
Subtyping

Thomas Wies
wies@mpi-sb.mpg.de

Seminar: Types and Programming Languages, WS 02/03
Pierce, ch. 15-18

OVERVIEW

- ① The subtype relation
- ② Typechecking
- ③ Extensions: references, casts
- ④ Case study: featherweight Java

FRAMEWORK

Language used in this talk:

simply typed lambda-calculus + records:

Example:

$r = (\lambda r : \{x : \text{Nat}, y : \text{Nat}\}. r) \{x = 1, y = 2\};$

▷ $r : \{x : \text{Nat}, y : \text{Nat}\}$

$r.x;$

▷ $1 : \text{Nat}$

in the context of imperative objects:

simply typed lambda-calculus + records + references

MOTIVATION

Problem:

Simply typed lambda-calculus is often too restrictive.

Example: $(\lambda r : \{x : \text{Nat}\}. r.x) \{x = 0, y = 1\}$ is not well-typed.

Intuition:

- *Subset semantics*: whenever S is a subset of T , then any term of type S should also be of type T .
- More general: whenever it is safe to use a term of type S in a context of type T , then S is a subtype of T , written $S <: T$.

In the example: $\{x : \text{Nat}, y : \text{Nat}\} <: \{x : \text{Nat}\}$

WHAT IS NEEDED?

- We have to extend our typing rules:

$$\frac{\Gamma \vdash t : S \quad S <: T}{\Gamma \vdash t : T} \text{ (T-SUB)}$$

- We have to formalize what $S <: T$ means.

Notice: Evaluation is not effected by the introduction of subtyping.

THE SUBTYPE RELATION

Top:

$S <: \text{Top}$ (S-TOP)

Reflexivity:

$S <: S$ (S-REFL)

Transitivity:

$$\frac{S <: U \quad U <: T}{S <: T} \text{ (S-TRANS)}$$

Arrow-Types:

$$\frac{T_1 <: S_1 \quad S_2 <: T_2}{S_1 \rightarrow S_2 <: T_1 \rightarrow T_2} \text{ (S-ARROW)}$$

THE SUBTYPE RELATION (2)

Record deepening:

$$\frac{\forall i : S_i <: T_i}{\{l_i : S_i^{i \in 1..n}\} <: \{l_i : T_i^{i \in 1..n}\}} \quad (\text{S-RCDDEPTH})$$

Record widening:

$$\{l_i : T_i^{i \in 1..n+k}\} <: \{l_i : T_i^{i \in 1..n}\} \quad (\text{S-RCDWIDTH})$$

Record permutation:

$$\frac{\{k_i : S_i^{i \in 1..n}\} \text{ is a permutation of } \{l_i : T_i^{i \in 1..n}\}}{\{k_i : S_i^{i \in 1..n}\} <: \{l_i : T_i^{i \in 1..n}\}} \quad (\text{S-RCDPERM})$$

EXAMPLE: A SUBTYPE DERIVATION

Derivation of:

$$\vdash (\lambda r : \{x : \text{Top}\}. r.x) \{x = 0, y = 1\} : \text{Top}$$

TYPE SAFETY

Type safety is preserved in presence of subtyping:

Theorem (Preservation): If $\Gamma \vdash t : T$ and $t \rightarrow t'$, then $\Gamma \vdash t' : T$

Theorem (Progress): If t is a closed, well-typed term, then either t is a value or $t \rightarrow t'$.

TYPECHECKING

- Algorithmic subtyping
- Algorithmic typing

A TYPECHECKER WITH SUBTYPING

How to implement a subtypechecker checking $S <: T$ for two types S and T ?

Problem:

$S <: S$ (S-REFL)

$$\frac{S <: U \quad U <: T}{S <: T} \text{ (S-TRANS)}$$

S and T match any types.

- Rules can be applied in any situation.
- The subtype relation considered so far can not be used to implement a subtypechecker directly.

Idea: Introduce an algorithmic subtype relation $\vdash S <: T$,
s.t. $\vdash S <: T$ iff $S <: T$

ALGORITHMIC SUBTYPING

Observations:

- Reflexivity is not needed for typechecking.
 - drop S-REF
- Transitivity is only needed for record-types.
 - merge record-rules in one single rule
 - drop S-TRANS

New rule for record-subtyping:

$$\frac{\{l_i^{i \in 1..n}\} \subseteq \{k_j^{j \in 1..m}\} \quad k_j = l_i \Rightarrow \vdash S_j <: T_i}{\vdash \{k_j : S_j^{j \in 1..m}\} <: \{l_i : T_i^{i \in 1..n}\}} \quad (\text{SA-RCD})$$

- S-TOP and S-ARROW do not change.

ALGORITHMIC TYPING

For typechecking we have a similar problem:

$$\frac{\Gamma \vdash t : S \quad S <: T}{\Gamma \vdash t : T} \text{ (T-SUB)}$$

t matches any term, hence the rule T-SUB fires on any term.

→ We need an algorithmic typing relation $\Gamma \vdash t : T$

ALGORITHMIC TYPING (2)

Observation:

- T-SUB is only needed to match the argument- and domain-types in application terms.
→ merge T-SUB into T-APP

New rule for applications:

$$\frac{\Gamma \vdash t_1 : T_{11} \rightarrow T_{12} \quad \Gamma \vdash t_2 : T_2 \quad \vdash T_2 <: T_{11}}{\Gamma \vdash t_1 t_2 : T_{12}} \text{ (TA-APP)}$$

Theorem: Algorithmic typing is sound and complete.

- ① if $\Gamma \vdash t : T$, then $\Gamma \vdash t : T$.
- ② if $\Gamma \vdash t : T$, then $\Gamma \vdash t : S$ for some $S <: T$.

SUBTYPING AND EXTENSIONS

- References
- Up- and down-casts

SUBTYPING AND REFERENCES

What conditions must hold in order to get $\text{Ref } S <: \text{Ref } T$?

Example:

$(\lambda r : \text{Ref } \{x : \text{Nat}, y : \text{Nat}\}. !r.x)$ $(\text{ref } \{y = 0\})$ will go wrong.

→ We need $S <: T$ in order to get safe dereferences.

$(\lambda r : \text{Ref } \{x : \text{Nat}, y : \text{Nat}\}. r := \{x = 1\}; !r.y)$ $(\text{ref } \{x = 0, y = 1\})$
will go wrong, too.

→ We also need $T <: S$ in order to get safe assignments.

Simple inference rule:

$$\frac{S <: T \quad T <: S}{\text{Ref } S <: \text{Ref } T} \text{ (S-REF)}$$

REFERENCES REFINED

Decompose Ref T in two new types

- Source T: capability to read from a reference cell
- Sink T: capability to write into a reference cell

and modify the typing rules for references accordingly:

$$\frac{\Gamma \mid \Sigma \vdash t : \text{Source } T}{\Gamma \mid \Sigma \vdash !t : T} \quad (\text{T-DEREF})$$

$$\frac{\Gamma \mid \Sigma \vdash t_1 : \text{Sink } T \quad \Gamma \mid \Sigma \vdash t_2 : T}{\Gamma \mid \Sigma \vdash t_1 := t_2 : \text{Unit}} \quad (\text{T-ASSIGN})$$

REFERENCES REFINED (2)

Now, subtyping for references is easy:

$$\frac{S <: T}{\text{Source } S <: \text{Source } T} \text{ (S-SOURCE)}$$

$$\frac{T <: S}{\text{Sink } S <: \text{Sink } T} \text{ (S-SINK)}$$

Ref is just a subtype of both Source and Sink:

$$\text{Ref } T <: \text{Source } T \text{ (S-REFSOURCE)}$$

$$\text{Ref } T <: \text{Sink } T \text{ (S-REFSINK)}$$

ASCRPTION AS A CASTING OPERATOR

Idea: use ascription operator t as T to perform type casts.

$$\frac{\Gamma \vdash t : T}{\Gamma \vdash t \text{ as } T : T} \text{ (T-ASCRIBE)}$$

Up-casts

Application: information-hiding.

Ascription + subsumption immediately gives us Up-casts.

Example: $\{x = 0, y = 1\}$ as $\{x : \text{Nat}\}$ is well-typed and y is hidden in the context of the ascribed term.

Notice: up-casts do not require ascription, they can be performed using lambda-terms, too.

ASCRPTION AS A CASTING OPERATOR (2)

Down-casts

Application: down-casts + Top provide simple form of polymorphism.

Example: container classes in Java.

Down-casts require an additional typing-rule:

$$\frac{\Gamma \vdash t : S}{\Gamma \vdash t \text{ as } T : T} \text{ (T-DOWNCAST)}$$

Problem: down-casts may be unsound.

Solution: Add dynamic type tests to the evaluation-rules for ascription.

COERCION SEMANTICS

Problem: subtyping may result in performance penalties on the low-level language implementation.

Example:

How to perform efficiently record-field accesses in the presence of a permutation rule?

Idea: *coercion semantics:* Use the type- and subtype-derivation trees to generate additional code for type conversions.

CASE STUDY: FEATHERWEIGHT JAVA

- Interfaces
- Inheritance
- Subtyping
- self and open recursion

IMPERATIVE OBJECTS

What are the essential features of imperative objects?

- Multiple representations
 - same interface, but different implementations
- Encapsulation and information hiding
 - internal state only accessible via interface
 - concrete representation hidden
- Inheritance
 - classes are used as templates for object instantiation
 - derived sub-classes can selectively share code with their super-classes

FEATURES OF IMPERATIVE OBJECTS (CONT'D.)

- Subtyping
 - objects of sub-classes can be used in any super-class context
- self and open recursion
 - methods are allowed to invoke other methods of the same object via `self` or `this`
 - in particular: super-classes may invoke methods declared in sub-classes (late-binding).

A SIMPLE JAVA EXAMPLE

Simple implementation of a counter in Java:

```
class Counter {  
    private int x;  
  
    public Counter() {super(); x=1;}  
  
    public int get () { return x;}  
  
    public void inc () { x++;}  
}
```

Question: How can we mimic this within the simply typed lambda-calculus with subtyping?

INTERFACES

The interfaces can be described by using record types:

- a label with functional type for each public method
- a label for each public instance variable with appropriate type

In the example:

`Counter = {get : Unit → Nat, inc : Unit → Unit};`

OBJECTS

A counter object can be implemented now by allocating the instance variable and constructing the method table:

```
c = let x = ref 1 in
  {get = λ_ : Unit. !x,
   inc = λ_ : Unit. x := succ(!x)};
```

▷ c : Counter

```
(c.inc unit; c.inc unit; c.get unit);
```

▷ 3 : Nat

A SIMPLE CLASS

Define a representation type for the instance variables:

CounterRep = {x : Ref Nat};

The counter class now abstracts over the counter representation:

counterClass =

λr : CounterRep.

{get = λ_ : Unit. !(r.x),

inc = λ_ : Unit. x := succ(!(r.x))};

▷ counterClass : CounterRep → Counter

New objects can be instantiated via an object generator:

newCounter =

λ_ : Unit. let r = {x = ref 1} in

counterClass r;

▷ newCounter : Unit → Counter

INHERITANCE IN JAVA

Example of an inherited class in Java:

```
class ResetCounter extends Counter {  
    public ResetCounter() {super();}  
  
    public void reset () { x = 1;}  
}
```

INHERITANCE

First we need to extend the counter interface:

$$\text{ResetCounter} = \{\text{get} : \text{Unit} \rightarrow \text{Nat}, \text{inc} : \text{Unit} \rightarrow \text{Unit}, \\ \text{reset} : \text{Unit} \rightarrow \text{Unit}\};$$

Then we can reuse counterClass within resetCounterClass:

$$\begin{aligned} \text{resetCounterClass} = \\ \lambda r : \text{CounterRep}. \\ \quad \text{let super} = \text{counterClass } r \text{ in} \\ \quad \{\text{get} = \text{super.get}, \\ \quad \text{inc} = \text{super.inc}, \\ \quad \text{reset} = \lambda_ : \text{Unit}. r.x := 1\}; \end{aligned}$$

▷ $\text{resetCounterClass} : \text{CounterRep} \rightarrow \text{ResetCounter}$

SUBTYPING

Record-subtyping provides all we need for subtyping between objects:

`ResetCounter <: Counter`

Hence any reset-counter can be used safely as a counter:

`rc = newResetCounter unit;`

▷ `rc : ResetCounter`

`inc3 = λc : Counter. c.inc unit; c.inc unit; c.inc unit;`

▷ `inc3 : Counter → Unit`

`(inc3 rc; rc.reset unit; inc3 rc; rc.get unit);`

▷ `4 : Nat`

CLASSES WITH self

Let us implement a new SetCounter class that provides a method to set the counter to a given amount:

$$\text{SetCounter} = \{\text{get} : \text{Unit} \rightarrow \text{Nat}, \text{set} : \text{Nat} \rightarrow \text{Unit}, \\ \text{inc} : \text{Unit} \rightarrow \text{Unit}\};$$

We use fixpoint recursion to introduce self:

```
setCounterClass =
  λr : CounterRep.
    fix
      (λself : SetCounter.
        {get = λ_ : Unit. !(r.x),
         set = λi : Nat. r.x := i,
         inc = λ_ : Unit. self.set (succ (self.get unit))});
▷ setCounterClass : CounterRep → SetCounter
```

OPEN RECURSION IN JAVA

Example of open recursion in Java:
(in Java all methods are late-bound)

```
class SetCounter extends Counter {  
    public SetCounter() {super();}  
  
    // set will be bound to a sub-class' method later  
    public void reset () { this.set 1}  
  
    public void set(int i ) { x = i ;}  
}
```

OPEN RECURSION IN JAVA (2)

```
class BackupCounter extends SetCounter {  
  
    private int b;  
  
    // bind super-class declaration of set to this one  
    public void set(int i ) { b = x; super.set i}  
  
    public void restore () { x = b;}  
  
}
```

OPEN RECURSION

There are several possibilities to implement open recursion:

- We can use fixpoint recursion again.
- We can use references.

We will use references here, which is the more efficient solution.

OPEN RECURSION VIA REFERENCES

New SetCounter interface:

$$\text{SetCounter} = \{\text{get} : \text{Unit} \rightarrow \text{Nat}, \text{inc} : \text{Unit} \rightarrow \text{Unit}, \\ \text{set} : \text{Nat} \rightarrow \text{Unit}, \text{reset} : \text{Unit} \rightarrow \text{Unit}\};$$

self is now a reference to a method table:

setCounterClass =

$$\lambda r : \text{SetCounterRep}. \lambda \text{self} : \text{Ref SetCounter}.$$
$$\text{let super} = \text{counterClass } r \text{ in}$$
$$\{\text{get} = \text{super.get},$$
$$\text{inc} = \text{super.inc},$$
$$\text{set} = \lambda i : \text{Nat}. r.x := i,$$
$$\text{reset} = \lambda _ : \text{Unit}. (!\text{self}).\text{set } 1\};$$

▷ setCounterClass : CounterRep → Ref SetCounter → SetCounter

OBJECT GENERATION FOR OPEN RECURSION

For object generation a dummy object must be allocated first:

`newSetCounter =`

```
  λ_ : Unit. let r = {x = ref 1} in
    let mTbl = ref {get = λ_ : Unit. 0,
                    inc = λ_ : Unit. unit,
                    set = λi : Nat. unit,
                    reset = λ_ : Unit. unit} in
      (mTbl := setCounterClass r mTbl); !mTbl;
```

▷ `newSetCounter : Unit → SetCounter`

BACKUPCOUNTER TYPES

The required BackupCounter interface, and representation:

$$\text{BackupCounter} = \{\text{get} : \text{Unit} \rightarrow \text{Nat}, \text{inc} : \text{Unit} \rightarrow \text{Unit}, \\ \text{set} : \text{Nat} \rightarrow \text{Unit}, \text{reset} : \text{Unit} \rightarrow \text{Unit}, \\ \text{restore} : \text{Unit} \rightarrow \text{Unit}\};$$
$$\text{BackupCounterRep} = \{x : \text{Nat}, b : \text{Nat}\};$$

FIRST ATTEMPT - FAILS!

Problem: S-REF does not allow the required subtyping:

BackupCounterClass =

λself : Ref BackupCounter.

λr : BackupCounterRep. let super = setCounterClass r **self** in

{get = super.get,

inc = super.inc,

set = λi : Nat. r.b :=!(r.x); super.set i,

reset = super.reset,

restore = λ_ : Unit. r.x :=!(r.b)};

▷ Error : parameter type mismatch

REFINED VERSION

Solution:

In `setCounterClass` only read-access to the method table is needed.

→ Use `Source SetCounter` instead of `Ref SetCounter`:

```
setCounterClass =
```

```
  λr : SetCounterRep.
```

```
    λself : Source SetCounter.
```

```
      let super = counterClass r in
```

```
        {get = super.get,
```

```
          inc = super.inc,
```

```
          set = λi : Nat. r.x := i,
```

```
          reset = λ_ : Unit. (!self).set 1};
```

```
▷ setCounterClass : CounterRep → Source SetCounter → SetCounter
```

REFINED VERSION (2)

Now the backup-counter class typechecks:

BackupCounterClass =

λself : Source BackupCounter.

λr : BackupCounterRep. let super = setCounterClass r self in

{get = super.get,

inc = super.inc,

set = λi : Nat. r.b =!(r.x); super.set i,

reset = super.reset,

restore = λ_ : Unit. r.x :=!(r.b)};

▷ backupCounterClass : BackupCounterRep →

Source BackupCounter → BackupCounter

CONCLUSION AND OUTLOOK

- Typing can be extended to respect subset-relations on types.
- Object-oriented language features can be expressed in the simply typed lambda-calculus.
- Subtyping introduces efficiency problems.
Possible solution: *coercion semantics*

Aspects not considered in this talk:

- Additional extensions: variants, lists, ...
- Additional types: the bottom type, joins and meets, ...
- Method-sharing between objects of the same class.
→ Bounded quantification *cf. Pierce, ch. 27*