# Abstractions for Type-Level Programming

Olivier Blanvillain Tuesday, 22 March 2022

# Example 1: database queries (2016)



## Example 1: database queries (2016)

Spark APIs are practically untyped:

class DataFrame {
 /\*\* Inner equi-join with another DataFrame on the given column.
 \* The join column will only appear once in the output. \*/
 def join(right: DataFrame, column: String): DataFrame
}

Revisited with type-level programming:

class DF[X] {
 def join[Y](df: DF[Y], col: String): DF[col+(X-col)++(Y-col)]
}

## Example 1: database queries (2016)

Spark APIs are practically untyped:

class DataFrame {
 /\*\* Inner equi-join with another DataFrame on the given column.
 \* The join column will only appear once in the output. \*/
 def join(right: DataFrame, column: String): DataFrame
}

Revisited with type-level programming:

class DF[X] {
 def join[Y](df: DF[Y], col: String): DF[col+(X-col)++(Y-col)]
}

### Example 2: regular expressions (2022)

```
Regular expressions in Scala's standard library:
val rational = new Regex("(\\d+)\\.?(\\d+)?")
rational.unapply("3.1415"): Option[Seq[String]]

Revisited with type-level programming:
rational.unapply("3.1415"): Option[(String, Option[String]))
class Regex(pattern: String) {
  def unapply(s: String): Option[GroupsOf[pattern]]
}
```

### Example 2: regular expressions (2022)

```
Regular expressions in Scala's standard library:
val rational = new Regex("(\\d+)\\.?(\\d+)?")
rational.unapply("3.1415"): Option[Seq[String]]

Revisited with type-level programming:
rational.unapply("3.1415"): Option[(String, Option[String])]
class Regex(pattern: String) {
  def unapply(s: String): Option[GroupsOf[pattern]]
}
```

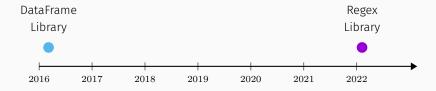
### Example 2: regular expressions (2022)

```
Regular expressions in Scala's standard library:
val rational = new Regex("(\\d+)\\.?(\\d+)?")
rational.unapply("3.1415"): Option[Seq[String]]

Revisited with type-level programming:
rational.unapply("3.1415"): Option[(String, Option[String])]

class Regex(pattern: String) {
  def unapply(s: String): Option[GroupsOf[pattern]]
}
```

### What changed between 2016 and 2022?



## What changed between 2016 and 2022?

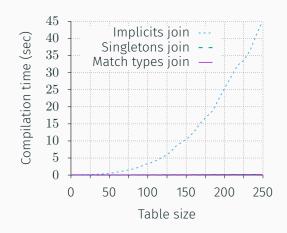
#### The Spark library uses a hack, implicits:

- · convoluted to use
- · slow to compile

The regex library uses a first-class language construct:

- · easier to use
- · faster to compile

# Compilation times for the type-level join operation







[1] Olivier Blanvillain. "Type-Safe Regular Expressions". In: Proc. ACM Scala Symposium. SCALA'22. Under submission.





[1] Georg Stefan Schmid, Olivier Blanvillain, Jad Hamza, and Viktor Kuncak. "Coming to Terms with Your Choices: An Existential Take on Dependent Types". In: Corr (2020). arXiv: 2011.07653.

# Singleton types

#### Definition

A singleton type is a type that contains exactly one value.

Scala has a few of those:

- $\cdot$  x.type, the type of the variable x (since forever)
- · 42, the type of the integer literal 42 (since 2016)
- · +, the type of integer addition (since 2020)

Question: How much more of the language can we represent as singleton types?

# Singleton types

#### Definition

A singleton type is a type that contains exactly one value.

Scala has a few of those:

- · x.type, the type of the variable x (since forever)
- 42, the type of the integer literal 42 (since 2016)
- +, the type of integer addition (since 2020)

Question: How much more of the language can we represent as singleton types?

# Singleton types

#### Definition

A singleton type is a type that contains exactly one value.

Scala has a few of those:

- · x.type, the type of the variable x (since forever)
- 42, the type of the integer literal 42 (since 2016)
- +, the type of integer addition (since 2020)

Question: How much more of the language can we represent as singleton types?

# Generalized singleton types: a proposal

#### Our proposal consists of 3 changes:

- new types for if-then-else, pattern matching, constructors, functions calls
- 2. "precise mode" of type inference
- 3. type evaluation, during subtyping

```
What's the type of if (x == 0) "zero" else "one"?

• String (Scala 2)

• "zero" | "one" (Scala 3)

• If[x.type == 0, "zero", "one"] (proposed)

• { if (x == 0) "zero" else "one" } (proposed, syntactic sugar)
```

```
What's the type of if (x == 0) "zero" else "one"?

• String (Scala 2)

• "zero" | "one" (Scala 3)

• If[x.type == 0, "zero", "one"] (proposed)

• { if (x == 0) "zero" else "one" } (proposed, syntactic sugar)
```

```
What's the type of if (x == 0) "zero" else "one"?

• String (Scala 2)

• "zero" | "one" (Scala 3)

• If[x.type == 0, "zero", "one"] (proposed)

• { if (x == 0) "zero" else "one" } (proposed, syntactic sugar)
```

```
What's the type of if (x == 0) "zero" else "one"?

• String (Scala 2)
• "zero" | "one" (Scala 3)
• If[x.type == 0, "zero", "one"] (proposed)
• { if (x == 0) "zero" else "one" } (proposed, syntactic sugar)
```

```
What's the type of if (x == 0) "zero" else "one"?

• String (Scala 2)

• "zero" | "one" (Scala 3)

• If[x.type == 0, "zero", "one"] (proposed)

• { if (x == 0) "zero" else "one" } (proposed, syntactic sugar)
```

Similarly, we add new types for other constructs: Match[], New[], Call[], TypeTest[], and the corresponding syntactic sugar.

```
What's the type of if (x == 0) "zero" else "one"?

• String (Scala 2)

• "zero" | "one" (Scala 3)

• If[x.type == 0, "zero", "one"] (proposed)

• { if (x == 0) "zero" else "one" } (proposed, syntactic sugar)
```

Similarly, we add new types for other constructs: Match[], New[], Call[], TypeTest[], and the corresponding syntactic sugar.

We need two modes of type inference, for backwards compatibility.

```
def int2str(x: Int) =
  if (x == 0) "zero" else "one"
```

We need two modes of type inference, for backwards compatibility.

```
def int2str(x: Int): String =
  if (x == 0) "zero" else "one"
```

We need two modes of type inference, for backwards compatibility.

```
dependent
def int2str(x: Int)
  : { if (x == 0) "zero" else "one" }
  = if (x == 0) "zero" else "one"
```

We need two modes of type inference, for backwards compatibility.

```
dependent
def int2str(x: Int)
  : If[x.type == 0, "zero", "one"]
  = if (x == 0) "zero" else "one"
```

We need two modes of type inference, for backwards compatibility.

```
dependent
def int2str(x: Int)
   : If[x.type == 0, "zero", "one"]
   = if (x == 0) "zero" else "one"
```

We need two modes of type inference, for backwards compatibility.

```
dependent
def int2str(x: Int)
  : If[x.type == 0, "zero", String]
  = if (x == 0) "zero" else readString()
```

We need two modes of type inference, for backwards compatibility.

```
dependent
def int2str(x: Int)
      : { if (x == 0) "zero" else (_: String) }
      = if (x == 0) "zero" else readString()
```

# "Precise mode" of type inference: list concatenation example

```
dependent def concat(xs: List, ys: List) <: List =
    xs match
    case x :: xs => x :: concat(xs, ys)
    case Nil => ys

dependent val l1 = "A" :: Nil
dependent val l2 = "B" :: Nil
dependent val l3 = concat(l1, l2)
l3: { "A" :: "B" :: Nil }
```

# "Precise mode" of type inference: list concatenation example

```
dependent def concat(xs: List, ys: List) <: List =
    xs match
    case x :: xs => x :: concat(xs, ys)
    case Nil => ys

val l1: List = "A" :: Nil
dependent val l2 = "B" :: Nil
dependent val l3 = concat(l1, l2)
l3: { concat(l1, l2) }
```

# "Precise mode" of type inference: list concatenation example

```
dependent def concat(xs: List, ys: List) <: List =
    xs match
    case x :: xs => x :: concat(xs, ys)
    case Nil => ys

dependent val l1 = "A" :: Nil
val l2: List = "B" :: Nil
dependent val l3 = concat(l1, l2)
l3: { "A" :: l2 }
```

### Type evaluation

During subtyping, we evaluate both sides:

Straightforward for if-then-else:

- eval(If[true, A, B]) = A
- eval(If[false, A, B]) = E

More interesting for pattern matching: we "desugar" pattern matching expressions into if-then-else & type tests.

### Type evaluation

During subtyping, we evaluate both sides:

```
· A <: B if eval(A) <: eval(B)
```

Straightforward for if-then-else:

- eval(If[true, A, B]) = A
- eval(If[false, A, B]) = B

More interesting for pattern matching: we "desugar" pattern matching expressions into if-then-else & type tests.

### Type evaluation

During subtyping, we evaluate both sides:

```
. A <: B if eval(A) <: eval(B)</pre>
```

Straightforward for if-then-else:

- eval(If[true, A, B]) = A
- eval(If[false, A, B]) = B

More interesting for pattern matching: we "desugar" pattern matching expressions into if-then-else & type tests.

### Type evaluation: type tests

We evaluate type tests using subtyping and type disjointness:

- eval(x.isInstanceOf[T]) = false if x.type and T are disjoint
- eval(x.isInstanceOf[T]) = x.isInstanceOf[T] otherwise

```
dependent val foo(x: Any) =
  x match
    case s: String => s
    case i: Int => i+1
foo(42): { 43 }
foo(readInt()): { (_: Int) + 1 }
```

```
dependent val foo(x: Any): {
   x match
    case s: String => s
    case i: Int => i+1
} =
   x match
   case s: String => s
   case i: Int => i+1

foo(42): { 43 }

foo(readInt()): { (_: Int) + 1 }
```

```
dependent val foo(x: Any): {
   if (x.isInstanceOf[String]) x.asInstanceOf[String]
   else if (x.isInstanceOf[Int]) x.asInstanceOf[Int] + 1
   else throw new MatchError()
} =
   x match
   case s: String => s
   case i: Int => i+1

foo(42): { 43 }
foo(readInt()): { (_: Int) + 1 }
```

```
dependent val foo(x: Any): {
   if (x.isInstanceOf[String]) x.asInstanceOf[String]
   else if (x.isInstanceOf[Int]) x.asInstanceOf[Int] + 1
   else throw new MatchError()
} =
   x match
   case s: String => s
   case i: Int => i+1

foo(42): { 43 }
foo(readInt()): { (_: Int) + 1 }
```

```
dependent val foo(x: Any): {
   if (x.isInstanceOf[String]) x.asInstanceOf[String]
   else if (x.isInstanceOf[Int]) x.asInstanceOf[Int] + 1
   else throw new MatchError()
} =
   x match
   case s: String => s
   case i: Int => i+1

foo(42): { 43 }
foo(readInt()): { (_: Int) + 1 }
```

## Generalized singletons: recap

We proposed a generalization of Scala's singleton types:

- 1. lift a subset of Scala's language constructs to the type level
- 2. add a "precise mode" of type inference
- 3. evaluate types during subtyping

### Part II: Match Types



### Part II: Match Types



[1] Olivier Blanvillain, Jonathan Immanuel Brachthäuser, Maxime Kjaer, and Martin Odersky. "Type-Level Programming with Match Types". In: <a href="https://example.com/Program.lang.">Proc. ACM Program. Lang.</a> POPL'22. ACM, 2022.

```
type Elem[X] = X match
  case String => Char
  case List[t] => Elem[t]
  case Any => X
```

```
type Elem[X] = X match
  case String => Char
  case List[t] => Elem[t]
  case Any => X
```

#### Some examples of reduction:

```
• Elem[String] =:= Char
```

```
• Elem[Int] =:= Int
```

- Elem[List[Int]] =:= Int
- Elem[Any] =:= Elem[Any]

```
type Elem[X] = X match
  case String => Char
  case List[t] => Elem[t]
  case Any => X
```

We evaluate type tests using subtyping and type disjointness:

- eval(X.isInstanceOf[T]) = true if X is a subtype of T
- eval(X.isInstanceOf[T]) = false if X and T are disjoint
- eval(X.isInstanceOf[T]) = X.isInstanceOf[T] otherwise

```
type Elem[X] = X match
  case String => Char
  case List[t] => Elem[t]
  case Any => X
```

Match types can be explained in terms of generalized singletons:

```
type Elem[X] = { (_: X) match
  case _: String => (_: Char)
  case _: List[t] => (_: Elem[t])
  case _: Any => (_: X)
}
```

```
type Elem[X] = X match
  case String => Char
  case List[t] => Elem[t]
  case Any => X
```

The type/term correspondence dictates match type semantics:

```
def elem[X](x: X): Elem[X] = x match
  case x: String => x.charAt(0)
  case x: List[t] => elem(x.head)
  case x: Any => x
```

```
type Elem[X] = X match
  case String => Char
  case List[t] => Elem[t]
  case Any => X
```

The type/term correspondence dictates match type semantics:

```
def elem[X](x: X): Elem[X] = x match
  case x: String => x.charAt(0)
  case x: List[t] => elem(x.head)
  case x: Any => x
elem[Any]("foo")
```

### **Formalization**

### System FM is an extension of System F<: with

- classes (defined externally)
- pattern matching
- match types

### Design goals:

- · explain the essence of match types
- give us confidence in our design
- · be as simple as possible

### **Formalization**

### System FM is an extension of System F<: with

- · classes (defined externally)
- pattern matching
- match types

### Design goals:

- $\cdot$  explain the essence of match types
- · give us confidence in our design
- · be as simple as possible

## System FM is parametric

System FM is parametrized by a set of externally defined classes:

 $FM(C, \Psi, \Xi)$ 

- · C := set of classes
- $\Psi \coloneqq$  class inheritance relation
- $\cdot \; \Xi \coloneqq \text{class disjointness relation}$

#### For example:

```
class A; class B extends A C = \{A, B\} \Psi = \{(B, A)\} \Xi = \varnothing class A; class D C = \{A, D\} \Psi = \varnothing \Xi = \{(A, D)\} class A; trait T C = \{A, T\} \Psi = \varnothing \Xi = \varnothing
```

## System FM is parametric

System FM is parametrized by a set of externally defined classes:

 $\mathsf{FM}(\mathsf{C},\,\Psi,\,\Xi)$ 

- $\cdot$  C := set of classes
- $\Psi :=$  class inheritance relation
- $\cdot \; \Xi \coloneqq \text{class disjointness relation}$

#### For example:

```
class A; class B extends A C = \{A, B\} \Psi = \{(B, A)\} \Xi = \varnothing class A; class D C = \{A, D\} \Psi = \varnothing \Xi = \{(A, D)\} class A; trait T C = \{A, T\} \Psi = \varnothing \Xi = \varnothing
```

# System FM's syntax

t ::=		T ::=	
X	variable	X	type variable
λx:T. t	abstraction	$T \rightarrow T$	type of functions
$\lambda X <: T. t$	type abstraction	∀X<:T. T	universal type
t t	application	Тор	maximum type
t T	type application	C	class
new C	constructor call	{ <i>new</i> C}	constructor singleton
t $match\{x:0\}$	$C \Rightarrow t$ } or t match expr.	$T match{T \Rightarrow T}$	or T match type
v ::=		Γ ::=	
λx:T. t	abstraction	Ø	empty context
$\lambda X <: T. t$	type abstraction	Γ,x:T	term binding
new C	constructor call	Γ,X<:T	type binding

## System FM's disjointness relation

$$\frac{(C_1,C_2) \in \Xi}{\Gamma \vdash \operatorname{disj}(C_1,C_2)} \qquad (D-\operatorname{XI}) \qquad \frac{(C_1,C_2) \notin \Psi}{\Gamma \vdash \operatorname{disj}(\{\operatorname{\textit{new}} C_1\},C_2)} \qquad (D-\operatorname{PSI})$$
 
$$\frac{\Gamma \vdash S <: U \quad \Gamma \vdash \operatorname{disj}(U,T)}{\Gamma \vdash \operatorname{disj}(S,T)} \qquad (D-\operatorname{SUB}) \qquad \Gamma \vdash \operatorname{disj}(\forall X <: T_1,T_2,C) \qquad (D-\operatorname{ALL})$$

### System FM's evaluation rules

We evaluate term-level matches by querying the class inheritance relation:

$$\frac{(\mathsf{C},\mathsf{C}_n) \in \Psi \quad \forall \, m < n. \; (\mathsf{C},\mathsf{C}_m) \notin \Psi}{new \, \mathsf{C} \quad match\{\mathsf{x}_i : \mathsf{C}_i \Rightarrow \mathsf{t}_i\} or \; \mathsf{t}_d \longrightarrow [\mathsf{x}_n \mapsto new \, \mathsf{C}] \mathsf{t}_n}$$

### System FM's evaluation rules

We evaluate term-level matches by querying the class inheritance relation:

$$\frac{(\mathsf{C},\mathsf{C}_n) \in \Psi \quad \forall \, m < n. \, (\mathsf{C},\mathsf{C}_m) \notin \Psi}{new \, \mathsf{C} \quad match\{\mathsf{x}_i : \mathsf{C}_i \Rightarrow \mathsf{t}_i\} or \, \, \mathsf{t}_d \longrightarrow [\mathsf{x}_n \mapsto new \, \mathsf{C}] \, \mathsf{t}_n}$$

We evaluate type-level matches using subtyping and disjointness:

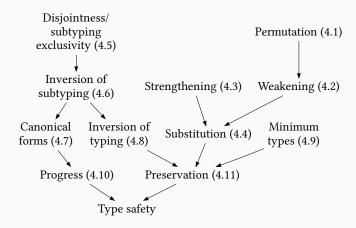
$$\frac{\Gamma \vdash T_s <: S_n \quad \forall m < n. \ \Gamma \vdash \text{disj}(T_s, S_m)}{\Gamma \vdash T_s \quad match\{S_i \Rightarrow T_i\} or \ T_d =:= T_n}$$

## System FM's type safety

We show type safety through the progress and preservation. Our proof comes in two versions:

- 1. Pen and paper (~30 pages)
- 2. Coq mechanization (~6000 LOC)

## Structure of the type safety proof



## Formalization vs implementation

Our implementation goes much further than System FM:

- recursion / non-termination
- type parameter bindings (case List[t])
- · empty types
- variance

### **Empty types**

#### Definition

An empty type is a type that contains no values.

Our system does not like empty types.

Nothing, in particular, is both subtype and disjoint from every type.

```
type M[X] = X match
  case Int => String
  case String => Int

class C:
  type X
  def f(bad: M[X & String]): Int = bad

class D extends C:
  type X = Int
```

### **Empty types**

#### Definition

An empty type is a type that contains no values.

Our system does not like empty types.

Nothing, in particular, is both subtype and disjoint from every type.

```
type M[X] = X match
  case Int => String
  case String => Int

tlass C:
  type X
  def f(bad: M[X & String]): Int = bad

tlass D extends C:
  type X = Int
```

### **Empty types**

#### Definition

An empty type is a type that contains no values.

Our system does not like empty types.

Nothing, in particular, is both subtype and disjoint from every type.

```
type M[X] = X match
  case Int => String
  case String => Int

class C:
  type X
  def f(bad: M[X & String]): Int = bad

class D extends C:
  type X = Int
```

### **Variance**

#### **Definitions**

```
F is covariant in T, written F[+T], if T1 <: T2 implies F[T1] <: F[T2]. G is contravariant in T, G[-T], if T1 <: T2 implies G[T1] >: G[T2].
```

We cannot prove disjointness when variance is involved, F[T1] and F[T2] always overlap:

We make an exception for covariant type that are used as class fields

```
case class Some[+A](value: A)
```

Some[String] and Some[Int] are disjoint, since there is no runtime value of type Some[Nothing].

### **Variance**

#### **Definitions**

```
F is covariant in T, written F[+T], if T1 <: T2 implies F[T1] <: F[T2]. G is contravariant in T, G[-T], if T1 <: T2 implies G[T1] >: G[T2].
```

We cannot prove disjointness when variance is involved, F[T1] and F[T2] always overlap:

```
\forall T1, T2. F[Nothing] <: F[T1] and F[Nothing] <: F[T2] 
\forall T1, T2. G[Any] <: G[T1] and G[Any] <: G[T2]
```

We make an exception for covariant type that are used as class fields

```
case class Some[+A](value: A)
```

Some[String] and Some[Int] are disjoint, since there is no runtime value of type Some[Nothing].

### **Variance**

#### **Definitions**

```
F is covariant in T, written F[+T], if T1 <: T2 implies F[T1] <: F[T2]. G is contravariant in T, G[-T], if T1 <: T2 implies G[T1] >: G[T2].
```

We cannot prove disjointness when variance is involved, F[T1] and F[T2] always overlap:

```
\forall T1, T2. F[Nothing] <: F[T1] and F[Nothing] <: F[T2] 
 \forall T1, T2. G[Any] <: G[T1] and G[Any] <: G[T2]
```

We make an exception for covariant type that are used as class fields:

```
case class Some[+A](value: A)
```

Some[String] and Some[Int] are disjoint, since there is no runtime value of type Some[Nothing].

### Match types: recap

We presented match types, a new language construct for type-level programming:

- · implemented in the Scala 3 compiler
- formalized in an extension of System F<:
- · already in active use

### Thank you!

