Monad Transformers and Modular Interpreters*

Sheng Liang  Paul Hudak  Mark Jones†

Yale University
Department of Computer Science
New Haven, CT 06520-8285

{liang,hudak,jones-mark}@cs.yale.edu

Abstract

We show how a set of building blocks can be used to construct programming language interpreters, and present implementations of such building blocks capable of supporting many commonly known features, including simple expressions, three different function call mechanisms (call-by-name, callby-value and lazy evaluation), references and assignment, nondeterminism, first-class continuations, and program tracing.

The underlying mechanism of our system is monad transformers, a simple form of abstraction for introducing a wide range of computational behaviors, such as state, I/O, continuations, and exceptions.

Our work is significant in the following respects. First, we have succeeded in designing a fully modular interpreter based on monad transformers that includes features missing from Steele’s, Espinosa’s, and Wand’s earlier efforts. Second, we have found new ways to lift monad operations through monad transformers, in particular difficult cases not achieved in Moggi’s original work. Third, we have demonstrated that interactions between features are reflected in liftings and that semantics can be changed by reordering monad transformers. Finally, we have implemented our interpreter in Gofer, whose constructor classes provide just the added power over Haskell’s type classes to allow precise and convenient expression of our ideas. This implementation includes a method for constructing extensible unions and a form of subtyping that is interesting in its own right.

1 Introduction and Related Work

This paper discusses how to construct programming language interpreters out of modular components. We will show how an interpreter for a language with many features can be composed from building blocks, each implementing a specific feature. The interpreter writer is able to specify the set of incorporated features at a very high level.

The motivation for building modular interpreters is to isolate the semantics of individual programming language features for the purpose of better understanding, simplifying, and implementing the features and their interactions. The lack of separability of traditional denotational semantics [19] has long been recognized. Algebraic approaches such as Mosses’ action semantics [16], and related efforts by Lee [13], Wand [23], Appel & Jim [1], Kelsey & Hudak [11], and others, attempt to solve parts of this problem, but fall short in several crucial ways.¹

A ground-breaking attempt to better solve the overall problem began with Moggi’s [15] proposal to use monads to structure denotational semantics. Wadler [21] popularized Moggi’s ideas in the functional programming community by showing that many type constructors (such as List) were monads and how monads could be used in a variety of settings, many with an “imperative” feel (such as in Peyton Jones & Wadler [17]). Wadler’s interpreter design, however, treats the interpreter monad as a monolithic structure which has to be reconstructed every time a new feature is added. More recently, Steele [18] proposed pseudomonads as a way to compose monads and thus build up an interpreter from smaller parts, but he failed to properly incorporate important features such as an environment and store, and struggled with restrictions in the Haskell [7] type system when trying to implement his ideas. In fact, pseudomonads are really just a special kind of monad transformer, first suggested by Moggi [15] as a potential way to leave a “hole” in a monad for further extension.

Returning to Moggi’s original ideas, Espinosa [4] nicely formulated in Scheme a system called Semantic Lego — the first modular interpreter based on monad transformers — and laid out the issues in lifting. Espinosa’s work reminded the programming language community (including us) — who had become distracted by the use of monads — that Moggi himself, responsible in many ways for the interest in monadic programming, had actually focussed more on the importance of monad transformers.

We begin by realizing the limitations of Moggi’s framework and Espinosa’s implementation, in particular the difficulty in dealing with complicated operations such as callcc, and investigate how common programming language fea-

¹Very recently, Cartwright and Fellesiesen [3] have independently proposed a modular semantics emphasizing a direct semantics approach, which seems somewhat more complex than ours; the precise relationship between the approaches is, however, not yet clear.
tures interact with each other. In so doing we are able to express more modularity and more language features than in previous work, solving several open problems that arose not only in Moggi’s work, but in Steele’s and Espinosa’s as well. Our work also shares results with Jones and Duponcheel’s [10] work on composing monads.

Independently, Espinosa [5] has continued working on monad transformers, and has also recognized the limitations of earlier approaches and proposed a solution quite different from ours. His new approach relies on a notion of “higher-order” monads (called situated monads) to relate different layers of monad transformers, and he has investigated the semantic implications of the order of monad transformer composition. It is not yet clear how his new approach relates to ours.

We use Gofer [8] syntax, which is very similar to Haskell’s, throughout the paper. We choose Gofer over Haskell because of its extended type system, and we choose a functional language over mathematical syntax for three reasons: (1) it is just about as concise as mathematical syntax, (2) it emphasizes the fact that our ideas are implementable (and thus have been debugged!), and (3) it shows how the relatively new idea of constructor classes [9] can be used to represent some rather complex typing relationships. Of course, monads can be expressed in a variety of other (higher-order) programming languages, in particular SML [14], whose type system is equally capable of expressing some of our ideas. The system could also be expressed in Scheme, but of course we would then lose the benefits of strong static type-checking. Our Gofer source code is available via anonymous ftp from nebula.cs.yale.edu in the directory pub/yale-fp/modular-interpreter.

To appreciate the extent of our results, Figure 1 gives the high-level definition of an interpreter, which is constructed in a modular way, and supports arithmetic, three different kinds of functions (call-by-name, call-by-value, and lazy), references and assignment, nondeterminism, first-class continuations, and tracing. The rest of the paper will provide the details of how the type declarations expand into a full interpreter and how each component is built. For now just note that OR is equivalent to the domain sum operator, and Term, Value and InterpM denote the source-level terms, runtime values, and supporting features (which can be regarded as the run-time system), respectively. Int and Fun are the semantic domains for integers and functions. TermA, TermF, TermR, TermL, TermT, TermC, TermN, etc. are the abstract syntax for arithmetic terms, function

\[ \text{type Value} = \text{OR Int (OR Fun (j))} \]

\[ \text{type InterpM} = \text{StateT Store (OR Int (OR Fun (j))) \text{EnvT Env (ContT Answer) StateT String (ErrorT (List ())))} \]

\[ \text{type Term} = \text{OR TermA - arithmetic (OR TermF - functions (OR TermR - assignment (OR TermL - lazy evaluation (OR TermT - tracing (OR TermC - callcc (TermN - nondeterminism))))))} \]

\[ \text{type InterpM} = \text{StateT Store - memory cells (EnvT Env - environment (ContT Answer - continuations (StateT String - trace output (ErrorT - error reporting (List - multiple results))))}} \]

\[ \text{interp :: Term \rightarrow InterpM Value} \]

\[ \text{fig: interp} \]

Figure 1: A modular interpreter

Although (for lack of space) we do not include any proofs, all constructs (monads, monad transformers and liftings) expressed as Gofer code have been verified to satisfy the necessary properties stated in this paper.

expressions, etc. Type constructors such as StateT and ContT are monad transformers; they add features, and are used to transform the monad List into the monad InterpM used by the interpreter.

To see how Term, Value, and InterpM constitute to modular interpreters, in the next section we will walk through some simple examples.

2 An Example

A conventional interpreter maps, say, a term, environment, and store, to an answer. In contrast, a monadic interpreter such as ours maps terms to computations, where the details of the environment, store, etc. are “hidden”. Specifically:

\[ \text{interp :: Term \rightarrow InterpM Value} \]

where “InterpM Value” is the interpreter monad of final answers.

What makes our interpreter modular is that all three components above — the term type, the value type, and the monad — are configurable. To illustrate, if we initially wish to have an interpreter for a small arithmetic language, we can fill in the definitions as follows:

\[ \text{type Value} = \text{OR Int ()} \]
\[ \text{type Term} = \text{TermA} \]
\[ \text{type InterpM} = \text{ErrorT Id} \]

The first line declares the answer domain to be the union of integers and the unit type (used as the base type). The second line defines terms as TermA, the abstract syntax for arithmetic operations. The final line defines the interpreter monad as a transformation of the identity monad \( \text{Id} \). The monad transformer ErrorT accounts for the possibility of errors; in this case, arithmetic exceptions.

At this point the interpreter behaves like a calculator:

\[ (\text{interp } ((1 + 4) \times 8)) \]
\[ 40 \]
\[ \text{ERROR: divide by 0} \]

Now if we wish to add function calls, we can extend the value domain with function types, add the abstract syntax for function calls to the term type, and apply the monad transformer EnvT to introduce an environment Env.

\[ \text{type Value} = \text{OR Int (OR Fun (j))} \]
\[ \text{type Term} = \text{OR TermF TermA} \]
\[ \text{type InterpM} = \text{EnvT Env (ErrorT Id)} \]

\[ ^3 \text{For lack of space, we omit the details of parsing and printing.} \]
Here is a test run:

\[
\begin{align*}
& (x + x + 4) 7 \\
& \text{ERROR: unbound variable: } x
\end{align*}
\]

By adding other features, we can arrive at (and go beyond) the interpreter in Figure 1. In the process of adding new source-level terms, whenever a new value domain (such as Boolean) is needed, we extend the Value type, and to add a new semantic feature (such as a store or continuation), we apply the corresponding monad transformer.

**Why monads?** In a sense, monads are nothing more than a good example of data abstraction. But they just happen to be a particularly good abstraction, and by using them in a disciplined (and appropriate) way, we generally obtain well-structured, modular programs. In our application, they are surprisingly useful for individually capturing the essence of a wide range of programming language features, while abstracting away from low-level details. Then with monad transformers we can put the individual features together, piece-by-piece in different orders, to create full-featured interpreters.

### 3 The Constructor Class System

For readers not familiar with the Gofe type system (in particular, constructor classes [9]), this section provides a motivating example.

Constructor classes support abstraction of common features among type constructors. Haskell, for example, provides the standard `map` function to apply a function to each element of a given list:

\[
\text{map} : (a \rightarrow b) \rightarrow [a] \rightarrow [b]
\]

Meanwhile, we can define similar functions for a wide range of other datatypes. For example:

```haskell
instance Functor Tree where
  map f Leaf = Leaf f
  map f (Node l r) = Node (map f l) (map f r)
```

The `mapTree` function has similar type and functionality to those of `map`. With this in mind, it seems a shame that we have to use different names for each of these variants. Indeed, Gofe allows type variables to stand for type constructors, on which the Haskell type class system has been extended to support overloading. To solve the problem with `map`, we can introduce a new constructor class `Functor` (in a categorical sense):

```haskell
class Functor f where
  map f = \( \) \rightarrow f a \rightarrow f b
```

Now the standard list (List) and the user-defined type constructor Tree are both instances of `Functor`:

```haskell
instance Functor List where
  map f [] = []
  map f (x : xs) = f x : map f xs
```

```haskell
instance Functor Tree where
  map f (Leaf x) = Leaf (f x)
  map f (Node l r) = Node (map f l) (map f r)
```

In building modular interpreters, we will find constructor classes extremely useful for dealing with multiple instances of monads and monad transformers (which are all type constructors).

### 4 Extensible Union Types

We begin with a discussion of a key idea in our framework: how values and terms may be expressed as extensible union types. (This facility has nothing to do with monads.)

The disjoint union of two types is captured by the datatype OR.

```haskell
data OR a b = L a | R b
```

where \( L \) and \( R \) are used to perform the conventional injection of a summand type into the union; conventional pattern-matching is used for projection. However, such injections and projections only work if we know the exact structure of the union; in particular, an extensible union may be arbitrarily nested, and we would like a single pair of injection and projection functions to work on all such constructions.

To achieve this, we define a type class to capture the summand/union type relationship, which we refer to as a “subtype” relationship:

```haskell
class SubType a b sup where
  inj :: sub \rightarrow sup -- injection
  prj :: sup \rightarrow Maybe sub -- projection
```

The `Maybe` datatype is used because the projection function may fail. We can now express the relationships that we desire:

```haskell
instance SubType a (OR a b) where
  inj = \( L \)
  prj (L x) = Just x
  prj _ = Nothing
```

```haskell
instance SubType a b \Rightarrow SubType a (OR c b) where
  inj = R \cdot inj
  prj (R a) = prj a
  prj _ = Nothing
```

Now we can see, for example, how the Value domain used in the interpreter example given earlier is actually constructed:

```haskell
type Value = OR Int (OR Fun ())
type Fun = InterpM Value \rightarrow InterpM Value
```

With these definitions the Gofe type system will infer that `Int` and `Function` are both “subtypes” of `Value`, and the coercion functions `inj` and `prj` will be generated automatically.\(^4\) (Note that the representation of a function is quite general — it maps computations to computations. As will be seen, this generality allows us to model both call-by-name and call-by-value semantics.)

\(^4\)We should point out here that most of the typing problems Steele encountered disappear with the use of our extensible union types; in particular, there is no need for Steele’s “towers” of datatypes.
5 The Interpreter Building Blocks

As in the example of Section 2, the Term type is also constructed as an extensible union (of subterm types). We define additionally a class InterpC to characterize the term types that we wish to interpret:

class InterpC t where
  interp :: t -> InterpM Value

The behavior of interp on unions of terms is given in the obvious way:

instance (InterpC t1, InterpC t2) => InterpC (OR t1 t2) where
  interp (L t) = interp t
  interp (R t) = interp t

The interp function mentioned in the opening example is just the method associated with the top-level type Term.

In the remainder of this section we define several representative interpreter building blocks, each an instance of class InterpC and written in a monadic style. We will more formally define monads later, but for now we note that the interpreter monad InterpM comes equipped with two basic operations:

unit :: a -> InterpM a
bind :: InterpM a -> (a -> InterpM b) -> InterpM b

Intuitively, InterpM a denotes a computation returning a result of type a. "Unit x" is a null computation that just returns x as result, whereas "m 'bind' k" runs m and passes the result to the rest of the computation k. As will be seen, besides unit and bind, each interpreter building block has several other operations that are specific to its purpose.

5.1 The Arithmetic Building Block

Our (very tiny) arithmetic sublanguage is given by:

data TermA = Num Int | Add Term Term

whose monadic interpretation is given by:

instance InterpC TermA where
  interp (Num x) = unitInj x
  interp (Add x y) = interp x 'bindPrj' \ i -> interp y 'bindPrj' \ j -> unitInj ((i+j)::Int)

unitInj :: String -> InterpM a
m 'bindPrj' k = unitInj k
  m 'bind' a -> m 'bind' k \ a -> m 'bindPrj' \ k =
  case (prj a) of
    Just x -> k x
    Nothing -> error "run-time type error"

err :: String -> InterpM a
  err = defined later

Note the simple use of inj and prj to inject/project the integer result into/out of the Value domain, regardless of how Value is eventually defined (unitInj and bindPrj make this a tad easier, and will be used later as well). Err is an operation for reporting errors to be defined later.

5.2 The Function Building Block

Our “function” sublanguage is given by:

    data TermF = Var Name
                  | LambdaN Name Term
                  | LambdaV Name Term
                  | App Term Term

which supports two kinds of abstractions, one for call-by-name, the other for call-by-value.

We assume a type Env of environments that associates variable names with computations (corresponding to the “closure” mode of evaluation [2]), and that has two operations:

lookupEnv :: Name -> Env -> Maybe (InterpM Value)
extendEnv :: (Name, InterpM Value) -> Env -> Env

The interpretation of the applicable sublanguage is then given in Figure 2.

The difference between call-by-value and call-by-name is clear: the former reduces the argument before evaluating the function body, whereas the latter does not. In a function application, the function itself is evaluated first, and bindPrj checks if it is indeed a function. The computation of e2 is packaged up with the current environment to form a closure, which is then passed to f. We could just as easily realize dynamic scoping by passing not the closure, but the computation of e2 alone.

When applying a call-by-value function, we build a computation which gets evaluated immediately upon entering the function body. Although semantically correct, this does not correspond to an efficient implementation. In practice, however, we expect that the presence of some kind of type information or a special syntax for call-by-value application will enable us to optimize away this overhead.

We note that Steele felt it unsatisfactory that his interpreter always had an environment argument, even though it was only used in the function building block. By abstracting environment-related operations as two functions (inEnv and rdEnv), we achieve exactly what Steele wished for.

5.3 The References and Assignment Building Block

A sublanguage of references and assignment is given by:

    data TermR = Ref Term
                | Deref Term
                | Assign Term Term

Given a heap of memory cells and three functions for managing it:

    allocLoc :: InterpM Loc
    lookupLoc :: Loc -> InterpM Value
    updateLoc :: (Loc, InterpM Value) -> InterpM ()

we can then give an appropriate interpretation to the new language features:
instance InterpC TermF where
interp (Var v) = rdEnv 'bind' \env →
case lookupEnv v \env of
  Just val → val
  Nothing → err ("unbound variable: " ++ v)
interp (LambdaN s t) = rdEnv 'bind' \env →
  unitInj (\arg → inEnv (extendEnv (s, arg) \env) (interp t))
interp (LambdaV s t) = rdEnv 'bind' \env →
  unitInj (\arg → arg 'bind' \v →
    inEnv (extendEnv (s, unit v) \env) (interp t))
interp (App e1 e2) = interp e1 'bindPrj' \f →
  rdEnv 'bind' \env →
  f (inEnv env (interp e2))

Figure 2: The function building block

instance InterpC TermR where
interp (Ref x) =
  interp x 'bindPrj' \loc →
  allocLoc 'bind' \loc →
  updateLoc (loc, unit val) 'bind' \w →
  unitInj loc
interp (Deref x) =
  interp x 'bindPrj' \loc →
  lookupLoc loc
interp (Assign lhs rhs) =
  interp lhs 'bindPrj' \loc →
  interp rhs 'bind' \val →
  updateLoc (loc, unit val) 'bind' \w →
  unit val

5.4 A Lazy Evaluation Building Block

Using this same heap of memory cells for references, we can implement “lazy” abstractions:

data TermL = LambdaL Name Term

whose operational semantics implies “caching” of results.

instance InterpC TermL where
interp (LambdaL s t) =
  rdEnv 'bind' \env →
  unitInj (\arg →
    allocLoc 'bind' \loc →
    updateLoc (loc, unit val) 'bind' \w →
    unitInj loc
  )

  let thunk = arg 'bind' \v →
  updateLoc (loc, unit val) 'bind' \w →
  unit val

  in
  updateLoc (loc, thunk) 'bind' \w →
  inEnv (extendEnv (s, lookupLoc loc) \env)
  (interp t)

Upon entering a lazy function, the interpreter first allocates a memory cell and stores a thunk (updatable closure) in it. When the argument is first evaluated in the function body, the interpreter evaluates the thunk and stores the result back into the memory cell, overwriting the thunk itself.

5.5 A Program Tracing Building Block

Given a function:

write :: String → InterpM ()

which writes a string output and continues the computation, we can define a “tracing” sublanguage, which attaches labels to expressions which cause a “trace record” to be invoked whenever that expression is evaluated:

data TermT = Trace String Term

instance InterpC TermT where
interp (Trace l t) =
  write ("enter " ++ l) 'bind' \_ →
  interp t 'bind' \v →
  write ("leave " ++ l ++ " with:" ++ show v) 'bind' \_ →
  unit v

Here we see that some of the features in Kishon et al.'s system [12] are easily incorporated into our interpreter.

5.6 The Continuation Building Block

First-class continuations can be included in our language with:

data TermC = CallCC

Using the callcc semantic function (to be defined later):

callcc :: ((a → InterpM b) → InterpM a) → InterpM a

we can give an interpretation for CallCC:

instance InterpC TermC where
interp CallCC = unitInj (\f →
  f 'bindPrj' \f' →
  callcc (\k → (f' (unitInj (\a → a 'bind' k)))))

CallCC is interpreted as a (strict) builtin function. Interp in this case does nothing more than inject and project values to the right domains.
5.7 The Nondeterminism Building Block

Our nondeterministic sublanguage is given by:

\[ \text{data TermN} = \text{Amb} [\text{Term}] \]

Given a function:

\[ \text{merge} : [\text{InterpM a}] \rightarrow \text{InterpM a} \]

which merges a list of computations into a single (nondeterministic) computation, nondeterminism interpretation can be expressed as:

\[ \text{instance InterpC TermN where} \]

\[ \text{interp (Amb f)} = \text{merge (map interp f)} \]

6 Monads With Operations

As mentioned earlier, particular monads have other operations besides \text{unit} and \text{bind}. Indeed, from the last section, it is clear that operations listed in Table 1 must be supported.

If we were building an interpreter in the traditional way, now is the time to set up the domains and implement the functions listed in the table. The major drawback of this monolithic approach is that we have to take into account all other features when we define an operation for one specific feature. When we define \text{callcc}, for example, we have to decide how it interacts with the store and environment etc. And if we later want to add more features, the semantic domains and all the functions in the table will have to be updated.

\text{Monad transformers}, on the other hand, allow us to individually capture the essence of language features. Furthermore, the concept of \text{lifting} allows us to account for the interactions between various features. These are the topics of the next two sections.

To simplify the set of operations somewhat, we note that both the store and output (used by the tracer) have to do with some notion of \text{state}. Thus we define \text{allocLoc}, \text{lookupLoc}, \text{updateLoc}, and \text{write} in terms of just one function:

\[ \text{update} :: (s \rightarrow s) \rightarrow \text{InterpM s} \]

for some suitably chosen \( s \). We can read the state by passing \text{update} the identity function, and change the state by passing it a state transformer. For example:

\[ \text{write } \text{msg} = \text{update } (\lambda \text{sofar } \rightarrow \text{sofar ++ msg}) \]

\[ '\text{bind}' \rightarrow \text{unit} \]

Table 1: Monad operations used by the interpreter

<table>
<thead>
<tr>
<th>Feature</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error handling</td>
<td>\text{err} : String \rightarrow \text{InterpM a}</td>
</tr>
<tr>
<td>Nondeterminism</td>
<td>\text{merge} : [\text{InterpM a}] \rightarrow \text{InterpM a}</td>
</tr>
<tr>
<td>Environment</td>
<td>\text{rdEnv} : \text{InterpM Env}</td>
</tr>
<tr>
<td></td>
<td>\text{inEnv} : \text{Env} \rightarrow \text{InterpM a} \rightarrow \text{InterpM a}</td>
</tr>
<tr>
<td>Store</td>
<td>\text{allocLoc} : \text{InterpM Int}</td>
</tr>
<tr>
<td></td>
<td>\text{lookupLoc} : Int \rightarrow \text{InterpM Value}</td>
</tr>
<tr>
<td></td>
<td>\text{updateLoc} : (\text{Int}, \text{InterpM Value}) \rightarrow \text{InterpM Int}</td>
</tr>
<tr>
<td>String output</td>
<td>\text{write} : String \rightarrow \text{InterpM ()}</td>
</tr>
<tr>
<td>Continuations</td>
<td>\text{callcc} : (\text{a} \rightarrow \text{InterpM b}) \rightarrow \text{InterpM a}</td>
</tr>
</tbody>
</table>

7 Monad Transformers

To get an intuitive feel for monad transformers, consider the merging of a state monad with an arbitrary monad, an example adapted from Jones’s constructor class paper [9]:

\[ \text{type StateT s m a} = s \rightarrow m (s,a) \]

Note that the type variable \( m \) above stands for a type constructor, a fact automatically determined by the Gofer kind inference system. It turns out that if \( m \) is a monad, so is “StateT s m”. \(^5\) “StateT s” is thus a monad transformer.

For example, if we substitute the identity monad:

\[ \text{type Id a} = a \]

for \( m \) in the above monad transformer, we arrive at:

\[ \text{StateT s Id a} = s \rightarrow \text{Id} (s,a) \]
\[ = s \rightarrow (s,a) \]

which is the standard state monad found, for example, in Wadler’s work [21].

The power of monad transformers is twofold. First, they add operations (i.e. introduce new features) to a monad. The \text{State} monad transformer above, for example, adds \text{state} \( s \) to the monad it is applied to, and the resulting monad accepts \text{update} as a legitimate operation on it.

Second, monad transformers compose easily. For example, applying both “StateT s” and “StateT t” to the identity monad, we get:

\[ \text{StateT t (StateT s Id) a} = t \rightarrow (\text{StateT s Id}) (t,a) \]
\[ = t \rightarrow s \rightarrow (s,(t,a)) \]

which is the expected type signature for transforming both \( s \) and \( t \). The observant reader will note, however, an immediate problem: in the resulting monad, which state does \text{update} act upon? In general, this is the problem of \text{lifting} monad operations through transformers, and will be addressed in detail later. But first we define monads and monad transformers more formally, and then describe monad transformers covering the features listed in Section 5.

We can formally define monads as follows:

\(^5\)In fact “StateT s m” is only legal in the current version of Gofer if \text{StateT} is a datatype rather than a type synonym. This does not limit our results, but does introduce superfluous data constructors that slightly complicate the presentation, so we will use \text{type} declarations as if they worked as \text{data} declarations.
The two functions map and join, together with unit provide an equivalent definition of monads, but are easily defined (as default methods) in terms of bind and unit.

To be a monad, bind and unit must satisfy the well-known Monad Laws [21]:

Left unit:

\[(\text{unit } a) \cdot \text{bind } k = k \ a\]

Right unit:

\[m \cdot \text{bind} \ \text{unit} = m\]

Associativity:

\[(m \cdot \text{bind } k) \cdot (k a \cdot \text{bind } h) = (m \cdot \text{bind } k) \cdot h\]

We define a monad transformer as any type constructor \(t\) such that if \(m\) is a monad (based on the above laws), so is \(\langle t \ m \rangle\). We can express this (other than the verification of the laws, which is generally undecidable) using the two-parameter constructor class \(\text{MonadT}\):

\[
\text{class } (\text{Monad } m, \text{ Monad } (t \ m)) \Rightarrow \text{MonadT } t \ m \ where
\]

\[
\text{lift } :: m \ a \rightarrow t \ m \ a
\]

The member function \(\text{lift}\) embeds a computation in monad \(m\) into monad \(\langle t \ m \rangle\). Furthermore, we expect a monad transformer to add features, without changing the nature of an existing computation. We introduce \(\text{Monad Transformer Laws}\) to capture the properties of \(\text{lift}\):

\[
\text{lift } \circ \text{unit}_m = \text{unit}_{t m}
\]

\[
\text{lift } (m \cdot \text{bind}_{t m}) = (\text{lift } \circ \text{unit}_{t m})\cdot (\text{lift } \cdot k)
\]

The above laws say that lifting a null computation results in a null computation, and that lifting a sequence of computations is equivalent to first lifting them individually, and then combining them in the lifted monad.

Specific monad transformers are described in the remainder of this section. Some of these (\(\text{StateT}, \text{ContT}, \text{and ErrorT}\)) appear in an abstract form in Moggi’s note [15]. The \(\text{environment}\) monad is similar to the \(\text{state reader}\) by Plotkern [22]. The \(\text{state and environment}\) monad transformers are related to ideas found in Jones and Duponcheel’s [9] [10] work.

### 7.1 State Monad Transformer

Recall the definition of state monad transformer \(\text{StateT}\):

\[
\text{type StateT } s \ m \ a = s \rightarrow m \ (s, a)
\]

Using instance declarations, we now wish to declare both that \(\text{StateT } s \ m\) is a monad (given \(m\) is a monad), and that \(\text{StateT } s\) is a monad transformer (for each of the state monad transformers defined in subsequent subsections, we will do exactly the same thing).

First, we establish the monad definition for \(\text{StateT } s \ m\), involving methods for \(\text{unit}\) and \(\text{bind}\):

\[
\text{instance Monad } m \Rightarrow \text{MonadT } (\text{StateT } s \ m) \ where
\]

\[
\text{unit } x = \lambda s \rightarrow \text{unit } (s, x)
\]

\[
m \cdot \text{bind}\ k = \lambda s \rightarrow \text{bind } k \ ((s, a) \rightarrow \text{unit } (s, s))
\]

Note that these definitions are not recursive; the constructor class system automatically infers that the \(\text{bind}\) and \(\text{unit}\) appearing on the right are for monad \(m\).

Next, we define “\(\text{StateT } s\)” as a monad transformer:

\[
\text{instance } (\text{Monad } m, \text{ MonadT } (\text{StateT } s \ m)) \Rightarrow \text{MonadT } (\text{StateT } s) \ m \ where
\]

\[
\text{lift } m = \lambda s \rightarrow m \cdot \text{bind } (x \rightarrow \text{unit } (s, x))
\]

Note that \(\text{lift}\) simply runs \(m\) in the new context, while preserving the state.

Finally, as explained earlier, a state monad must support the operation \(\text{update}\). To keep things modular, we define a class of state monads:

\[
\text{class Monad } m \Rightarrow \text{StateMonad } s \ m \ where
\]

\[
\text{update } :: (s \rightarrow t) \rightarrow m \ s
\]

In particular, “\(\text{StateT } s\)” transforms any monad into a state monad, where “\(\text{update}\)” applies \(f\) to the state, and returns the old state:

\[
\text{instance } \text{Monad } m \Rightarrow \text{StateMonad } s (\text{StateT } s \ m) \ where
\]

\[
\text{update } f = \lambda s \rightarrow \text{unit } (f \ s, s)
\]

#### 7.2 Environment Monad Transformer

“\(\text{EnvT } r\)” transforms any monad into an environment monad. The definition of \(\text{bind}\) tells us that two subsequent computation steps run under the same environment \(r\). (Compare this with the state monad, where the second computation is run in the state returned by the first computation.) \(\text{Lift}\) just performs a computation — which cannot depend on the environment — and ignores the environment. \(\text{InEnv}\) ignores the environment carried inside the monad, and performs the computation in a given environment.

\[
\text{type EnvT } r \ m \ a = r \rightarrow m \ a
\]

\[
\text{instance Monad } m \Rightarrow \text{Monad } (\text{EnvT } r \ m) \ where
\]

\[
\text{unit } a = \lambda r \rightarrow \text{unit } a
\]

\[
m \cdot \text{bind}\ k = \lambda r \rightarrow m \ r \cdot \text{bind}\ \lambda a \rightarrow k \ a \ r
\]

\[
\text{instance } (\text{Monad } m, \text{ Monad } (\text{EnvT } r m)) \Rightarrow \text{MonadT } (\text{EnvT } r) \ m \ where
\]

\[
\text{lift } m = \lambda r \rightarrow m
\]

\[
\text{class Monad } m \Rightarrow \text{EnvMonad } r \ m \ where
\]

\[
\text{inEnv } :: \text{Env} ightarrow m \ a ightarrow m \ a
\]

\[
\text{rdEnv } :: m \text{ env}
\]

\[
\text{instance Monad } m \Rightarrow \text{EnvMonad } r (\text{EnvT } r \ m) \ where
\]

\[
inEnv r m = \lambda e \rightarrow m \ r
\]

\[
\text{rdEnv } = \lambda r \rightarrow \text{unit } r
\]

#### 7.3 Error Monad Transformer

\(\text{Monad Error}\) completes a series of computations if all succeed, or aborts as soon as an error occurs. The monad transformer \(\text{ErrorT}\) transforms a monad into an error monad.

\[
\text{data Error } a = \text{ Ok } a | \text{ Error } \text{String}
\]

\[
\text{type ErrorT } m \ a = m (\text{Error } a)
\]
instance Monad m ⇒ Monad (ErrorT m) where
  unit = unit · Ok
  m'bind' k = m'bind' (\a → (Ok x) → k x
                       (Error msg) → unit (Error msg)

case a of
  instance (Monad m, Monad (ErrorT m)) ⇒
  MonadT ErrorT m where
  lift = map unit

class Monad m ⇒ ErrMonad m where
  err :: String → m a

instance Monad m ⇒ ErrMonad (ErrorT m) where
  err = unit · Error

7.4 Continuation Monad Transformer

We define the continuation monad transformer as:

\[
\text{type } \text{ContT ans } m a = (a \rightarrow m \text{ans}) \rightarrow m \text{ans}
\]

instance Monad m ⇒ Monad (ContT ans m) where
  unit x = \k → k x
  m'bind' f = \k → m (\a → f a k)

ContT introduces an additional continuation argument (of type \(a \rightarrow m \text{ans}\)), and by the above definitions of \text{unit} and \text{bind}, all computations in monad “\text{ContT ans m}” are carried out in a continuation passing style.

Lift for “\text{Cont ans m}” turns out to be the same as \text{bind} for \(m\). (It is easy to see this from the type signature.) “\text{Callcc \ f}” invokes the computation in \(f\), passing it a continuation that once applied, throws away the current continuation (denoted as \(\lambda\)) and invokes the captured continuation \(k\).

instance (Monad m, Monad (ContT ans m)) ⇒
  MonadT (ContT ans m) m where
  lift = bind

class Monad m ⇒ ContMonad m where
  callcc :: ((a → m b) → m a) → m a

instance Monad m ⇒ ContMonad (ContT ans m) where
  callcc f = \k → f (\a → \a → k a) k

7.5 The List Monad

Jones and Duponcheel [10] have shown that lists compose with special kinds of monads called commutative monads. It is not clear, however, if lists compose with arbitrary monads. Since many useful monads (e.g. state, error and continuation monads) are not commutative, we cannot define a list monad transformer — one which adds the operation \text{merge} to any monad.

Fortunately, every other monad transformer we have considered in this paper takes arbitrary monads. We thus use lists as the base monad, upon which other transformers can be applied.

instance Monad List where
  unit x = [x]
  m'bind' k = \(x : xs) → k x ++ (xs \ 'bind' \ k)

class Monad m ⇒ ListMonad m where
  merge :: [m a] → m a

instance ListMonad List where
  merge = concat

8 Lifting Operations

We have introduced monad transformers that add useful operations to a given monad, but have not addressed how these operations can be carried through other layers of monad transformers, or equivalently, how a monad transformer lifts existing operations within a monad.

Lifting an operation \(f\) in monad \(m\) through a monad transformer \(t\) results in an operation whose type signature can be derived by substituting all occurrences of \(m\) in the type of \(f\) with \(t\). For example, lifting “\\text{inEnv} :: r → m a → m a” through \(t\) results in an operation with type “\(r → t m a → t m a\)”.

Given the types of operations in monad \(m\):

\[
\tau :: A \\
\tau → \tau \\
(\tau, \tau) → \tau \\
m \tau
\]

\(\prod_\tau\) is the mapping of types across the monad transformer \(t\):

\[
\begin{align*}
[A]_\tau &= A \\
[\tau]_\tau &= \tau \\
[(\tau, \tau)]_\tau &= (\tau_\tau, \tau_\tau) \\
[m \tau]_\tau &= t m \tau
\end{align*}
\]

Moggi [15] studied the problem of lifting under a categorical context. The objective was to identify liftable operations from their type signatures. Unfortunately, many useful operations such as \text{merge}, \text{inEnv} and \text{callcc} failed to meet Moggi’s criteria, and were left unsolved.

We individually consider how to lift these difficult cases. This allows us to make use of their definitions (rather than just the types), and find ways to lift them through all monad transformers studied so far.

This is exactly where monad transformers provide us with an opportunity to study how various programming language features interact. The easy-to-lift cases correspond to features that are independent in nature, and the more involved cases require a deeper analysis of monad structures in order to clarify the semantics.

An unfortunate consequence of our approach is that as we consider more monad transformers, the number of possible liftings grows quadratically. It seems, however, that there are not too many different kinds of monad transformers (although there may be many instances of the same monad transformer such as \text{StateT}). What we introduced so far are able to model almost all commonly known features of sequential languages. Even so, not all of them are strictly necessary. The environment, for example, can be simulated using a state monad:

instance (Monad m, StateMonad r m) ⇒
  EnvMonad r m where
  inEnv r m = update (\(x → r\)) \'(bind' \(\cdot\) →
                       m \'bind' \(\cdot\) → update (\(x → \pi\)) \'(bind' \(\cdot\) →
                       unit
  rdEnv = update id


Also, as is well known, error reporting can be implemented using \texttt{callcc}.

### 8.1 Correctness Criteria

The basic requirement of lifting is that any program which does not use the added features should behave in the same way after a monad transformer is applied. The monad transformer laws introduced in Section \ref{sec:monad-transformers} are meant to guarantee such property for lifting a single computation. Most monad operations, however, have more general types. To deal with operations on arbitrary types, we extend Moggi's corresponding categorical approach, and define $\mathcal{L}_\tau$ as the \textit{natural lifting} of operations of type $\tau$ along the monad transformer $k$:

\[
\begin{align*}
\mathcal{L}_\tau &::= \tau \Rightarrow \left[\tau\right]_k \\
\mathcal{L}_\tau i &::= \text{id} \\
\mathcal{L}_{\tau_1 \times \tau_2} &::= \left\{ f \to f' \text{ such that } f(a,b) = \mathcal{L}_{\tau_1} a \cdot \mathcal{L}_{\tau_2} b \right\} \\
\mathcal{L}_{\tau_1} (\tau_2) &::= \left(\mathcal{L}_{\tau_1} \cdot \mathcal{L}_{\tau_2}\right) \\
\mathcal{L}_m \tau &::= \text{lift} \cdot \left(\text{map } \mathcal{L}_\tau\right)
\end{align*}
\]

Constant types (such as \texttt{Integer}) and type variables do not depend on any particular monad. (See cases 1 and 2.) On the other hand, we expect a lifted function, when applied to a value lifted from the domain of the original function, to return the lifting of the result of applying the original function to the unlifted value. This relationship is precisely captured by equation 3, which corresponds to the following commuting diagram:

\[
\begin{array}{c}
\left[\tau_1\right]_k \xrightarrow{f'} \left[\tau_2\right]_k \\
\mathcal{L}_{\tau_1} \xrightarrow{f} \mathcal{L}_{\tau_2}
\end{array}
\]

The lifting of tuples is straightforward. Finally, the \texttt{lift} operator come with the monad transformer lifts computations expressed in monad types. Note that $\mathcal{L}_\tau$ is mapped to the result of the computation, which may involve other computations.

Note that the above does not provide a Gofer definition for an overloaded lifting function $\mathcal{L}$. The “such that” clause in the third equation specifies a constraint, rather than a definition of $f'$. In practice, we first find out by hand how to lift an operation through a certain (or a class of) monad transformer, and then use the above equations to verify that such a lifting is indeed natural. Generally we require operations to be lifted naturally — although as will be seen, certain unnatural liftings change the semantics in interesting ways.

### 8.2 Easy Cases

\texttt{Err} and \texttt{update} are handled by \texttt{lift}, whereas \texttt{merge} benefits from \texttt{List} being the base monad.

\[
\text{instance (ErrMonad m, MonadT t m)} \Rightarrow \text{ErrMonad (t m) where}
\]
\[
\text{err} = \text{lift} \cdot \text{err}
\]

\[
\text{instance (MonadT t m, MonadT (t m)} \Rightarrow \text{MonadT (t m) where}
\]
\[
\text{update} = \text{lift} \cdot \text{update}
\]

\[
\text{instance MonadT t List} \Rightarrow \text{ListMonad (t List) where}
\]
\[
\text{merge} = \text{join} \cdot \text{lift}
\]

### 8.3 Lifting \texttt{callcc}

The following lifting of \texttt{callcc} through \texttt{EnvT} discards the \textit{current} environment $\tau'*$ upon invoking the captured continuation $k$. The execution will continue in the environment $\tau$ captured when \texttt{callcc} was first invoked.

\[
\text{instance (MonadT (ErrT r m, ContMonad m) \Rightarrow \text{ContMonad (EnvT r m) where}}
\]
\[
\text{- } \text{callcc} :: \left((a \to r \to m \ b) \to r \to m \ a\right) \to r \to m \ a
\]
\[
\text{callcc } f = \left(\lambda r \to \text{callcc} \left(k \cdot f \ (\lambda a \to \lambda r' \to k \ a) \ r\right)\right)
\]

The Appendix shows that if we flip the order of monad transformers and apply \texttt{ContT} to “\texttt{EnvT env m}” — in which case no lifting of \texttt{callcc} will be necessary — the current environment will be passed to the continuation. (We will see how to fix this by carefully recovering the environment when we lift \texttt{inEnv} in a moment.)

In general we can swap the order of some monad transformers (such as between \texttt{StateT} and \texttt{EnvT}), but doing so to others (such as \texttt{ContT}) may affect semantics. This is consistent with Filinski’s observations [6], and, in practice, provides us an opportunity to fine tune the resulting semantics.

In lifting \texttt{callcc} through “\texttt{StateT s}”, we have a choice of passing either the current state $s_1$ or the captured state $s_0$. The former is the usual semantics for \texttt{callcc}, and the latter is useful in Tolmach and Appel’s approach to debugging [20].

\[
\text{instance (MonadT (StateT s m, ContMonad m) \Rightarrow \text{ContMonad (StateT s m) where}}
\]
\[
\text{- } \text{callcc} :: \left((a \to s \to m \ b, s) \to s \to m \ a\right)
\]
\[
\text{callcc } f = \left(\lambda s \to \text{callcc} \left(k \cdot f \ (\lambda a \to \lambda s_1 \to k \ (s_1, a)) \ s_1\right)\right)
\]

The above shows the usual \texttt{callcc} semantics, and can be changed to the “debugging” version by instead passing $(s_0, a)$ to $k$.

The lifting of \texttt{inEnv} through \texttt{ErrorT} can be found in the Appendix.

### 8.4 Lifting \texttt{inEnv}

We only consider lifting \texttt{inEnv} through \texttt{ContT} here; the Appendix shows how to lift \texttt{inEnv} through other monad transformers.

\[
\text{instance (MonadT (ContT ans) m, EnvMonad r m) \Rightarrow \text{EnvMonad r (ContT ans) where}}
\]
\[
\text{inEnv } r \ c = \left(\lambda k \to \text{rdEnv } \text{"bind"} \ (\lambda c \to \text{inEnv } r \ (c \ (\text{inEnv } o \cdot k))\right)
\]
\[
\text{rdEnv } = \text{lift} \cdot \text{rdEnv}
\]

We restore the environment before invoking the continuation, sort of like popping arguments off the stack. On the other hand, an interesting (but not natural) way to lift \texttt{inEnv} is:
instance (MonadT (ContT ans) m, EnvMonad r m) ⇒ EnvMonad r (ContT ans m)

where

\[
\begin{align*}
\text{inEnv} r c & \equiv \lambda k \rightarrow \text{inEnv} r (c \ k) \\
\text{rdEnv} & \equiv \text{lift rdEnv}
\end{align*}
\]

Here the environment is not restored when \( c \) invokes \( k \), and thus reflects the history of dynamic execution.

9 Conclusions

We have shown how a modular monadic interpreter can be designed using two key ideas: extensible union types and monad transformers, and implemented using constructor classes. A key technical problem that we had to overcome was the lifting of operations through monads. Our approach also helps to clarify the interactions between various programming language features.

This paper realized Moggi’s idea of a modular presentation of denotational semantics for complicated languages, and is much cleaner than the traditional approach [19]. On the practical side, our results provide new insights into designing and implementing programming languages, in particular, extensible languages, which allow the programmer to specify new features on top of existing ones.

Acknowledgements

We thank Dan Rabin, Zhong Shao, Rajiv Mirani and anonymous referees for helpful suggestions.

References


A  The Ordering of \textit{ContT} and \textit{EnvT}

It is interesting to compare the following two \textit{callcc} functions on monad $M$ and $N$, both composed from \textit{“ContT ans”} and \textit{“EnvT m”}, but in different order.

Case 1:

\[
\text{type } M a = \text{ContT ans (EnvT r m) a} = (a \rightarrow r \rightarrow m \text{ ans}) \rightarrow r \rightarrow m \text{ ans}
\]

\[
\text{callcc } f = \lambda k \rightarrow f (\lambda a \rightarrow \lambda r \rightarrow k a) k
\]

(eta convert \text{\texttt{\_}} and \text{\texttt{\_}})

\[
= \lambda k \rightarrow \lambda r \rightarrow f (\lambda a \rightarrow \lambda r' \rightarrow k a r') k r
\]

Case 2:

\[
\text{type } M a = \text{EnvT r (ContT ans m) a} = r \rightarrow (a \rightarrow m \text{ ans}) \rightarrow m \text{ ans}
\]

\[
\text{callcc } f = \lambda r \rightarrow \text{callcc} (\lambda k \rightarrow f (\lambda a \rightarrow \lambda r' \rightarrow k a) r)
\]

\[
= \lambda r \rightarrow \lambda k \rightarrow (\lambda k' \rightarrow f (\lambda a \rightarrow \lambda r' \rightarrow k' a) r')
\]

\[
(\lambda a \rightarrow \lambda r \rightarrow k a) k
\]

\[
= \lambda r \rightarrow \lambda k \rightarrow f (\lambda a \rightarrow \lambda r' \rightarrow \lambda r \rightarrow k a) r k
\]

From the expansion of type $M$ in case 1, we can see that both result and environment are passed to the continuation. When \textit{callcc} invokes a continuation, it passes the current, rather than the captured continuation. The \textit{callcc} function in case 2 works in the opposite way.

B  Lifting \textit{Callcc through ErrorT}

\[
\text{instance (MonadT ErrorT m, ContMonad m) } \Rightarrow \text{ContMonad (ErrorT m) where}
\]

\[
\text{callcc } : (a \rightarrow m \text{ (Error a)}) \rightarrow m \text{ (Error a)}
\]

\[
\text{callcc } f = \text{callcc} (\lambda k \rightarrow f (\lambda a \rightarrow k \text{ (Ok a)}))
\]

C  Lifting \textit{InEnv through EnvT, StateT and ErrorT}

\[
\text{instance (MonadT (EnvT s) m, EnvMonad r m) } \Rightarrow \text{EnvMonad r (EnvT s m) where}
\]

\[
\text{inEnv } r m = \lambda s \rightarrow \text{inEnv } r (m s)
\]

\[
\text{rdEnv } = \text{lift rdEnv}
\]

\[
\text{instance (MonadT (StateT s) m, EnvMonad r m) } \Rightarrow \text{EnvMonad r (StateT s m) where}
\]

\[
\text{inEnv } r m = \lambda s \rightarrow \text{inEnv } r (m s)
\]

\[
\text{rdEnv } = \text{lift rdEnv}
\]

A function of type “\texttt{m a \rightarrow m a}” maps “\texttt{m (Error a)}” to “\texttt{m (Error a)}”, thus \texttt{inEnv} stays the same after being lifted through \texttt{ErrorT}.