, 2, ... }
  - Boolean = { false, true }
  - Int → Int = { f : Int -> Int | f is computable }

Types as Sets

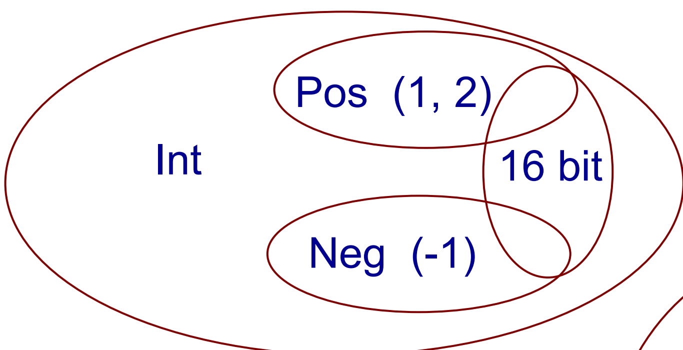
- Sets so far were disjoint

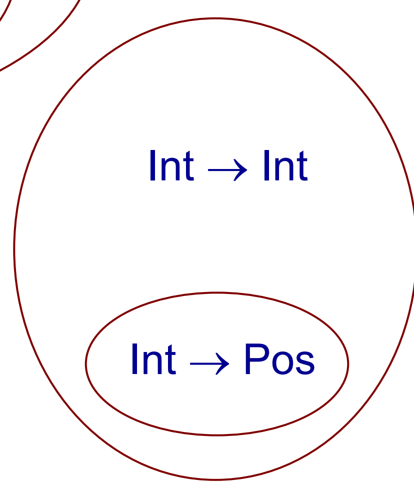


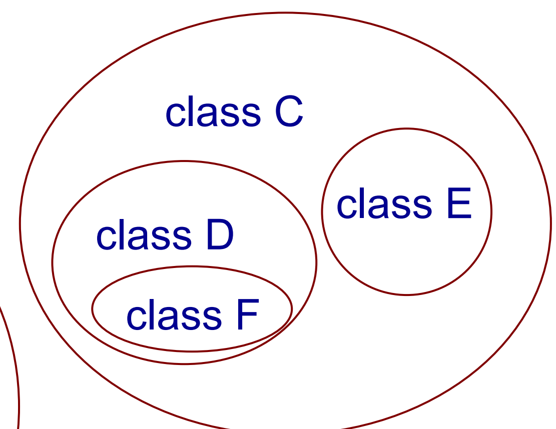


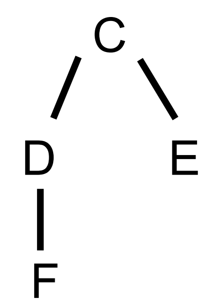
- Sets can overlap

C represents not only declared C, but all possible extensions as well









F extends D,
D extends C

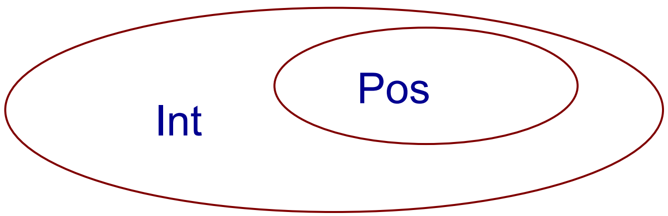
SUBTYPING

Subtyping

- Subtyping corresponds to subset
- Systems with subtyping have non-disjoint sets
- $T_1 <: T_2$ means T_1 is a subtype of T_2
  - corresponds to $T_1 \subseteq T_2$ in sets of values
- Rule for subtyping: analogous to set reasoning

In terms of sets

$$\frac{\Gamma \vdash e : T_1 \qquad T_1 <: T_2}{\Gamma \vdash e : T_2} \qquad\qquad \frac{e \in T_1 \qquad T_1 \subseteq T_2}{e \in T_2}$$



Types for Positive and Negative Ints

Int = { ... , -2, -1, 0, 1, 2, ... }

Pos = { 1, 2, ... } (not including zero)

Neg = { ..., -2, -1 } (not including zero)

types:

Pos <: Int
Neg <: Int

$$\frac{\Gamma \vdash \mathrm{x}\colon \mathrm{Pos} \qquad \Gamma \vdash \mathrm{y}\colon \mathrm{Pos}}{\Gamma \vdash \mathrm{x} + \mathrm{y}\colon \mathrm{Pos}}$$

$$\frac{\Gamma \vdash \mathrm{x}\colon \mathrm{Pos} \qquad \Gamma \vdash \mathrm{y}\colon \mathrm{Neg}}{\Gamma \vdash \mathrm{x} * \mathrm{y}\colon \mathrm{Neg}}$$

$$\frac{\Gamma \vdash \mathrm{x}\colon \mathrm{Pos} \qquad \Gamma \vdash \mathrm{y}\colon \mathrm{Pos}}{\Gamma \vdash \mathrm{x} \,/\, \mathrm{y}\colon \mathrm{Pos}}$$

sets:

Pos $\subseteq$ Int
Neg $\subseteq$ Int

$$\frac{\mathrm{x} \in \mathrm{Pos} \qquad \mathrm{y} \in \mathrm{Pos}}{\mathrm{x} + \mathrm{y} \in \mathrm{Pos}}$$

$$\frac{\mathrm{x} \in \mathrm{Pos} \qquad \mathrm{y} \in \mathrm{Neg}}{\mathrm{x} * \mathrm{y} \in \mathrm{Neg}}$$

(y not zero)

$$\frac{\mathrm{x} \in \mathrm{Pos} \qquad \mathrm{y} \in \mathrm{Pos}}{\mathrm{x} \,/\, \mathrm{y} \in \mathrm{Pos}}$$

(x/y well defined)

Rules for Neg, Pos, Int

$$\frac{\Gamma \vdash x: \mathrm{Pos} \quad \Gamma \vdash y: \mathrm{Neg}}{\Gamma \vdash x + y: ???}$$

$$\frac{\Gamma \vdash x: \mathrm{Pos} \quad \Gamma \vdash y: \mathrm{Neg}}{\Gamma \vdash x * y: ???}$$

$$\frac{\Gamma \vdash x: \mathrm{Pos} \quad \Gamma \vdash y: \mathrm{Int}}{\Gamma \vdash x + y: ???}$$

$$\frac{\Gamma \vdash x: \mathrm{Pos} \quad \Gamma \vdash y: \mathrm{Int}}{\Gamma \vdash x * y: ???}$$

More Rules

$$\frac{\Gamma \vdash \text{x: Neg} \qquad \Gamma \vdash \text{y: Neg}}{\Gamma \vdash \text{x * y: Pos}}$$

$$\frac{\Gamma \vdash \text{x: Neg} \qquad \Gamma \vdash \text{y: Neg}}{\Gamma \vdash \text{x + y: Neg}}$$

More rules for division?

$$\frac{\Gamma \vdash \text{x: Neg} \qquad \Gamma \vdash \text{y: Neg}}{\Gamma \vdash \text{x / y: Pos}}$$

$$\frac{\Gamma \vdash \text{x: Pos} \qquad \Gamma \vdash \text{y: Neg}}{\Gamma \vdash \text{x / y: Neg}}$$

$$\frac{\Gamma \vdash \text{x: Int} \qquad \Gamma \vdash \text{y: Neg}}{\Gamma \vdash \text{x / y: Int}}$$

Making Rules Useful

- Let x be a variable

$$\frac{\Gamma \vdash \text{x: Int} \qquad \Gamma \oplus \{(x, Pos)\} \vdash e_1 : T \qquad \Gamma \vdash e_2 : T}{\Gamma \vdash (\text{if } (\text{x} > 0)\ e_1 \text{ else } e_2)\text{: T}}$$

$$\frac{\Gamma \vdash \text{x: Int} \qquad \Gamma \vdash e_1 : T \qquad \Gamma \oplus \{(x, Neg)\} \vdash e_2 : T}{\Gamma \vdash (\text{if } (\text{x} >= 0)\ e_1 \text{ else } e_2)\text{: T}}$$

```
var x : Int
var y : Int
if (y > 0) {
  if (x > 0) {
    var z : Pos = x * y
    res = 10 / z
} }
```

type system proves: no division by zero

Subtyping Example

```
def f(x:Int) : Pos = {
  if (x < 0) -x else x+1
}
var p : Pos
var q : Int
q = f(p)
```

Does this statement type check?

Given:

$$\mathrm{Pos} <: \mathrm{Int}$$

$$\Gamma \vdash \mathrm{f}: \mathrm{Int} \to \mathrm{Pos}$$

$$\dfrac{(\mathrm{q},\ \mathrm{Int}) \in \Gamma \qquad \dfrac{\dfrac{\dfrac{\mathrm{p}:\ \mathrm{Pos} \qquad \mathrm{Pos} <:\ \mathrm{Int}}{\mathrm{p}:\ \mathrm{Int}} \qquad \mathrm{f}:\ \mathrm{Int} \to \mathrm{Pos}}{\mathrm{f(p)}:\ \mathrm{Pos}} \qquad \mathrm{Pos} <:\ \mathrm{Int}}{\mathrm{f(p)}:\ \mathrm{Int}}}{\mathrm{q{=}f(p)}:\ \mathrm{void}}$$

Subtyping Example

```
def f(x:Pos) : Pos = {
  if (x < 0) -x else x+1
}
var p : Int
var q : Int
q = f(p)
```

Does this statement type check?

does not type check

What Pos/Neg Types Can Do

```
def multiplyFractions(p1 : Int, q1 : Pos, p2 : Int, q2 : Pos) : (Int,Pos) {
  (p1*q1, q1*q2)
}
def addFractions(p1 : Int, q1 : Pos, p2 : Int, q2 : Pos) : (Int,Pos) {
  (p1*q2 + p2*q1, q1*q2)
}
def printApproxValue(p : Int, q : Pos) = {
  print(p/q)   // no division by zero
}
```

More sophisticated types can track intervals of numbers and ensure that a program does not crash with an array out of bounds error.

Subtyping and Product Types

Subtyping for Products

T_1 <: T_2 implies for all e:

$$\frac{\Gamma \vdash e : T_1}{\Gamma \vdash e : T_2}$$

Type for a tuple:

$$\frac{x : T_1 \qquad y : T_2}{(x, y) : T_1 \times T_2}$$

$$\frac{\dfrac{x : T_1 \qquad T_1 <: T_1'}{x : T_1'} \qquad \dfrac{y : T_2 \qquad T_2 <: T_2'}{y : T_2'}}{(x, y) : T_1' \times T_2'}$$

So, we might as well add:

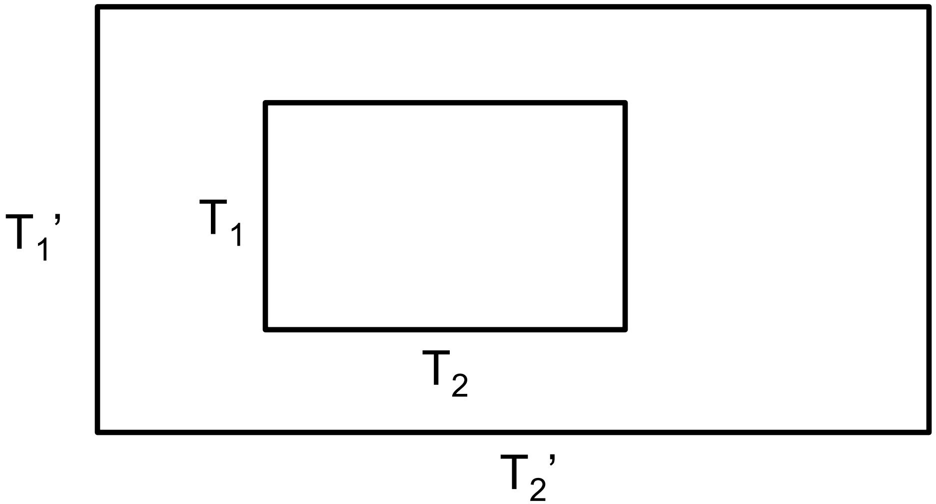
$$\frac{T_1 <: T_1' \qquad T_2 <: T_2'}{T_1 \times T_2 <: T_1' \times T_2'}$$

covariant subtyping for pair types denoted (T_1, T_2) or Pair[T_1, T_2]

Analogy with Cartesian Product

$$\frac{T_1 <: T_1' \qquad T_2 <: T_2'}{T_1 \times T_2 <: T_1' \times T_2'}$$

$$\frac{T_1 \subseteq T_1' \qquad T_2 \subseteq T_2'}{T_1 \times T_2 \subseteq T_1' \times T_2'}$$



A x B = { (a, b) | a $\in$ A, b $\in$ B}

Subtyping and Function Types

Subtyping for Function Types

$T_1 <: T_2$ implies for all e:

$$\frac{\Gamma \vdash e : T_1}{\Gamma \vdash e : T_2}$$

$$\frac{\overbrace{T_1' <: T_1 \quad \dots \quad T_n' <: T_n}^{\text{contravariance}} \qquad \overbrace{T <: T'}^{\text{covariance}}}{(T_1 \times \ \dots \ \times T_n \to T) <: (T_1' \times \ \dots \ \times T_n' \to T')}$$

Consequence:

$$\dfrac{\dfrac{\Gamma \vdash m : T_1 \times \ \dots \ \times T_n \to T \qquad \dfrac{\Gamma \vdash e_1 : T_1' \quad T_1' <: T_1}{\Gamma \vdash e_1 : T_1} \qquad \dfrac{\Gamma \vdash e_n : T_n' \quad T_n' <: T_n}{\Gamma \vdash e_n : T_n}}{\Gamma \vdash m(e_1, \ \dots \ , e_n) : T} \qquad T <: T'}{\Gamma \vdash m(e_1, \ \dots \ , e_n) : T'}$$

as if Γ |- m: $T'_1 \times \dots \times T_n' \to T'$

Function Space as Set

A function type is a set of functions (function space) defined as follows:

$T_1 \to T_2 = \{ f \mid \forall x. (x \in T_1 \to f(x) \in T_2)\}$

contravariance because
$x \in T_1$ is left of implication

We can prove

$$\frac{T_1' \subseteq T_1 \qquad T_2 \subseteq T_2'}{T_1 \to T_2 \subseteq T_1' \to T_2'}$$

Proof

$T_1 \rightarrow T_2 = \{ f \mid \forall x. (x \in T_1 \rightarrow f(x) \in T_2)\}$

$$\frac{T_1' \subseteq T_1 \qquad T_2 \subseteq T_2'}{T_1 \rightarrow T_2 \subseteq T_1' \rightarrow T_2'}$$

Subtyping for Classes

- Class C contains a collection of methods

- For class sub-typing, we require that methods named the same are subtypes

Example

class C {
  def m(x : T_1) : T_2 = {...}
}
class D **extends** C {
  override def m(x : T'_1) : T'_2 = {...}
}

D <: C so need to have $(T'_1 \rightarrow T'_2) <: (T_1 \rightarrow T_2)$

Therefore, we need to have:

$T'_2 <: T_2$ (result behaves like the class)

$T_1 <: T'_1$ (**argument behaves opposite**)

Mutable and Immutable Fields

- We view field var f: T as two methods
  - getF : T   T
  - setF(x:T): void $T \rightarrow void$
- For val f: T (immutable): we have only getF

Could we allow this?

```
class A {}        class B extends A {…}          B <: A
class C {
  val x : A = …
}
class D extends C {
  override val x : B = …
}
```

Because B <: A, this is a valid way for D to extend C (D <: C)

Substitution principle:

If someone uses z:D thinking it is z:C, the fact that they read z.x and obtain B as a specific kind of A is not a problem.

What if x is a **var** ?

```
class A {}        class B extends A {...}        B <: A

class C {
  var x : A = ...
}
class D extends C {
  override var x : B = ...                       ?!?
}
```

If we now imagine the setter method (i.e. field assignment), in the first case the setter has type (A -> void) and in the second (B -> void). By contravariance (A -> void) <: (B -> void) so we cannot have D <: C

Soundness of Types

ensuring that a type system
is not broken

Example: *Tootool 0.1* Language

Tootool is a rural community in the central east part of the Riverina [New South Wales, Australia]. It is situated by road, about 4 kilometres east from French Park and 16 kilometres west from The Rock.
Tootool Post Office opened on 1 August 1901 and closed in 1966. [Wikipedia]

unsound

Type System for *Tootool 0.1*

Pos <: Int

Neg <: Int

$$\frac{\Gamma \vdash x:\ T \qquad \Gamma \vdash e:\ T}{\Gamma \vdash (x = e):\ \text{void}}$$

assignment

$$\frac{\Gamma \vdash e:\ T \qquad \Gamma \vdash T <:\ T'}{\Gamma \vdash e:\ T'}$$

subtyping

does it type check?

```
def intSqrt(x:Pos) : Pos = { ...}
var p : Pos
var q : Neg
var r : Pos
q = -5
p = q
r = intSqrt(p)
```

Γ = {(p, Pos), (q, Neg), (r, Pos), (intSqrt, Pos $\to$ Pos)}

Runtime error: intSqrt invoked with a negative argument!

$$\frac{\dfrac{p:\ \text{Pos} \qquad \text{Pos} <:\ \text{Int}}{p:\ \text{Int}} \qquad \dfrac{q:\ \text{Neg} \qquad \text{Neg} <:\ \text{Int}}{q:\ \text{Int}}}{(p{=}q):\ \text{void}}$$

What went wrong in *Tootool 0.1* ?

Pos <: Int

Neg <: Int

$$\frac{\Gamma \vdash x:\ T \qquad \Gamma \vdash e:\ T}{\Gamma \vdash (x = e):\ \text{void}}$$ assignment

$$\frac{\Gamma \vdash e:\ T \qquad \Gamma \vdash T <:\ T'}{\Gamma \vdash e:\ T'}$$ subtyping

does it type check? – yes

```
def intSqrt(x:Pos) : Pos = { ...}
var p : Pos
var q : Neg
var r : Pos
q = -5
p = q
r = intSqrt(p)
```

Γ= {(p, Pos), (q, Neg), (r, Pos), (intSqrt, Pos $\to$ Pos)}

Runtime error: intSqrt invoked with a negative argument!

x must be able to store any value from T

e can have any value from T

$$\frac{? \qquad \Gamma \vdash e:\ T}{\Gamma \vdash (x = e):\ \text{void}}$$

Cannot use Γ |- e:T to mean "x promises it can store any e $\in$ T"

Recall Our Type Derivation

Pos <: Int

Neg <: Int

$$\frac{\Gamma \vdash x:\ T \qquad \Gamma \vdash e:\ T}{\Gamma \vdash (x = e):\ \text{void}}$$ assignment

$$\frac{\Gamma \vdash e:\ T \qquad \Gamma \vdash T <:\ T'}{\Gamma \vdash e:\ T'}$$ subtyping

does it type check? – yes

```
def intSqrt(x:Pos) : Pos = { ...}
var p : Pos
var q : Neg
var r : Pos
q = -5
p = q
r = intSqrt(p)
```

¡ = {(p, Pos), (q, Neg), (r, Pos), (intSqrt, Pos → Pos)}

Runtime error: intSqrt invoked with a negative argument!

Values from p are integers. But p did not promise to store all kinds of integers/ Only positive ones!

$$\cfrac{\cfrac{\text{p: Pos} \qquad \text{Pos <: Int}}{\text{p: Int}} \qquad \cfrac{\text{q: Neg} \qquad \text{Neg <: Int}}{\text{q: Int}}}{\text{(p=q): void}}$$

Corrected Type Rule for Assignment

Pos <: Int

Neg <: Int

$$\cancel{\frac{\Gamma \vdash x\colon\ T \qquad \Gamma \vdash e\colon\ T}{\Gamma \vdash (x = e)\colon\ void}}$$ assignment

$$\frac{\Gamma \vdash e\colon\ T \qquad \Gamma \vdash T <\colon\ T'}{\Gamma \vdash e\colon\ T'}$$ subtyping

does it type check? – yes

```
def intSqrt(x:Pos) : Pos = { ...}
var p : Pos
var q : Neg
var r : Pos
q = -5
p = q
r = intSqrt(p)
```

¡ = {(p, Pos), (q, Neg), (r, Pos), (intSqrt, Pos $\rightarrow$ Pos)}

does not type check

x must be able to store any value from T

e can have any value from T

$$\frac{(x, T) \in \Gamma \qquad \Gamma \vdash e\colon\ T}{\Gamma \vdash (x = e)\colon\ void}$$

Γ stores declarations (promises)

How could we ensure that some other programs will not break?

Type System Soundness

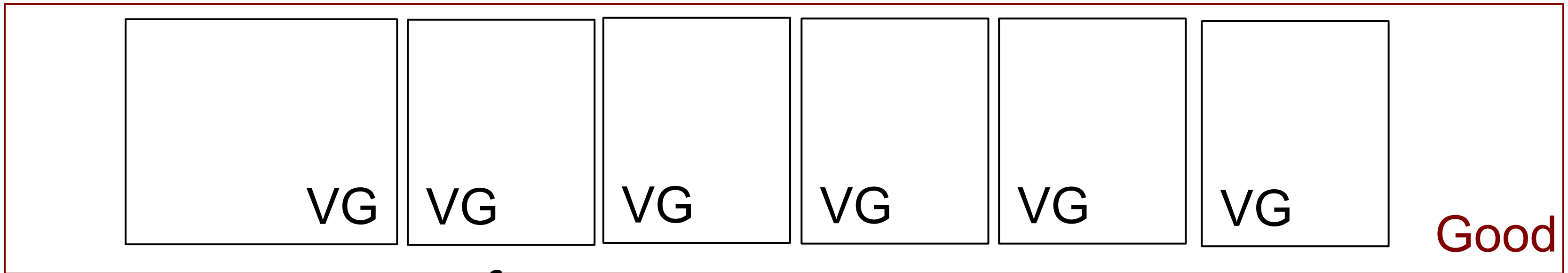
Proving Soundness of Type Systems

- Goal of a sound type system:
  - if the program type checks, then it never “crashes”
  - crash = some precisely specified bad behavior

    e.g. invoking an operation with a wrong type
    - dividing one string by another string “cat” / “frog
    - trying to multiply a Window object by a File object

    e.g. not dividing an integer by zero
- Never crashes: no matter how long it executes
  - proof is done by induction on program execution

Proving Soundness by Induction



- Program moves from state to state
- Bad state = state where program is about to exhibit a bad operation (“cat” / “frog”)
- Good state = state that is not bad
- To prove:
  program type checks → states in all executions are good
- Usually need a *stronger inductive hypothesis*;
  some notion of very good (VG) state such that:
  program type checks → program’s initial state is very good
  state is very good → next state is also very good
  state is very good → state is good (not about to crash)

A Simple Programming Language

Program State

```
var x : Pos
var y : Int
var z : Pos
x = 3
y = -5
z = 4
x = x + z
y = x / z
z = z + x
```

position in source

Initially, all variables have value 1

values of variables:

x = 1

y = 1

z = 1

Program State

```
var x : Pos
var y : Int
var z : Pos
x = 3
y = -5
z = 4
x = x + z
y = x / z
z = z + x
```

position in source

values of variables:

x = 3

y = 1

z = 1

Program State

```
var x : Pos
var y : Int
var z : Pos
x = 3
y = -5
z = 4
x = x + z
y = x / z
z = z + x
```

position in source

values of variables:

x = 3

y = -5

z = 1

Program State

var x : Pos
var y : Int
var z : Pos
x = 3
y = -5
z = 4
x = x + z
y = x / z
z = z + x

position in source

values of variables:
x = 3
y = -5
z = 4

Program State

```
var x : Pos
var y : Int
var z : Pos
x = 3
y = -5
z = 4
x = x + z
y = x / z
z = z + x
```

position in source

values of variables:

x = 7

y = -5

z = 4

Program State

```
var x : Pos
var y : Int
var z : Pos
x = 3
y = -5
z = 4
x = x + z
y = x / z
z = z + x
```

position in source

values of variables:

x = 7

y = 1

z = 4

formal description of such program execution
is called operational semantics

Operational semantics

Operational semantics gives meaning to programs by describing how the program state changes as a sequence of steps.

- big-step semantics: consider the effect of entire blocks
- <u>small-step semantics</u>: consider individual steps (e.g. z = x + y)

V:	set of variables in the program
pc:	integer variable denoting the program counter
$g: V \to Int$	function giving the values of program variables
(g, pc)	program state

Then, for each possible statement in the program we define how it changes the program state.

Example: `z = x+y`

$(g, pc) \to (g', pc + 1)$ s. t. $g' = g[z := g(x)+g(y)]$

Type Rules of Simple Language

Programs:

var x_1 : Pos
var x_2 : Int
...
var x_n : Pos

variable declarations
var x: Pos (strictly positive)
or
var x: Int

followed by

$x_i = x_j$
$x_p = x_q + x_r$
$x_a = x_b / x_c$
...
$x_p = x_q + x_r$

statements of one of the forms
1) x_i=k
2) $x_i = x_j$
3) $x_i = x_j / x_k$
4) $x_i = x_j + x_k$

(No complex expressions)

Type rules:
Γ = { (x_1, Pos),
(x_2, Int),
...
(x_n, Pos)}
Pos <: Int

$$\frac{(x,T) \in \Gamma \qquad \Gamma \vdash e : T}{\Gamma \vdash (x = e) : void}$$

$$\frac{\Gamma \vdash x : T \qquad T <: T'}{\Gamma \vdash x : T'}$$

$$\frac{(x,T) \in \Gamma}{\Gamma \vdash x : T} \qquad \frac{e_1 : Int \qquad e_2 : Int}{e_1 + e_2 : Int}$$

$$\frac{e_1 : Int \qquad e_2 : Pos}{e_1 / e_2 : Int} \qquad \frac{e_1 : Pos \qquad e_2 : Pos}{e_1 + e_2 : Pos}$$

$$\frac{}{\text{k: Pos}} \qquad \frac{}{\text{-k: Int}}$$

Bad State: About to Divide by Zero (Crash)

var x : Pos
var y : Int
var z : Pos
x = 1
y = -1
z = x + y
x = x + z
y = x / z
z = z +

position in source

values of variables:
x = 1
y = -1
z = 0

Definition: state is *bad* if the next instruction is of the form
$x_i = x_j / x_k$ and x_k has value 0 in the current state.

Good State: Not (Yet) About to Divide by Zero

```
var x : Pos
var y : Int
var z : Pos
x = 1
y = -1
z = x + y
x = x + z
y = x / z
z = z + x
```

position in source

values of variables:

x = 1

y = -1

z = 1

Good

Definition: state is *good* if it is not *bad.*

Definition: state is *bad* if the next instruction is of the form $x_i = x_j / x_k$ and x_k has value 0 in the current state.

Good State: Not (Yet) About to Divide by Zero

var x : Pos
var y : Int
var z : Pos
x = 1
y = -1
z = x + y
x = x + z
y = x / z
z = z + x

position in source

values of variables:
x = 1
y = -1
z = 0

Good

Definition: state is *good* if it is not *bad.*

Definition: state is *bad* if the next instruction is of the form $x_i = x_j / x_k$ and x_k has value 0 in the current state.

Moved from Good to Bad in One Step!

Being good is not preserved by one step, not inductive!

It is very local property, does not take future into account.

```
var x : Pos
var y : Int
var z : Pos
x = 1
y = -1
z = x + y
x = x + z
y = x / z
z = z + x
```

position in source

values of variables:

x = 1

y = -1

z = 0

Bad

Definition: state is *good* if it is not *bad.*

Definition: state is *bad* if the next instruction is of the form $x_i = x_j / x_k$ and x_k has value 0 in the current state.

Being Very Good: A Stronger Inductive Property

Pos = { 1, 2, 3, ... }

var x : Pos
var y : Int
var z : Pos
x = 1
y = -1
z = x + y
x = x + z
y = x / z
z = z + x

This state is already not *very good.*
We took future into account.

position in source

values of variables:
x = 1
y = -1
z = 0 $\notin$ Pos

Definition: state is *good* if it is not about to divide by zero.

Definition: state is *very good* if each variable belongs to the domain determined by its type (if z:Pos, then z is strictly positive).

If you are a little typed program, what will your parents teach you?

If you *type check* and succeed:

- you will be *very good* from the start
- if you are *very good*, then you will remain *very good* in the next step
- If you are *very good*, you will not *crash*

Hence, please type check, and you will never crash!

Soundnes proof = defining “very good” and checking the properties above.

Definition of Simple Language

Programs:

var x_1 : Pos
var x_2 : Int
...
var x_n : Pos

variable declarations
var x: Pos
or
var x: Int

followed by

$x_i = x_j$
$x_p = x_q + x_r$
$x_a = x_b / x_c$
...
$x_p = x_q + x_r$

statements of one of the forms
1) x_i=k
2) $x_i = x_j$
3) $x_i = x_j / x_k$
4) $x_i = x_j + x_k$

(No complex expressions)

Type rules:

Γ = { (x_1, Pos),
(x_2, Int),
...
(x_n, Pos)}

Pos <: int

$$\frac{(x,T) \in \Gamma \qquad \Gamma \vdash e : T}{\Gamma \vdash (x = e) : void}$$

$$\frac{\Gamma \vdash x : T \qquad T <: T'}{\Gamma \vdash x : T'}$$

$$\frac{(x,T) \in \Gamma}{\Gamma \vdash x : T} \qquad \frac{e_1 : Int \qquad e_2 : Int}{e_1 + e_2 : Int}$$

$$\frac{e_1 : Int \qquad e_2 : Pos}{e_1 / e_2 : Int} \qquad \frac{e_1 : Pos \qquad e_2 : Pos}{e_1 + e_2 : Pos}$$

$$\frac{}{\text{k: Pos}} \qquad \frac{}{\text{-k: Int}}$$

Checking Properties in Our Case

Holds: in initial state, variables are =1

- If you *type check* and succeed:
  - you will be *very good* from the start.
  - if you are *very good*, then you will remain *very good* in the next step
  - If you are *very good*, you will not *crash.*

$1 \in \text{Pos}$
$1 \in \text{Int}$

If next state is x / z, type rule ensures z has type Pos
Because state is very good, it means $z \in \text{Pos}$
so z is not 0, and there will be no crash.

Definition: state is *very good* if each variable belongs to the domain determined by its type (if z:Pos, then z is strictly positive).

Example Case 1

Assume each variable belongs to its type.

var x : Pos
var y : Pos
var z : Pos
y = 3
z = 2
z = x + y
x = x + z
y = x / z
z = z + x

position in source

values of variables:
x = 1
y = 3
z = 2

the next statement is: z=x+y
where x,y,z are declared Pos.

Goal: prove that again each variable belongs to its type.

- variables other than z did not change, so belong to their type
- z is sum of two positive values, so it will have positive value

Example Case 2

Assume each variable belongs to its type.

var x : Pos
var y : Int
var z : Pos
y = -5
z = 2
z = x + y
x = x + z
y = x / z
z = z + x

position in source

values of variables:
x = 1
y = -5
z = 2

the next statement is: z=x+y
where x,z declared Pos, y declared Int

Goal: prove that again each variable belongs to its type.
this case is impossible, because z=x+y would not type check

How do we know it could not type check?

Must Carefully Check Our Type Rules

var x : Pos
var y : Int
var z : Pos
y = -5
z = 2
z = x + y
x = x + z
y = x / z
z = z + x

Conclude that the only types we can derive are:

x : Pos, x : Int
y : Int
x + y : Int

Cannot type check
z = x + y in this environment.

Type rules:
Γ = { (x_1, Pos),
(x_2, Int),
…
(x_n, Pos)}
Pos <: int

$$\frac{(x,T) \in \Gamma \qquad \Gamma \vdash e : T}{\Gamma \vdash (x = e) : void}$$

$$\frac{\Gamma \vdash x : T \qquad T <: T'}{\Gamma \vdash x : T'}$$

$$\frac{(x,T) \in \Gamma}{\Gamma \vdash x : T} \qquad \frac{e_1 : Int \qquad e_2 : Int}{e_1 + e_2 : Int}$$

$$\frac{e_1 : Int \qquad e_2 : Pos}{e_1 / e_2 : Int} \qquad \frac{e_1 : Pos \qquad e_2 : Pos}{e_1 + e_2 : Pos}$$

$$\frac{}{\text{k: Pos}} \qquad \frac{}{\text{-k: Int}}$$

We would need to check all cases
(there are many, but they are easy)

Back to the start

$$\frac{}{\text{k: Pos}} \qquad \frac{}{\text{-k: Int}}$$

$$\boxed{\frac{\Gamma \vdash x : T \qquad \Gamma \vdash e : T}{\Gamma \vdash (x = e) : void}}$$

$$\frac{\Gamma \vdash x : T \qquad T <: T'}{\Gamma \vdash x : T'}$$

$$\frac{(x, T) \in \Gamma}{\Gamma \vdash x : T}$$

$$\frac{e_1 : Int \qquad e_2 : Int}{e_1 + e_2 : Int}$$

$$\frac{e_1 : Int \qquad e_2 : Pos}{e_1 / e_2 : Int}$$

$$\frac{e_1 : Pos \qquad e_2 : Pos}{e_1 + e_2 : Pos}$$

Does the proof still work?

If not, where does it break?

Remark

- We used in examples Pos <: Int

- Same examples work if we have

class Int { ... }
class Pos **extends** Int { ... }

and is therefore relevant for OO languages

What if we want more complex types?

```
class A { }
class B extends A {
  void foo() { }
}
class Test {
  public static void main(String[]
args) {
    B[] b = new B[5];
    A[] a;
    a = b;
    System.out.println("Hello,");
    a[0] = new A();
    System.out.println("world!");
    b[0].foo();
  }
}
```

- Should it type check?
- Does this type check in Java?
  - can you run it?
- Does this type check in Scala?

What if we want more complex types?

Suppose we add to our language a reference type:

class Ref[T](var content : T)

Programs:

var x_1 : Pos
var x_2 : Int
var x_3 : Ref[Int]
var x_4 : Ref[Pos]

x = y
x = y + z
x = y / z
x = y + z.content
x.content = y

Exercise 1:
Extend the type rules to use with Ref[T] types.
Show your new type system is sound.

Exercise 2:
Can we use the subtyping rule?
If not, where does the proof break?

$$\frac{T <: T'}{\mathrm{Ref}[T] <: \mathrm{Ref}[T']}$$

Simple Parametric Class

class Ref[T](var content : T)

Can we use the subtyping rule

$$\frac{T <: T'}{Ref[T] <: Ref[T']} \qquad \frac{Pos <: Int}{Ref[Pos] <: Ref[Int]}$$

$$\left.\begin{array}{l}\text{var x : Ref[Pos]}\\ \text{var y : Ref[Int]}\\ \text{var z : Int}\end{array}\right\} \Gamma$$

x.content = 1

y.content = -1

y = x

y.content = 0

z = z / x.content

type checks

$$\frac{\dfrac{\Gamma \vdash x : Ref[Pos]}{(x, Ref[Int]) \in \Gamma} \qquad \Gamma \vdash y : Ref[Int]}{(y{=}x){:}void}$$

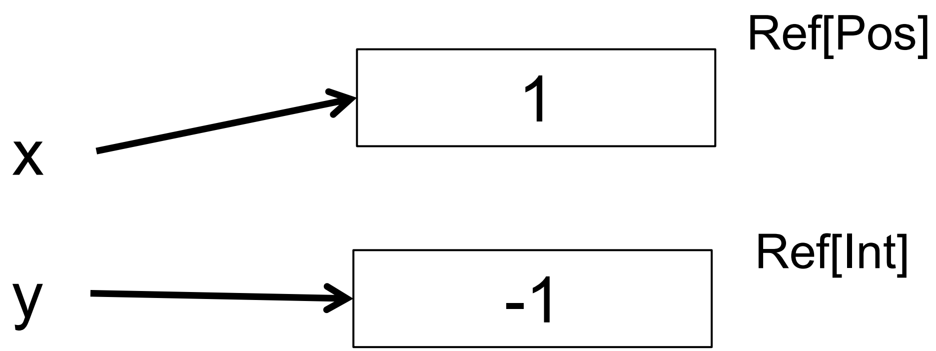
Simple Parametric Class

class Ref[T](var content : T)

Can we use the subtyping rule

$$\frac{T <: T'}{\mathrm{Ref}[T] <: \mathrm{Ref}[T']}$$

```
var x : Ref[Pos]
var y : Ref[Int]
var z : Int
x.content = 1
y.content = -1
y = x
y.content = 0
z = z / x.content
```



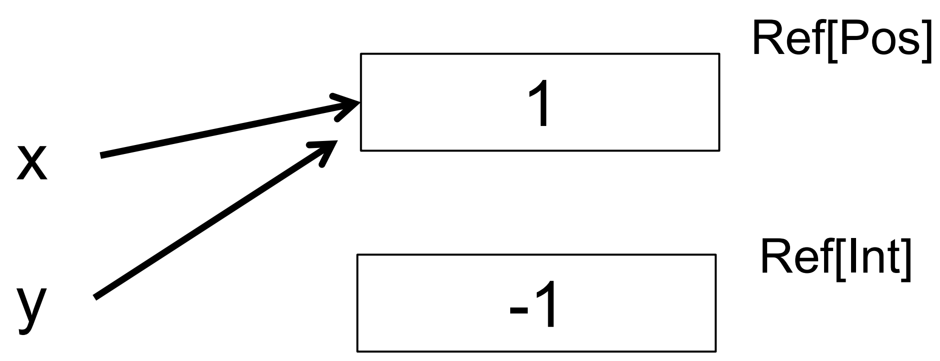
Simple Parametric Class

class Ref[T](var content : T)

Can we use the subtyping rule

$$\frac{T <: T'}{\mathrm{Ref}[T] <: \mathrm{Ref}[T']}$$

```
var x : Ref[Pos]
var y : Ref[Int]
var z : Int
x.content = 1
y.content = -1
y = x
y.content = 0
z = z / x.content
```



Simple Parametric Class

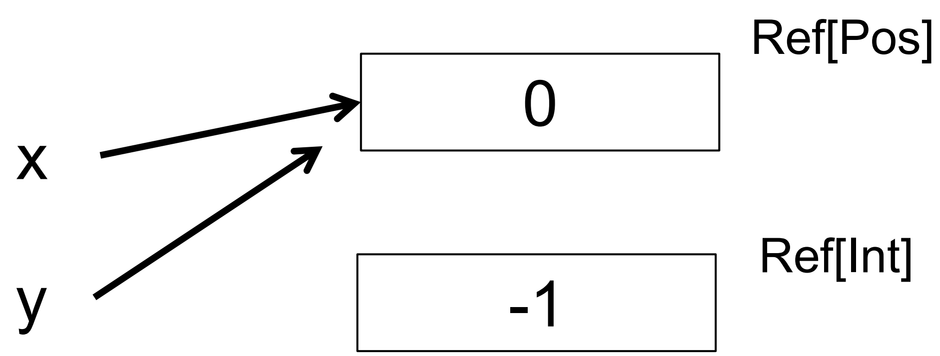
```
class Ref[T](var content : T)
```

Can we use the subtyping rule

$$\frac{T <: T'}{Ref[T] <: Ref[T']}$$

```
var x : Ref[Pos]
var y : Ref[Int]
var z : Int
x.content = 1
y.content = -1
y = x
y.content = 0
z = z / x.content
```

CRASHES



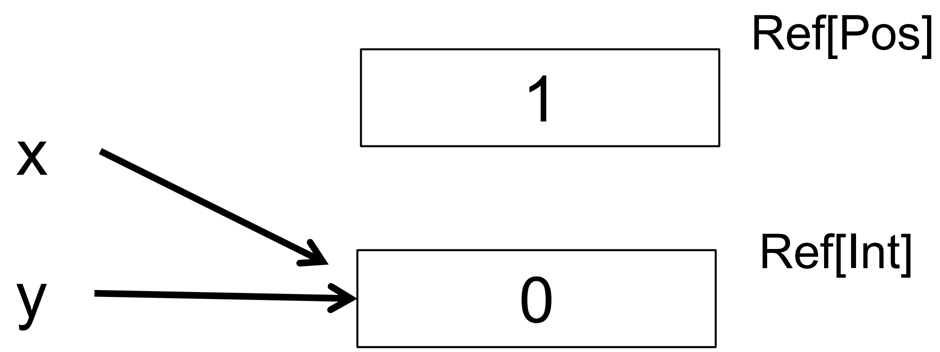
Analogously

class Ref[T](var content : T)

Can we use the converse subtyping rule

$$\frac{\mathrm{T} <: \mathrm{T'}}{\mathrm{Ref[T']} <: \mathrm{Ref[T]}}$$

```
var x : Ref[Pos]
var y : Ref[Int]
var z : Int
x.content = 1
y.content = -1
x = y
y.content = 0
z = z / x.content
```



CRASHES

Mutable Classes do not Preserve Subtyping

```
class Ref[T](var content : T)
```

Even if T <: T',

Ref[T] and Ref[T'] are unrelated types

```
var x : Ref[T]
var y : Ref[T']
...
x = y   ←—— type checks only if T=T'
...
```

Same Holds for Arrays, Vectors, all mutable containers

Even if T <: T',

Array[T] and Array[T'] are unrelated types

```
var x : Array[Pos](1)
var y : Array[Int](1)
var z : Int
x[0] = 1
y[0] = -1
y = x
y[0] = 0
z = z / x[0]
```

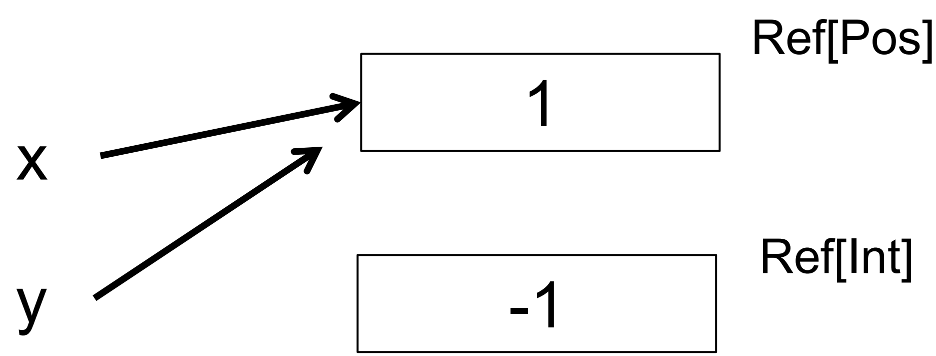
Case in Soundness Proof Attempt

class Ref[T](var content : T)

Can we use the subtyping rule

$$\frac{T <: T'}{\mathrm{Ref}[T] <: \mathrm{Ref}[T']}$$

var x : Ref[Pos]
var y : Ref[Int]
var z : Int
x.content = 1
y.content = -1
y = x
y.content = 0
z = z / x.content



prove each variable belongs to its type:
variables other than y did not change.. (?!)

Mutable vs Immutable Containers

- Immutable container, Coll[T]
  - has methods of form e.g. get(x:A) : T
  - if T <: T', then Coll[T'] has get(x:A) : T'
  - we have $(A \rightarrow T) <: (A \rightarrow T')$
    covariant rule for functions, so Coll[T] <: Coll[T']
- Write-only data structure have
  - setter-like methods, set(v:T) : B
  - if T <: T', then Container[T'] has set(v:T) : B
  - would need $(T \rightarrow B) <: (T' \rightarrow B)$
    contravariance for arguments, so Coll[T'] <: Coll[T]
- Read-Write data structure need both,
  so they are invariant, no subtype on Coll if T <: T'