

A Semantic Foundation for Gradual Set-theoretic Types

GPL Award Accessit

Victor Lanvin

June 7th, 2022

Gradual Typing

- Goal: have both **static** and **dynamic** typing in the same language.
- How: by adding a **dynamic type**, denoted “?”.

Gradual Typing

- Goal: have both **static** and **dynamic** typing in the same language.
- How: by adding a **dynamic type**, denoted “?”.
- **Allows for a trade-off between safety and programming productivity.**

Gradual Typing

- Goal: have both **static** and **dynamic** typing in the same language.
- How: by adding a **dynamic type**, denoted “?”.
- **Allows for a trade-off between safety and programming productivity.**

The transition is **gradual**:

`? ≍ ? → ? ≍ Int → ? ≍ Int → Bool`

Set-Theoretic Types

- **Types** with **connectives** (\vee , \wedge , \neg).

Set-Theoretic Types

- **Types** with **connectives** (\vee , \wedge , \neg).

$(\text{Int} \rightarrow \text{Int}) \wedge (\text{Bool} \rightarrow \text{Bool}) = \text{overloaded function}$

Set-Theoretic Types

- **Types** with **connectives** (\vee , \wedge , \neg).

`(Int -> Int) \wedge (Bool -> Bool)` = overloaded function

`if x then 3 else true : Int \vee Bool`

Set-Theoretic Types

- **Types** with **connectives** (\vee , \wedge , \neg).

`(Int -> Int) \wedge (Bool -> Bool)` = overloaded function

`if x then 3 else true : Int \vee Bool`

`(x = e \in Int)? true : x : Bool \vee \neg Int`

Set-Theoretic Types

- **Types** with **connectives** (\vee , \wedge , \neg).

`(Int -> Int) \wedge (Bool -> Bool)` = overloaded function

`if x then 3 else true : Int \vee Bool`

`(x = e \in Int)? true : x : Bool \vee \neg Int`

- **Powerful** but often syntactically heavy.

Set-Theoretic Types

- **Types** with **connectives** (\vee , \wedge , \neg).

`(Int -> Int) \wedge (Bool -> Bool)` = overloaded function

`if x then 3 else true : Int \vee Bool`

`(x = e \in Int)? true: x : Bool \vee \neg Int`

- **Powerful** but often syntactically heavy.

- In **Semantic subtyping**:

Types \simeq Sets of values

Subtyping \simeq Set-containment

Motivating Example (1/2)

Let's write a map, that can work on both arrays and lists depending on a condition:

```
let map (condition : Bool) (f :  $\alpha$  ->  $\beta$ ) (data : ) : =
```

Motivating Example (1/2)

Let's write a map, that can work on both arrays and lists depending on a condition:

```
let map (condition : Bool) (f :  $\alpha$  ->  $\beta$ ) (data : ) :    =  
  if condition then  
    List.map f data  
  else  
    Array.map f data
```

Motivating Example (1/2)

Let's write a map, that can work on both arrays and lists depending on a condition:

```
let map (condition : Bool) (f :  $\alpha$  ->  $\beta$ ) (data : ?) :    =  
  if condition then  
    List.map f data  
  else  
    Array.map f data
```

Motivating Example (1/2)

Let's write a map, that can work on both arrays and lists depending on a condition:

```
let map (condition : Bool) (f :  $\alpha$  ->  $\beta$ ) (data : ?) : ? =  
  if condition then  
    List.map f data  
  else  
    Array.map f data
```

Motivating Example (1/2)

Let's write a map, that can work on both arrays and lists depending on a condition:

```
let map (condition : Bool) (f :  $\alpha$  ->  $\beta$ ) (data : ?) : ? =  
  if condition then  
    List.map f data  
  else  
    Array.map f data
```

Runtime **checks** or **casts** are then inserted **automatically** by the compiler.

Motivating Example (1/2)

Let's write a map, that can work on both arrays and lists depending on a condition:

```
let map (condition : Bool) (f :  $\alpha$  ->  $\beta$ ) (data : ?) : ? =  
  if condition then  
    List.map f data < $\alpha$  list>  
  else  
    Array.map f data
```

Runtime **checks** or **casts** are then inserted **automatically** by the compiler.

Motivating Example (1/2)

Let's write a map, that can work on both arrays and lists depending on a condition:

```
let map (condition : Bool) (f :  $\alpha$  ->  $\beta$ ) (data : ?) : ? =  
  if condition then  
    List.map f data < $\alpha$  list>  
  else  
    Array.map f data < $\alpha$  array>
```

Runtime **checks** or **casts** are then inserted **automatically** by the compiler.

Motivating Example (1/2)

Let's write a map, that can work on both arrays and lists depending on a condition:

```
let map (condition : Bool) (f :  $\alpha$  ->  $\beta$ ) (data : ?) : ? =  
  if condition then  
    List.map f data < $\alpha$  list>  
  else  
    Array.map f data < $\alpha$  array>
```

Runtime **checks** or **casts** are then inserted **automatically** by the compiler.

This is however very **unsafe**, as it accepts a **string** for example.

Motivating Example (2/2)

```
let map condition f
  (data : ( $\alpha$  list  $\vee$   $\alpha$  array) ) =
  if condition then
    List.map f data
  else
    Array.map f data
```

Motivating Example (2/2)

```
let map condition f
  (data : ( $\alpha$  list  $\vee$   $\alpha$  array)  $\wedge$  ?) =
  if condition then
    List.map f data
  else
    Array.map f data
```

Motivating Example (2/2)

```
let map condition f
  (data : ( $\alpha$  list  $\vee$   $\alpha$  array)  $\wedge$  ?) =
  if condition then
    List.map f data
  else
    Array.map f data
```

- By **subtyping**, $(\alpha \text{ list } \vee \alpha \text{ array}) \wedge ? \leq ?$.

Motivating Example (2/2)

```
let map condition f
  (data : ( $\alpha$  list  $\vee$   $\alpha$  array)  $\wedge$  ?) =
  if condition then
    List.map f data
  else
    Array.map f data
```

- By **subtyping**, $(\alpha \text{ list } \vee \alpha \text{ array}) \wedge ? \leq ?$.
- Can only be used with **lists or arrays**.
- No need for **manual type checks**.

Motivating Example (2/2)

```
let map condition f
  (data : ( $\alpha$  list  $\vee$   $\alpha$  array)  $\wedge$  ?) =
  if condition then
    List.map f data
  else
    Array.map f data
```

- By **subtyping**, $(\alpha \text{ list } \vee \alpha \text{ array}) \wedge ? \leq ?$.
- Can only be used with **lists or arrays**.
- No need for **manual type checks**.
- We want to infer **all non-gradual types** (including the return type).

Motivating Example (2/2)

```
let map (condition : Bool) f
  (data : ( $\alpha$  list  $\vee$   $\alpha$  array)  $\wedge$  ?) =
  if condition then
    List.map f data
  else
    Array.map f data
```

- By **subtyping**, $(\alpha \text{ list } \vee \alpha \text{ array}) \wedge ? \leq ?$.
- Can only be used with **lists or arrays**.
- No need for **manual type checks**.
- We want to infer **all non-gradual types** (including the return type).

Motivating Example (2/2)

```
let map condition (f :  $\alpha \rightarrow \beta$ )  
  (data : ( $\alpha$  list  $\vee$   $\alpha$  array)  $\wedge$  ?) =  
  if condition then  
    List.map f data  
  else  
    Array.map f data
```

- By **subtyping**, $(\alpha \text{ list } \vee \alpha \text{ array}) \wedge ? \leq ?$.
- Can only be used with **lists or arrays**.
- No need for **manual type checks**.
- We want to infer **all non-gradual types** (including the return type).

Motivating Example (2/2)

```
let map condition f
  (data : ( $\alpha$  list  $\vee$   $\alpha$  array)  $\wedge$  ?) :  $\beta$  list  $\vee$   $\beta$  array =
  if condition then
    List.map f data
  else
    Array.map f data
```

- By **subtyping**, $(\alpha \text{ list } \vee \alpha \text{ array}) \wedge ? \leq ?$.
- Can only be used with **lists or arrays**.
- No need for **manual type checks**.
- We want to infer **all non-gradual types** (including the return type).

How it is Usually Done

1. Define a **subtype-consistency** relation \lesssim .

How it is Usually Done

1. Define a **subtype-consistency** relation \lesssim .

This relation is not transitive! $? \lesssim \tau \lesssim ?$ for all τ

How it is Usually Done

1. Define a **subtype-consistency** relation \lesssim .

This relation is not transitive! $? \lesssim \tau \lesssim ?$ for all τ

2. Embed this relation into typing rules.

$$\frac{\Gamma \vdash e_1 : \tau_1 \rightarrow \tau'_1 \quad \Gamma \vdash e_2 : \tau_2 \quad \tau_2 \lesssim \tau_1}{\Gamma \vdash e_1 e_2 : \tau'_1}$$

How it is Usually Done

1. Define a **subtype-consistency** relation \lesssim .

This relation is not transitive! $? \lesssim \tau \lesssim ?$ for all τ

2. Embed this relation into typing rules.

$$\frac{\Gamma \vdash e_1 : \tau_1 \quad \Gamma \vdash e_2 : \tau_2 \quad \tau_2 \lesssim \text{dom}(\tau_1)}{\Gamma \vdash e_1 e_2 : \tau_1 \circ \tau_2}$$

This gets even more complicated with set-theoretic types!

Our (First) Approach

1. Translate gradual types to **static types** (types without **?**) **with variables**.

Our (First) Approach

1. Translate gradual types to **static types** (types without ?) **with variables**.
2. Define **transitive** relations on gradual types, and in particular “**precision**” which contains the **essence of gradual typing**.

Our (First) Approach

1. Translate gradual types to **static types** (types without ?) **with variables**.
2. Define **transitive** relations on gradual types, and in particular “**precision**” which contains the **essence of gradual typing**.
3. Embed precision into **more and more complex systems** (Hindley-Milner, with subtyping, and with semantic subtyping).

Our (First) Approach

1. Translate gradual types to **static types** (types without ?) **with variables**.
2. Define **transitive** relations on gradual types, and in particular “**precision**” which contains the **essence of gradual typing**.
3. Embed precision into **more and more complex systems** (Hindley-Milner, with subtyping, and with semantic subtyping).

Important remark: this translation is **only used** to define and compute relations, and **is not done in the source program**.

Subtyping

Subtyping only allows us to move **inside** the dynamic world, or **inside** the static world. It **does not** allow crossing the barrier.

Subtyping

Subtyping only allows us to move **inside** the dynamic world, or **inside** the static world. It **does not** allow crossing the barrier.

As opposed to consistent subtyping, it is **transitive**:

$$? \leq ? \quad ? \not\leq \text{Int} \quad \text{Int} \not\leq ?$$

Subtyping

Subtyping only allows us to move **inside** the dynamic world, or **inside** the static world. It **does not** allow crossing the barrier.

As opposed to consistent subtyping, it is **transitive**:

$$? \leq ? \quad ? \not\leq \text{Int} \quad \text{Int} \not\leq ?$$

It can be used to handle unions and intersections, by **simply plugging-in** the static version of **semantic subtyping**:

$$? \leq ? \vee \text{Int} \quad \text{Int} \wedge ? \leq ?$$

Precision is what allows us to **cross the barrier** from the dynamic world into the static world (**and only this way**).

Precision is what allows us to **cross the barrier** from the dynamic world into the static world (**and only this way**).

? \preceq τ for every τ

? \rightarrow ? \preceq $\tau_1 \rightarrow \tau_2$ for every τ_1, τ_2

Precision

Precision is what allows us to **cross the barrier** from the dynamic world into the static world (**and only this way**).

$? \preceq \tau$ for every τ

$? \rightarrow ? \preceq \tau_1 \rightarrow \tau_2$ for every τ_1, τ_2

And it is **transitive**:

$? \preceq ? \rightarrow ? \preceq ? \rightarrow \text{Int} \preceq \text{Int} \rightarrow \text{Int}$

Precision

Precision is what allows us to **cross the barrier** from the dynamic world into the static world (**and only this way**).

$? \preceq \tau$ for every τ

$? \rightarrow ? \preceq \tau_1 \rightarrow \tau_2$ for every τ_1, τ_2

And it is **transitive**:

$? \preceq ? \rightarrow ? \preceq ? \rightarrow \text{Int} \preceq \text{Int} \rightarrow \text{Int}$

Therefore it can be embedded into a type system as a **subsumption-like** rule: **materialization**.

$$\frac{}{\Gamma, x : \tau \vdash x : \tau} \qquad \frac{\Gamma, x : \tau_1 \vdash e : \tau_2}{\Gamma \vdash \lambda x. e : \tau_1 \rightarrow \tau_2}$$
$$\frac{\Gamma \vdash e_1 : \tau_1 \rightarrow \tau_2 \quad \Gamma \vdash e_2 : \tau_1}{\Gamma \vdash e_1 e_2 : \tau_2}$$

Declarative Type Systems

$$\frac{}{\Gamma, x : \tau \vdash x : \tau} \qquad \frac{\Gamma, x : \tau_1 \vdash e : \tau_2}{\Gamma \vdash \lambda x. e : \tau_1 \rightarrow \tau_2}$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \rightarrow \tau_2 \quad \Gamma \vdash e_2 : \tau_1}{\Gamma \vdash e_1 e_2 : \tau_2}$$

$$\frac{\Gamma \vdash e : \tau_1 \quad \tau_1 \preceq \tau_2}{\Gamma \vdash e : \tau_2} \qquad \frac{\Gamma \vdash e : \tau_1 \quad \tau_1 \leq \tau_2}{\Gamma \vdash e : \tau_2}$$

$$\frac{}{\Gamma, x : \forall \vec{\alpha}. \tau \vdash x : \tau \{ \vec{\alpha} := \vec{t} \}} \quad \frac{\Gamma, x : \tau_1 \vdash e : \tau_2}{\Gamma \vdash \lambda x. e : \tau_1 \rightarrow \tau_2}$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \rightarrow \tau_2 \quad \Gamma \vdash e_2 : \tau_1}{\Gamma \vdash e_1 e_2 : \tau_2}$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \quad \Gamma, x : \text{Gen}_{\Gamma}(\tau_1) \vdash e_2 : \tau}{\Gamma \vdash \text{let } x = e_1 \text{ in } e_2 : \tau}$$

$$\frac{}{\Gamma, x : \forall \vec{\alpha}. \tau \vdash x : \tau \{ \vec{\alpha} := \vec{t} \}} \qquad \frac{\Gamma, x : \tau_1 \vdash e : \tau_2}{\Gamma \vdash \lambda x. e : \tau_1 \rightarrow \tau_2}$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \rightarrow \tau_2 \quad \Gamma \vdash e_2 : \tau_1}{\Gamma \vdash e_1 e_2 : \tau_2}$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \quad \Gamma, x : \text{Gen}_\Gamma(\tau_1) \vdash e_2 : \tau}{\Gamma \vdash \text{let } x = e_1 \text{ in } e_2 : \tau}$$

$$\frac{\Gamma \vdash e : \tau_1 \quad \tau_1 \preceq \tau_2}{\Gamma \vdash e : \tau_2}$$

$$\frac{}{\Gamma, x : \forall \vec{\alpha}. \tau \vdash x : \tau \{ \vec{\alpha} := \vec{t} \}} \qquad \frac{\Gamma, x : \tau_1 \vdash e : \tau_2}{\Gamma \vdash \lambda x. e : \tau_1 \rightarrow \tau_2}$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \rightarrow \tau_2 \quad \Gamma \vdash e_2 : \tau_1}{\Gamma \vdash e_1 e_2 : \tau_2}$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \quad \Gamma, x : \text{Gen}_\Gamma(\tau_1) \vdash e_2 : \tau}{\Gamma \vdash \text{let } x = e_1 \text{ in } e_2 : \tau}$$

$$\frac{\Gamma \vdash e : \tau_1 \quad \tau_1 \preceq \tau_2}{\Gamma \vdash e : \tau_2} \qquad \frac{\Gamma \vdash e : \tau_1 \quad \tau_1 \leq \tau_2}{\Gamma \vdash e : \tau_2}$$

$$\frac{}{\Gamma, x : \forall \vec{\alpha}. \tau \vdash x : \tau \{ \vec{\alpha} := \vec{t} \}} \qquad \frac{\Gamma, x : \tau_1 \vdash e : \tau_2}{\Gamma \vdash \lambda x. e : \tau_1 \rightarrow \tau_2}$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \rightarrow \tau_2 \quad \Gamma \vdash e_2 : \tau_1}{\Gamma \vdash e_1 e_2 : \tau_2}$$

$$\frac{\Gamma \vdash e_1 : \tau_1 \quad \Gamma, x : \text{Gen}_\Gamma(\tau_1) \vdash e_2 : \tau}{\Gamma \vdash \text{let } x = e_1 \text{ in } e_2 : \tau}$$

$$\frac{\Gamma \vdash e : \tau_1 \quad \tau_1 \preceq \tau_2}{\Gamma \vdash e : \tau_2} \qquad \frac{\Gamma \vdash e : \tau_1 \quad \tau_1 \leq \tau_2}{\Gamma \vdash e : \tau_2}$$

And as a bonus, we get the **static gradual guarantee** for free!

Theorem

For every type $\tau \in \text{GTypes}$, there exists $t_1, t_2 \in \text{STypes}$ such that:

$$\tau \preceq t_1 \text{ and } \tau \preceq t_2$$

$$\forall \tau' \in \text{GTypes}. \tau \preceq \tau' \implies t_1 \leq \tau' \leq t_2$$

Theorem

For every type $\tau \in \text{GTypes}$, there exists $t_1, t_2 \in \text{STypes}$ such that:

$$\tau \preceq t_1 \text{ and } \tau \preceq t_2$$

$$\forall \tau' \in \text{GTypes}. \tau \preceq \tau' \implies t_1 \leq \tau' \leq t_2$$

We write $t_1 = \tau^\downarrow$ and $t_2 = \tau^\uparrow$.

Theorem

For every type $\tau \in \text{GTypes}$, there exists $t_1, t_2 \in \text{STypes}$ such that:

$$\tau \preceq t_1 \text{ and } \tau \preceq t_2$$

$$\forall \tau' \in \text{GTypes}. \tau \preceq \tau' \implies t_1 \leq \tau' \leq t_2$$

We write $t_1 = \tau^\downarrow$ and $t_2 = \tau^\uparrow$.

$$(\text{?} \rightarrow \text{?})^\uparrow = 0 \rightarrow \mathbb{1} \quad (\text{?} \rightarrow \text{?})^\downarrow = \mathbb{1} \rightarrow 0$$

Extremal Materializations

Theorem

For every type $\tau \in \text{GTypes}$, there exists $t_1, t_2 \in \text{STypes}$ such that:

$$\tau \preceq t_1 \text{ and } \tau \preceq t_2$$

$$\forall \tau' \in \text{GTypes}. \tau \preceq \tau' \implies t_1 \leq \tau' \leq t_2$$

We write $t_1 = \tau^\downarrow$ and $t_2 = \tau^\uparrow$.

$$(\text{?} \rightarrow \text{?})^\uparrow = 0 \rightarrow \mathbb{1} \quad (\text{?} \rightarrow \text{?})^\downarrow = \mathbb{1} \rightarrow 0$$

These types are computed in **linear time**!

An Equivalent Representation of Gradual Types

We show the following:

$$\tau_1 \leq \tau_2 \iff \begin{cases} \tau_1^{\Downarrow} \leq \tau_2^{\Downarrow} \\ \tau_1^{\Uparrow} \leq \tau_2^{\Uparrow} \end{cases}$$

An Equivalent Representation of Gradual Types

We show the following:

$$\tau_1 \leq \tau_2 \iff \begin{cases} \tau_1^{\Downarrow} \leq \tau_2^{\Downarrow} \\ \tau_1^{\Uparrow} \leq \tau_2^{\Uparrow} \end{cases} \quad \tau_1 \preceq \tau_2 \iff \begin{cases} \tau_1^{\Downarrow} \leq \tau_2^{\Downarrow} \\ \tau_2^{\Uparrow} \leq \tau_1^{\Uparrow} \end{cases}$$

An Equivalent Representation of Gradual Types

We show the following:

$$\tau_1 \leq \tau_2 \iff \begin{cases} \tau_1^{\Downarrow} \leq \tau_2^{\Downarrow} \\ \tau_1^{\Uparrow} \leq \tau_2^{\Uparrow} \end{cases} \quad \tau_1 \preceq \tau_2 \iff \begin{cases} \tau_1^{\Downarrow} \leq \tau_2^{\Downarrow} \\ \tau_2^{\Uparrow} \leq \tau_1^{\Uparrow} \end{cases}$$

Moreover, we have that **for every gradual type** τ ,

$$\tau \simeq \tau^{\Downarrow} \vee (? \wedge \tau^{\Uparrow})$$

An Equivalent Representation of Gradual Types

We show the following:

$$\tau_1 \leq \tau_2 \iff \begin{cases} \tau_1^\Downarrow \leq \tau_2^\Downarrow \\ \tau_1^\Uparrow \leq \tau_2^\Uparrow \end{cases} \quad \tau_1 \preceq \tau_2 \iff \begin{cases} \tau_1^\Downarrow \leq \tau_2^\Downarrow \\ \tau_2^\Uparrow \leq \tau_1^\Uparrow \end{cases}$$

Moreover, we have that **for every gradual type** τ ,

$$\tau \simeq \tau^\Downarrow \vee (? \wedge \tau^\Uparrow)$$

We can use this representation to **lift operators** to gradual types!

$$\text{dom}(\tau) \stackrel{\text{def}}{=} \text{dom}(\tau^\Uparrow) \vee (? \wedge \text{dom}(\tau^\Downarrow))$$

Conclusion

Your favorite typing rules + Materialization + Subsumption = Your gradual type system

Your favorite typing rules + Materialization + Subsumption = Your gradual type system

We present:

1. A simple method of **declaratively adding** gradual typing to any existing type system.

Your favorite typing rules + Materialization + Subsumption = Your gradual type system

We present:

1. A simple method of **declaratively adding** gradual typing to any existing type system.
2. A **set-theoretic interpretation** of gradual types that has considerable consequences.

Your favorite typing rules + Materialization + Subsumption = Your gradual type system

We present:

1. A simple method of **declaratively adding** gradual typing to any existing type system.
2. A **set-theoretic interpretation** of gradual types that has considerable consequences.
3. The **algorithmic systems** for our GTLC with set-theoretic types.

Your favorite typing rules + Materialization + Subsumption = Your gradual type system

We present:

1. A simple method of **declaratively adding** gradual typing to any existing type system.
2. A **set-theoretic interpretation** of gradual types that has considerable consequences.
3. The **algorithmic systems** for our GTLC with set-theoretic types.
4. **Denotational semantics** for several calculi, including CDuce, and a GTLC without set-theoretic types.

- Fully **unify** our logical approach and our denotational semantics.

- Fully **unify** our logical approach and our denotational semantics.
- Sound and complete **type inference** for gradual set-theoretic types.

- Fully **unify** our logical approach and our denotational semantics.
- Sound and complete **type inference** for gradual set-theoretic types.
- Add **more features** to our calculus, such as intersection types for functions.

- Fully **unify** our logical approach and our denotational semantics.
- Sound and complete **type inference** for gradual set-theoretic types.
- Add **more features** to our calculus, such as intersection types for functions.
- A **denotational semantics** for a cast calculus with set-theoretic types.