Modular implicits

Leo White Frédéric Bour January 4, 2015

Abstract

A common criticism of OCaml is its lack of support for ad-hoc polymorphism. The classic example of this is OCaml's separate addition operators for integers (+) and floating-point numbers (+.). Another example is the need for type-specific printing functions (print_int, print_string, etc.) rather than a single print function which works across multiple types. Taking inspiration from Scala's implicits and "Modular Type Classes" by Dreyer et al. [2] we propose a system for ad-hoc polymorphism in OCaml based on using modules as type-directed implicit parameters. We describe the design of this system, and compare it to systems for ad-hoc polymorphism in other languages.

1 Introduction

A common criticism of OCaml is its lack of support for ad-hoc polymorphism. The classic example of this is OCaml's separate addition operators for integers (+) and floating-point numbers (+.). Another example is the need for type-specific printing functions (print_int, print_string, etc.) rather than a single print function which works across multiple types.

1.1 Type classes

Type classes in Haskell [6] have proved an effective mechanism for supporting adhoc polymorphism. Type classes provide a form of constrained polymorphism, allowing constraints to be placed on type variables. For example, the show function has the following type:

```
show :: Show a => a -> String
```

This indicates that the type variable a can only be instantiated with types which obey the constraint Show a. These constraints are called *type classes*. The Show type class is defined as¹:

```
class Show a where
  show :: a -> String
```

¹Some methods of Show have been omitted for simplicity.

which specifies a list of methods which must be provided in order for a type to meet the Show constraint. The method implementations for a particular type are specified by defining an *instance* of the type class. For example, the instance of Show for Int is defined as:

```
instance Show Int where
    show = showSignedInt
```

Type classes in Haskell have a property called *coherence* – for each pair of type and type class there is at most one instance in the entire program. For example, there is at most one instance of **Show Int**, defining two such instances will result in a compiler error. Section 4.5.1 describes why this property cannot hold in OCaml.

Type classes are implemented using a type-directed implicit parameter passing mechanism. Essentially, each constraint on a type is treated as a parameter containing a dictionary of the methods of the type class. This parameter is implicit passed by the compiler using the appropriate type class instance.

1.2 Implicits

Implicits in Scala [4] provide similar capabilities to type classes via direct support for type-directed implicit parameter passing. Parameters can be marked implicit which then allows them to be omitted from method calls. For example, a show function could be specified as:

```
def show[T](x : T)(implicit s : Showable[T]): String
where Showable[T] is a regular Scala type defined as:
trait Showable[T] { def show(x: T): String }
The show function can be called just like any other:
show(7)(IntShowable)
```

However, the second argument can also be elided, in which case its value is selected from the set of definitions in scope which have been marked implicit. For example, given the following definition in scope:

```
implicit object IntShowable extends Showable[Int] {
  def show(x: Int) = x.toString
}
```

show can be called on integers without specifying the second parameter – which
will automatically be inserted as IntShowable because it has the required type
Showable[Int]:

```
show(7)
```

Unlike Haskell type classes, Scala implicits have no coherence property. You can define multiple implicit objects of type Showable[Int] in your program without causing an error. Instead Scala issues an error if the resolution of an

implicit parameter is ambiguous. For example, if two implicit objects of type Showable[Int] are in scope when show is applied to an Int then the compiler will emit an ambiguity error.

1.3 Modular implicits

Dreyer et al. [2] describe modular type classes, a type class system which uses ML module types as type classes and ML modules as type class instances. Taking inspiration from this, we propose a system for ad-hoc polymorphism in OCaml based on passing implicit module parameters to functions based on their module type. We argue that, due to the incompatibility of OCaml with the coherence requirements of type classes, implicits are a more appropriate basis for such a system.

We also argue that using modules as the basis for our implicits, rather than regular types, is a natural fit for implementing ad-hoc polymorphism. It also allows us to support advanced type class features such as associated types and constructor classes.

The rest of this paper is structured as follows: Section 2 describes the definition of functions with implicit module parameters, Section 3 describes which module definitions are available for selection as implicit parameters, Section 4 describes the search procedure for resolving implicit parameters, Section 5 discusses related work, and Section 6 discusses future work.

2 Implicit parameters

We propose adding a new kind of function parameter (implicit M : S) where M is an identifier and S is a module type. For example, the generic show function is written as follows:

```
module type Show = sig
  type t
  val show : t -> string
end

let show (implicit S : Show) x = S.show x
The type of show is written:
(implicit S : Show) -> S.t -> string
```

When the show function is applied the system attempts to fill-in the implicit S parameter by finding a module with the correct type. For example,

```
show [1.2; 2.3; 4.5]
```

will cause the compiler to search for a module with type Show with type t = float list, and pass it as the implicit argument to the show function.

2.1 Inferring implicit parameters

With type classes, the constraints on a function can be inferred from uses of other constrained functions within the function's definition. For example, if a show_twice function uses the show function:

```
show_twice x = show x ++ show x
```

then Haskell will infer that show_twice has type Show a => a -> String.

With modular implicits, all implicit parameters must be explicitly given by the programmer. Implicit parameters cannot be inferred from uses of other functions with implicit parameters, so **show_twice** would be written:

```
let show_twice (implicit S : Show) (x : S.t) =
  show x ^ show x
```

If the implicit parameter was removed from the definition of show_twice:

```
let show_twice x = show x ^ show x
```

then the implicit parameter to **show** would be resolved in the scope of the definition of **show_twice**, resulting in a function that was monomorphic in the type of x (and probably an ambiguity error).

These two function definitions behave differently because it matters in which scope an implicit parameter is resolved. Whereas, for type classes, it does not matter in which scope a resolution is performed because there can only be a single instance of the correct type in the whole program (due to coherence). This enables type class systems to treat both definitions as equivalent by inferring the presence of a constraint on the type of show_twice.

2.2 Higher-rank implicit parameters

Note that functions with implicit parameters can themselves be used as parameters to functions. For example, the following function:

```
let show_3 (sh : (implicit S : Show) -> S.t -> string) =
  sh 3
```

takes a function sh which expects an implicit parameter and applies it too 3. This can be applied to show:

```
show_3 show
```

In the above example the value of the implicit parameter to **show** is looked up in the scope of the *body* of **show_3**. This makes it quite different from the following example:

```
let show_3 (sh : int -> string) = sh 3
[...]
show_3 (fun (x : int) -> show x)
```

where the implicit parameter to show is looked up in the scope of the *call* to show_3.

This feature means that implicit parameters provide support for higher-rank polymorphism. For example, we can define the following function:

```
let show_stuff (sh : (implicit S : Show) -> S.t -> string) =
   sh 3 ^ " " ^ sh "hello"
```

This function takes a function that expects an implicit parameter S of type Show, and uses it once with S.t equal to int and once with S.t equal to string. This requires sh to be polymorphic in S.t, which, since s is a parameter, constitutes higher-rank polymorphism.

It is worth noting that OCaml already provides some support for higher-rank polymorphism through polymorphic record fields and methods.

2.3 Higher-kinded implicit parameters

Implicit parameters also provide support for higher-kinded polymorphism. For example, given a module type definition for monads:

```
module type Monad = sig
  type 'a t
  val return : 'a -> 'a t
  val bind : 'a t -> ('a -> 'b t) -> 'b t
end
```

We can define functions that are polymorphic in the monad they operate on:

```
let return (implicit M : Monad) x = M.return x
let (>>=) (implicit M : Monad) m k = M.bind m k
which have the following types:
val return : (implicit M : Monad) -> 'a -> 'a M.t
```

These types are polymorphic in the higher-kinded type constructor ${\tt M.t.}$

This corresponds to constructor classes in Haskell, and would introduce higher-kinded polymorphism into OCaml's core language. Previously, higher-kinded polymorphism was only supported in OCaml's verbose module language, so this feature is a genuine improvement in OCaml's expressivity.

2.4 Explicit implicit parameters

As with Scala implicits, searching for an implicit value can fail due to ambiguity (Section 4.1). In such cases, the implicit parameter can be given explicitly using the following syntax:

```
show (implicit ShowInt) 6
```

where ShowInt is a module which matches the signature Show with type t = int.

3 Implicit declarations

3.1 Implicit modules

Our proposal introduces a new form of module declaration, which makes the module available for use as an implicit parameter. These implicit declarations are regular module definitions annotated with the **implicit** keyword. For example, an implicit module for showing integers could be defined:

```
implicit module ShowInt = struct
  type t = int
  let show = string_of_int
end
```

This declaration both declares a regular module ShowInt and makes it available for use in implicit arguments.

Note that only "top-level" implicit modules are considered in scope: if module N contains an implicit sub-module M then M will not be in scope unless N has been opened. For convenience, we provide the syntax:

```
open implicit N
```

which brings the implicit sub-modules modules of N into the implicit scope.

3.2 Implicit parameters

In addition to implicit module declarations the system also considers any implicit parameters currently in scope. For example

```
let print (implicit S : Show) (x : S.t) =
  print_string (show x)
```

will use the implicit parameter S as the implicit parameter of the show function. Note that there is guaranteed to be no ambiguity here because the type S.t is abstract.

3.3 Implicit functors

Modules for use as implicit arguments can also be constructed using functors. For example, an implicit functor for showing lists could be defined:

```
implicit functor ShowList (S:Show) = struct
  type t = S.t list
  let show l = string_of_list S.show l
end
```

This means that if the system is looking for an implicit module with type Show with type t = int list, then it can instead look for an implicit module with type Show with type t = int, and apply ShowList to it to create the required module.

Such functors can be used multiple times whilst resolving a single implicit parameter, for example

```
show [ [1; 2; 3]; [4; 5; 6] ]
```

will resolve the implicit parameter of show to ShowList(ShowList(ShowInt))).

4 Resolving implicit parameters

4.1 Ambiguity

It is important that uses of implicit parameters are not ambiguous. If there are two implicit modules with type S in scope then using a function of type (implicit M:S) $\rightarrow \ldots$ is an error. It is not sufficient that the search for implicits finds a solution; it must also ensure that the solution is unique.

4.2 Backtracking

Unlike Haskell type classes, our system considers functor arguments when deciding if a resolution is ambiguous. For example, given a module type and function for printing:

```
module type Print = sig
  type t
  val print : t -> unit
end

let print (implicit P : Print) x = P.print x

with the following implicit definitions in scope:
implicit module PrintFloat = struct
  type t = float
  let print = print_float
end

implicit functor PrintOfShow (S : Show) = struct
  type t = S.t
  let print (x : t) = print_string (show x)
end
```

Then the ambiguity of print 1.5 depends on whether or not there exists an implicit module with type Show with type t = float.

In Haskell, these definitions are always considered ambiguous or, with overlapping instances enabled, they are always considered unambiguous. By checking the existence of a module of type Show with type t = float to use as the argument S of PrintOfShow, our system is able to decide ambiguity more accurately. Handling such definitions can be seen as a form of backtracking. The PrintOfShow functor is selected first but if an implicit module for its argument S cannot be found the search backtracks and instead selects PrintFloat.

4.3 Termination

Implicit functors can be used multiple times whilst resolving a single implicit parameter so we must be careful of non-termination in the resolution procedure. For example, a functor which tries to define how to show a type in terms of how to show that type:

```
implicit functor ShowIt (S:Show) = struct
  type t = S.t
  let show = show
end
```

is obviously not well-founded.

Type classes ensure the termination of resolution through a number of restrictions on instance declarations. However, termination of an implicit parameter resolution depends on the scope in which the resolution is performed. For this reason, our system places restrictions on the behaviour of the resolution directly and reports an error only when a resolution which breaks these restrictions is actually attempted.

To prevent non-terminating resolution all recursive uses of the same implicit functor the module type of the nested use must be strictly smaller than the module type of the outer use. Strictly smaller is defined point-wise on type definitions within the module type: all type definitions must be smaller or equal, and at least one type definition must be strictly smaller.

For example, if we use a functor to resolve a parameter of module type Show with type t = int list then any recursive use of that functor to resolve an argument must be for module type Show with type t = int because that is the only module type which is strictly smaller than the original.

Non-terminating resolutions are really a subclass of ambiguous resolutions: if the resolution procedure does not terminate then we do not know whether there may be multiple possible resolutions of an implicit parameter. For this reason, failure to meet the termination restrictions must be treated as an error, we cannot simply ignore the non-terminating possibilities and continue to look for an alternative resolution.

4.4 Existentials

OCaml modules can contain abstract types. This means that searches can be existentially quantified – we can ask for a type which can be shown (Show) rather than how to show a specific type (Show with type t = int).

Support for existential searches gives us similar features to Haskell's associated types. We can search for a module based on a subset of the types it

contains and the search will fill-in the remaining types for us. For example, consider a module type for arrays and function to create such arrays:

```
module type Array = sig
  type t
  type elem
  [...]
end

val create : (implicit A:Array) -> A.elem -> int -> A.array
```

where t is the type of an array and elem is the type of an array element. create can be used without specifying the array type being created:

```
let x = create true 5
```

This will search for an implicit Array with type elem = bool. When one is found x will correctly be given the associated t type. This allows specialised array types for different element types (e.g. bit vectors to represent arrays of bools).

4.5 Coherence

A key distinction between type classes and implicits is that with type classes you can only define a single instance of a type class for each type. In our system this would be equivalent to only allowing a single implicit module declaration for a given module type within a given program. This property is sometimes called coherence.

4.5.1 Coherence and abstraction

It would not be possible to have coherence in OCaml because type aliases can be made abstract. Consider the following example:

```
module F (X : Show) = struct
  implicit module S = X
end

implicit module ShowInt = struct
  type t = int
  let show = string_of_int

F(struct
  type t = int
  let show _ = "An int"
end)
```

The functor F defines an implicit Show module for the abstract type X.t, whilst the implicit module ShowInt is for the type int. However, F is later applied to a module where t is an alias for int. This violates coherence but this violation is hidden by abstraction.

Whilst it may seem that such cases can be detected by peering through abstractions, this is not possible in general and defeats the entire purpose of abstraction. Fundamentally, coherence is not a modular property and cannot be respected by a language with full support for modular abstraction.

4.5.2 A benefit of coherence

An often cited example of why coherence is useful is the union function for sets in Haskell. The Ord type class defines an ordering for a type²:

```
class Ord a where
  (<=) :: a -> a -> Bool
```

this ordering is used to create sets implemented as binary trees:

```
data Set a
empty :: Set a
insert :: Ord a => a -> Set a -> Set a
delete :: Ord a => a -> Set a -> Set a
```

The union function gives the union of two sets:

```
union :: Ord a => Set a -> Set a -> Set a
```

Efficiently implementing this union requires both sets to have been created using the same ordering. This is ensured by coherence, since there is only one instance of Ord a for each a, and all sets of type Set a must have been created using it.

4.5.3 An alternative to coherence

In terms of our system, Haskell's union function would have type:

but without coherence it is not safe for us to give union this type since we cannot ensure that all sets of a given type are using the same ordering.

The issue is that the **set** type is only parametrised by the type of its elements, when it should really be parametrised by the ordering used to create it. Traditionally, this problem is solved in OCaml by using applicative functors:

```
module Set (0 : Ord) : sig
  type elt
  type t
  val empty : t
  val add : elt -> t -> t
  val remove : elt -> t -> t
```

 $^{^2\}mathrm{Some}$ details of \mathtt{Ord} are omitted for simplicity

```
val union : t -> t -> t
[...]
end
```

When applied to an Ord argument O, the Set functor produces a module containing the following functions:

```
val empty : Set(0).t
val add : elt -> Set(0).t -> Set(0).t
val remove : elt -> Set(0).t -> Set(0).t
val union : Set(0).t -> Set(0).t -> Set(0).t
```

We can use the same approach with implicits by giving our polymorphic set operations the following types:

The type for sets is now Set(0).t which is parametrised by the ordering module 0, ensuring that union is only applied to sets using the same ordering.

5 Related work

5.1 Type classes

Haskell type classes [6] are the classic formalised method for ad-hoc polymorphism. They have been replicated in a number of other programming languages (e.g. Coq's type classes, Agda's instance arguments, Rust's traits). The key difference between approaches based on type class and approaches based on implicits is the property of coherence. One notable difference, due to coherence, is that type class constraints can be inferred whilst implicit parameters must always be given explicitly.

Since coherence is not possible in a language, like OCaml, which supports modular abstraction, type classes are not always a viable choice. Coherence is also not always desirable: the restriction to a single instance per type is not compositional and can force users to create additional types to work around it.

Other differences between our proposal and type classes include support for backtracking during parameter resolution – allowing for more precise detection of ambiguity, and resolution based on any type defined within the module rather than on a single specific type.

5.2 Implicits

Scala implicits [4] are an obvious inspiration for this work. They provide implicit parameters on functions, which are selected from the scope of the call site based on their type. In Scala these parameters have regular Scala types, whilst we propose using module types. Scala's object system has many properties in common with a module system, so advanced features such as associated types are still possible despite Scala's implicits being based on regular types.

Scala's implicits have a much more complicated notion of scope than our proposal. This seems to be aimed at fitting implicits into Scala's object-oriented approach: for example allowing implicits to be searched for in companion objects of the class of the implicit parameter. This makes it more difficult to answer the question "Where is the implicit parameter coming from?", in turn making it more difficult to reason about code. Our proposal simply uses regular lexical scope when searching for an implicit parameter.

Scala supports overlapping implicit instances. If an implicit parameter is resolved to more than one definition, rather than give an ambiguity error, a complex set of rules gives an ordering between definitions, and a most specific definition will be selected. An ambiguity error is only given if multiple definitions are considered equally specific. This can be useful, but makes reasoning about implicit parameters much more difficult: to know which definition is selected you must know all the definitions in the current scope. Our proposal always gives an ambiguity error if multiple implicit modules are available.

In addition to implicit parameters, Scala also supports implicit conversions. If a method is not available on an object's type the implicit scope is searched for a function to convert the object to a type on which the method is available. This feature greatly increases the complexity of finding a method's definition, and is not supported in our proposal.

Chambart et al. have proposed [1] adding support for implicits to OCaml using regular OCaml types for implicit parameters. Our proposal instead uses module types for implicit parameters. This allows our system to support more advanced features including associated types and higher-kinds. The module system also seems a more natural fit for ad-hoc polymorphism due to its direct support for signatures.

The implicit calculus [5] provides a minimal and general calculus of implicits which could serve as a basis for formalising many aspects of our proposal.

5.3 Canonical structures

In addition to type classes, Coq also supports a mechanism for ad-hoc polymorphism called *canonical structures*. Type classes and implicits provide a mechanism to resolve a value based on type information. Coq, being dependently typed, already uses unification to resolve values from type information, so canonical structures support ad-hoc polymorphism by providing additional ad-hoc rules that are applied during unification.

Canonical structures rely on coherence, but unlike type classes they do not

operate on a single specific type: ad-hoc unification rules are created for every type or term defined in the structure. Canonical structures also support backtracking of their search due to the backtracking built into Coq's unification.

5.4 Concepts

Gregor et al. [3] describe *concepts*, a system for ad-hoc polymorphism in $C++^3$. C++ has traditionally used simple overloading to support ad-hoc polymorphism restricted to monomorphic uses. C++ also supports parametric polymorphism through templates. However, overloading within templates is re-resolved after template instantiation. This means that the combination of overloading and templates provides full ad-hoc polymorphism. However, since much type-checking is delayed until template instantiation, C++ has very slow compilation times and very confusing error messages.

Concepts provide a disciplined approach to full ad-hoc polymorphism through an approach similar to type classes and implicits. Like type classes, a new kind of type is used to constrain parametric type variables. New concepts are defined using a <code>concept</code> construct, and instances of a concept are defined using the <code>concept_map</code> construct. Like implicits, concepts do not have a coherence property.

Concepts allow overlapping instances, using C++'s complex overloading rules to resolve ambiguities. Concepts also attempt to infer instances for types based on their definition if the required <code>concept_map</code> is not defined. These features can be useful, but make reasoning about implicit parameters much more difficult. Our proposal requires all implicit modules to be explicit and always gives an ambiguity error if multiple implicit modules are available.

5.5 Modular type classes

Dreyer et al. [2] describe modular type classes, a type class system which uses ML module types as type classes and ML modules as type class instances. This system sticks closely to the design of Haskell type classes, in particular it gives ambiguity and termination errors at the point when modules are made implicit, and infers type class constraints from uses of constrained functions. In order to do this it requires a number of undesirable restrictions:

- All implicit modules must define a type named t and resolution is always done based on this type.
- Modules may only be made implicit at the top-level, they cannot be introduced within a module or a value definition.
- Only module definitions are permitted at the top-level, all value definitions must be contained within a sub-module.

 $^{^3{\}rm This}$ should not be confused with more recent "concepts lite" proposals, due for inclusion in the next C++ standard

• All top-level module definitions must have an explicit signature.

By basing our model on implicits rather than type classes we avoid such restrictions, and the result is a cleaner design.

Our proposal also includes higher-rank implicit parameters, higher-kinded implicit parameters and resolution based on multiple types (the equivalent of multi-parameter type classes). These are not included in the design of modular type classes.

6 Future work

This paper focuses on the design of modular implicits and the properties of the resolution procedure. It does not describe the details of the type system, and giving a formal description of this type system is left as future work. In particular, the typing of higher-rank and higher-kinded implicit parameters, whilst related to previously published type systems, is novel enough to be worthy of a full formal treatment.

A prototype implementation based on OCaml 4.02.0 has been created and is available through the OPAM package manager⁴. Future work is needed to bring this prototype up to production quality.

Acknowledgements

We would like to thank Jeremy Yallop and Anil Madhavapeddy for helpful discussions.

References

- [1] Pierre Chambart & Grégoire Henry (2012): Experiments in Generic Programming. OCaml Users and Developers Workshop.
- [2] Derek Dreyer, Robert Harper, Manuel M. T. Chakravarty & Gabriele Keller (2007): Modular type classes. In Martin Hofmann & Matthias Felleisen, editors: Proceedings of the 34th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2007, Nice, France, January 17-19, 2007, ACM, pp. 63–70, doi:10.1145/1190216.1190229. Available at http://doi.acm.org/10.1145/1190216.1190229.
- [3] Douglas Gregor, Jaakko Järvi, Jeremy G. Siek, Bjarne Stroustrup, Gabriel Dos Reis & Andrew Lumsdaine (2006): Concepts: linguistic support for generic programming in C++. In Peri L. Tarr & William R. Cook, editors: Proceedings of the 21th Annual ACM SIGPLAN Conference on Object-Oriented Programming, Systems, Languages,

⁴opam switch 4.02.0+modular-implicits

- and Applications, OOPSLA 2006, October 22-26, 2006, Portland, Oregon, USA, ACM, pp. 291-310, doi:10.1145/1167473.1167499. Available at http://doi.acm.org/10.1145/1167473.1167499.
- [4] Bruno C. d. S. Oliveira, Adriaan Moors & Martin Odersky (2010): Type classes as objects and implicits. In William R. Cook, Siobhán Clarke & Martin C. Rinard, editors: Proceedings of the 25th Annual ACM SIGPLAN Conference on Object-Oriented Programming, Systems, Languages, and Applications, OOPSLA 2010, October 17-21, 2010, Reno/Tahoe, Nevada, USA, ACM, pp. 341-360, doi:10.1145/1869459.1869489. Available at http://doi.acm.org/10.1145/1869459.1869489.
- [5] Bruno C. d. S. Oliveira, Tom Schrijvers, Wontae Choi, Wonchan Lee & Kwangkeun Yi (2012): The implicit calculus: a new foundation for generic programming. In Jan Vitek, Haibo Lin & Frank Tip, editors: ACM SIGPLAN Conference on Programming Language Design and Implementation, PLDI '12, Beijing, China June 11 16, 2012, ACM, pp. 35-44, doi:10.1145/2254064.2254070. Available at http://doi.acm.org/10.1145/2254064.2254070.
- [6] Philip Wadler & Stephen Blott (1989): How to Make ad-hoc Polymorphism Less ad-hoc. In: Conference Record of the Sixteenth Annual ACM Symposium on Principles of Programming Languages, Austin, Texas, USA, January 11-13, 1989, ACM Press, pp. 60-76, doi:10.1145/75277.75283. Available at http://doi.acm.org/10.1145/75277.75283.