COMP4075: Lecture 8

Introduction to Monads

Henrik Nilsson

University of Nottingham, UK

A Blessing and a Curse

- The ***BIG*** advantage of ***pure*** functional programming is

  "everything is explicit;"

  i.e., flow of data manifest, no side effects. Makes it a lot easier to understand large programs.
- The ***BIG*** problem with ***pure*** functional programming is

  "everything is explicit."

  Can add a lot of clutter, make it hard to maintain code

Conundrum

"Shall I be pure or impure?" (Wadler, 1992)

- Absence of effects
  - facilitates understanding and reasoning
  - makes lazy evaluation viable
  - allows choice of reduction order, e.g. parallel
  - enhances modularity and reuse.
- Effects (state, exceptions, . . .) can
  - help making code concise
  - facilitate maintenance
  - improve the efficiency.

Answer to Conundrum: Monads (1)

- Monads bridges the gap: allow effectful programming in a pure setting.
- Key idea: ***Computational types***: an object of type $M\,A$ denotes a ***computation*** of an object of type A.
- ***Thus we shall be both pure and impure, whatever takes our fancy!***
- Monads originated in Category Theory.
- Adapted by
  - Moggi for structuring denotational semantics
  - Wadler for structuring functional programs

Answer to Conundrum: Monads (2)

Monads

- promote ***disciplined*** use of effects since the type reflects which effects can occur;
- allow great ***flexibility*** in tailoring the effect structure to precise needs;
- support ***changes*** to the effect structure with minimal impact on the overall program structure;
- allow integration into a pure setting of ***real*** effects such as
  - I/O
  - mutable state.

This Lecture

Pragmatic introduction to monads:

- Effectful computations
- Identifying a common pattern
- Monads as a ***design pattern***

Example 1: A Simple Evaluator

$$
\begin{aligned}
\mathbf{data}\ \mathit{Exp} = {} & \mathit{Lit\ Integer} \\
\mid {} & \mathit{Add\ Exp\ Exp} \\
\mid {} & \mathit{Sub\ Exp\ Exp} \\
\mid {} & \mathit{Mul\ Exp\ Exp} \\
\mid {} & \mathit{Div\ Exp\ Exp}
\end{aligned}
$$

$$
\begin{aligned}
&\mathit{eval} :: \mathit{Exp} \to \mathit{Integer} \\
&\mathit{eval}\ (\mathit{Lit}\ n) \qquad\ \ = n \\
&\mathit{eval}\ (\mathit{Add}\ e1\ e2) = \mathit{eval}\ e1 + \mathit{eval}\ e2 \\
&\mathit{eval}\ (\mathit{Sub}\ e1\ e2)\ = \mathit{eval}\ e1 - \mathit{eval}\ e2 \\
&\mathit{eval}\ (\mathit{Mul}\ e1\ e2) = \mathit{eval}\ e1 * \mathit{eval}\ e2 \\
&\mathit{eval}\ (\mathit{Div}\ e1\ e2)\ = \mathit{eval}\ e1\ \text{`}\mathit{div}\text{`}\ \mathit{eval}\ e2
\end{aligned}
$$

Making the Evaluator Safe (1)

$$
\begin{aligned}
&\mathbf{data}\ \mathit{Maybe}\ a = \mathit{Nothing} \mid \mathit{Just}\ a \\
&\mathit{safeEval} :: \mathit{Exp} \to \mathit{Maybe\ Integer} \\
&\mathit{safeEval}\ (\mathit{Lit}\ n) \qquad\ \ = \mathit{Just}\ n \\
&\mathit{safeEval}\ (\mathit{Add}\ e1\ e2) = \\
&\quad \mathbf{case}\ \mathit{safeEval}\ e1\ \mathbf{of} \\
&\quad\quad \mathit{Nothing} \to \mathit{Nothing} \\
&\quad\quad \mathit{Just}\ n1\ \ \to \mathbf{case}\ \mathit{safeEval}\ e2\ \mathbf{of} \\
&\qquad\qquad\qquad\quad \mathit{Nothing} \to \mathit{Nothing} \\
&\qquad\qquad\qquad\quad \mathit{Just}\ n2\ \ \to \mathit{Just}\ (n1 + n2)
\end{aligned}
$$

Making the Evaluator Safe (2)

$$
\begin{array}{l}
\mathit{safeEval}\ (\mathit{Sub}\ e1\ \ e2) = \\
\quad \textbf{case}\ \mathit{safeEval}\ e1\ \textbf{of} \\
\quad\quad \mathit{Nothing} \rightarrow \mathit{Nothing} \\
\quad\quad \mathit{Just}\ n1\ \ \rightarrow \textbf{case}\ \mathit{safeEval}\ e2\ \textbf{of} \\
\quad\quad\quad\quad\quad\quad \mathit{Nothing} \rightarrow \mathit{Nothing} \\
\quad\quad\quad\quad\quad\quad \mathit{Just}\ n2\ \ \rightarrow \mathit{Just}\ (n1 - n2)
\end{array}
$$

Making the Evaluator Safe (3)

$$
\begin{array}{l}
\mathit{safeEval}\ (\mathit{Mul}\ e1\ \ e2) = \\
\quad \textbf{case}\ \mathit{safeEval}\ e1\ \textbf{of} \\
\quad\quad \mathit{Nothing} \rightarrow \mathit{Nothing} \\
\quad\quad \mathit{Just}\ n1\ \ \rightarrow \textbf{case}\ \mathit{safeEval}\ e2\ \textbf{of} \\
\quad\quad\quad\quad\quad\quad \mathit{Nothing} \rightarrow \mathit{Nothing} \\
\quad\quad\quad\quad\quad\quad \mathit{Just}\ n2\ \ \rightarrow \mathit{Just}\ (n1 * n2)
\end{array}
$$

Making the Evaluator Safe (4)

$$
\begin{array}{l}
\mathit{safeEval}\ (\mathit{Div}\ e1\ \ e2) = \\
\quad \textbf{case}\ \mathit{safeEval}\ e1\ \textbf{of} \\
\quad\quad \mathit{Nothing} \rightarrow \mathit{Nothing} \\
\quad\quad \mathit{Just}\ n1\ \ \rightarrow \textbf{case}\ \mathit{safeEval}\ e2\ \textbf{of} \\
\quad\quad\quad\quad\quad\quad \mathit{Nothing} \rightarrow \mathit{Nothing} \\
\quad\quad\quad\quad\quad\quad \mathit{Just}\ n2\ \ \rightarrow \\
\quad\quad\quad\quad\quad\quad\quad \textbf{if}\ n2 \equiv 0 \\
\quad\quad\quad\quad\quad\quad\quad \textbf{then}\ \mathit{Nothing} \\
\quad\quad\quad\quad\quad\quad\quad \textbf{else}\ \mathit{Just}\ (n1\ \text{`}\mathit{div}\text{`}\ n2)
\end{array}
$$

Any Common Pattern?

Clearly a lot of code duplication!
Can we factor out a common pattern?

We note:

- ***Sequencing*** of evaluations (or ***computations***).
- If one evaluation fails, fail overall.
- Otherwise, make result available to following evaluations.

Sequencing Evaluations

$$\begin{aligned} &\mathit{evalSeq} :: \mathit{Maybe\ Integer} \\ &\qquad \to (\mathit{Integer} \to \mathit{Maybe\ Integer}) \\ &\qquad \to \mathit{Maybe\ Integer} \\ &\mathit{evalSeq}\ \mathit{ma}\ f = \mathbf{case}\ \mathit{ma}\ \mathbf{of} \\ &\qquad\qquad \mathit{Nothing} \to \mathit{Nothing} \\ &\qquad\qquad \mathit{Just}\ a \to f\ a \end{aligned}$$

Exercise 1: Refactoring *safeEval*

Rewrite *safeEval*, case *Add*, using *evalSeq*:

```
safeEval (Add e1 e2) =
    case safeEval e1 of
        Nothing -> Nothing
        Just n1 ->
            case safeEval e2 of
                Nothing -> Nothing
                Just n2 -> Just (n1 + n2)
evalSeq ma f =
    case ma of
        Nothing -> Nothing
        Just a  -> f a
```


Exercise 1: Solution

$$\begin{aligned} &\mathit{safeEval} :: \mathit{Exp} \to \mathit{Maybe\ Integer} \\ &\mathit{safeEval}\ (\mathit{Add}\ e1\ e2) = \\ &\quad \mathit{evalSeq}\ (\mathit{safeEval}\ e1) \\ &\qquad (\lambda n1 \to \mathit{evalSeq}\ (\mathit{safeEval}\ e2) \\ &\qquad\qquad (\lambda n2 \to \mathit{Just}\ (n1 + n2))) \end{aligned}$$

or

$$\begin{aligned} &\mathit{safeEval} :: \mathit{Exp} \to \mathit{Maybe\ Integer} \\ &\mathit{safeEval}\ (\mathit{Add}\ e1\ e2) = \\ &\quad \mathit{safeEval}\ e1\ `\mathit{evalSeq}`\ \lambda n1 \to \\ &\quad \mathit{safeEval}\ e2\ `\mathit{evalSeq}`\ \lambda n2 \to \\ &\quad \mathit{Just}\ (n1 + n2) \end{aligned}$$

Refactored Safe Evaluator (1)

$$\begin{aligned} &\mathit{safeEval} :: \mathit{Exp} \to \mathit{Maybe\ Integer} \\ &\mathit{safeEval}\ (\mathit{Lit}\ n) = \mathit{Just}\ n \\ &\mathit{safeEval}\ (\mathit{Add}\ e1\ e2) = \\ &\quad \mathit{safeEval}\ e1\ `\mathit{evalSeq}`\ \lambda n1 \to \\ &\quad \mathit{safeEval}\ e2\ `\mathit{evalSeq}`\ \lambda n2 \to \\ &\quad \mathit{Just}\ (n1 + n2) \\ &\mathit{safeEval}\ (\mathit{Sub}\ e1\ e2) = \\ &\quad \mathit{safeEval}\ e1\ `\mathit{evalSeq}`\ \lambda n1 \to \\ &\quad \mathit{safeEval}\ e2\ `\mathit{evalSeq}`\ \lambda n2 \to \\ &\quad \mathit{Just}\ (n1 - n2) \end{aligned}$$

Refactored Safe Evaluator (2)

$$
\begin{aligned}
&safeEval\ (Mul\ e1\ \ e2) = \\
&\quad safeEval\ e1\ \text{`}evalSeq\text{`}\ \lambda n1 \rightarrow \\
&\quad safeEval\ e2\ \text{`}evalSeq\text{`}\ \lambda n2 \rightarrow \\
&\quad Just\ (n1 * n2) \\
&safeEval\ (Div\ e1\ \ e2) = \\
&\quad safeEval\ e1\ \text{`}evalSeq\text{`}\ \lambda n1 \rightarrow \\
&\quad safeEval\ e2\ \text{`}evalSeq\text{`}\ \lambda n2 \rightarrow \\
&\quad \textbf{if}\ n2 \equiv 0 \\
&\quad \textbf{then}\ Nothing \\
&\quad \textbf{else}\ Just\ (n1\ \text{`}div\text{`}\ n2)
\end{aligned}
$$

`Maybe` Viewed as a Computation (1)

- Consider a value of type `Maybe a` as denoting a ***computation*** of a value of type `a` that ***may fail***.
- When sequencing possibly failing computations, a natural choice is to fail overall once a subcomputation fails.
- I.e. ***failure is an effect***, implicitly affecting subsequent computations.
- Let's generalize and adopt names reflecting our intentions.

`Maybe` Viewed as a Computation (2)

Successful computation of a value:

$$
\begin{aligned}
&mbReturn :: a \rightarrow Maybe\ a \\
&mbReturn = Just
\end{aligned}
$$

Sequencing of possibly failing computations:

$$
\begin{aligned}
&mbSeq :: Maybe\ a \rightarrow (a \rightarrow Maybe\ b) \rightarrow Maybe\ b \\
&mbSeq\ ma\ f = \textbf{case}\ ma\ \textbf{of} \\
&\qquad\qquad Nothing \rightarrow Nothing \\
&\qquad\qquad Just\ a \quad \rightarrow f\ a
\end{aligned}
$$

`Maybe` Viewed as a Computation (3)

Failing computation:

$$
\begin{aligned}
&mbFail :: Maybe\ a \\
&mbFail = Nothing
\end{aligned}
$$

The Safe Evaluator Revisited

$$
\begin{aligned}
&\mathit{safeEval} :: \mathit{Exp} \to \mathit{Maybe\ Integer} \\
&\mathit{safeEval}\ (\mathit{Lit}\ n) = \mathit{mbReturn}\ n \\
&\mathit{safeEval}\ (\mathit{Add}\ e1\ e2) = \\
&\quad \mathit{safeEval}\ e1\ `\mathit{mbSeq}`\ \lambda n1 \to \\
&\quad \mathit{safeEval}\ e2\ `\mathit{mbSeq}`\ \lambda n2 \to \\
&\quad \mathit{mbReturn}\ (n1 + n2) \\
&\ldots \\
&\mathit{safeEval}\ (\mathit{Div}\ e1\ e2) = \\
&\quad \mathit{safeEval}\ e1\ `\mathit{mbSeq}`\ \lambda n1 \to \\
&\quad \mathit{safeEval}\ e2\ `\mathit{mbSeq}`\ \lambda n2 \to \\
&\quad \mathbf{if}\ n2 \equiv 0\ \mathbf{then}\ \mathit{mbFail}\ \mathbf{else}\ \mathit{mbReturn}\ (n1\ `\mathit{div}`\ n2)))
\end{aligned}
$$

Example 2: Numbering Trees

$$
\begin{aligned}
&\mathbf{data}\ \mathit{Tree}\ a = \mathit{Leaf}\ a \mid \mathit{Node}\ (\mathit{Tree}\ a)\ (\mathit{Tree}\ a) \\
&\mathit{numberTree} :: \mathit{Tree}\ a \to \mathit{Tree\ Int} \\
&\mathit{numberTree}\ t = \mathit{fst}\ (\mathit{ntAux}\ t\ 0) \\
&\quad \mathbf{where} \\
&\qquad \mathit{ntAux} :: \mathit{Tree}\ a \to \mathit{Int} \to (\mathit{Tree\ Int}, \mathit{Int}) \\
&\qquad \mathit{ntAux}\ (\mathit{Leaf}\ \_)\ n \qquad = (\mathit{Leaf}\ n, n + 1) \\
&\qquad \mathit{ntAux}\ (\mathit{Node}\ t1\ t2)\ n = \\
&\qquad\quad \mathbf{let}\ (t1', n') = \mathit{ntAux}\ t1\ n \\
&\qquad\quad \mathbf{in\ let}\ (t2', n'') = \mathit{ntAux}\ t2\ n' \\
&\qquad\qquad \mathbf{in}\ (\mathit{Node}\ t1'\ t2', n'')
\end{aligned}
$$

Observations

- Repetitive pattern: threading a counter through a ***sequence*** of tree numbering ***computations***.
- It is very easy to pass on the wrong version of the counter!

Can we do better?

Stateful Computations (1)

- A ***stateful computation*** consumes a state and returns a result along with a possibly updated state.
- The following type synonym captures this idea:

  $$\mathbf{type}\ S\ a = \mathit{Int} \to (a, \mathit{Int})$$

  (Only Int state for the sake of simplicity.)
- A value (function) of type $S\ a$ can now be viewed as denoting a stateful computation computing a value of type a.

Stateful Computations (2)

- When sequencing stateful computations, the resulting state should be passed on to the next computation.
- I.e. ***state updating is an effect***, implicitly affecting subsequent computations. (As we would expect.)

Stateful Computations (3)

Computation of a value without changing the state (For ref.: $S\ a \quad = \quad Int \to (a, Int)$):

$$sReturn :: a \to S\ a$$
$$sReturn\ a = \lambda n \to (a, n)$$

Sequencing of stateful computations:

$$sSeq :: S\ a \to (a \to S\ b) \to S\ b$$
$$sSeq\ sa\ f = \lambda n \to$$
$$\quad \mathbf{let}\ (a, n') = sa\ n$$
$$\quad \mathbf{in}\ f\ a\ n'$$

Stateful Computations (4)

Reading and incrementing the state (For ref.: $S\ a \quad = \quad Int \to (a, Int)$):

$$sInc :: S\ Int$$
$$sInc = \lambda n \to (n, n + 1)$$

Numbering trees revisited

$$\mathbf{data}\ Tree\ a = Leaf\ a \mid Node\ (Tree\ a)\ (Tree\ a)$$
$$numberTree :: Tree\ a \to Tree\ Int$$
$$numberTree\ t = fst\ (ntAux\ t\ 0)$$
$$\quad \mathbf{where}$$
$$\quad\quad ntAux :: Tree\ a \to S\ (Tree\ Int)$$
$$\quad\quad ntAux\ (Leaf\ \_) =$$
$$\quad\quad\quad sInc\ `sSeq`\ \lambda n \to sReturn\ (Leaf\ n)$$
$$\quad\quad ntAux\ (Node\ t1\ t2) =$$
$$\quad\quad\quad ntAux\ t1\ `sSeq`\ \lambda t1' \to$$
$$\quad\quad\quad ntAux\ t2\ `sSeq`\ \lambda t2' \to$$
$$\quad\quad\quad sReturn\ (Node\ t1'\ t2')$$

Observations

- The "plumbing" has been captured by the abstractions.
- In particular:
  - counter no longer manipulated directly
  - no longer any risk of "passing on" the wrong version of the counter!

Comparison of the examples

- Both examples characterized by sequencing of effectful computations.
- Both examples could be neatly structured by introducing:
  - A type denoting computations
  - A function constructing an effect-free computation of a value
  - A function constructing a computation by sequencing computations
- In fact, both examples are instances of the general notion of a ***MONAD***.

Monads in Functional Programming

A monad is represented by:

- A type constructor

  $$M :: * \to *$$

  $M\ T$ represents computations of value of type T.
- A polymorphic function

  $$return :: a \to M\ a$$

  for lifting a value to a computation.
- A polymorphic function

  $$(\gg\!=) :: M\ a \to (a \to M\ b) \to M\ b$$

  for sequencing computations.

Exercise 2: $join$ and $fmap$

Equivalently, the notion of a monad can be captured through the following functions:

$$\begin{aligned} return &:: a \to M\ a \\ join &:: (M\ (M\ a)) \to M\ a \\ fmap &:: (a \to b) \to M\ a \to M\ b \end{aligned}$$

$join$ "flattens" a computation, $fmap$ "lifts" a function to map computations to computations.

Define $join$ and $fmap$ in terms of $(\gg\!=)$ (and $return$), and $(\gg\!=)$ in terms of $join$ and $fmap$.

$$(\gg\!=) :: M\ a \to (a \to M\ b) \to M\ b$$

Exercise 2: Solution

$$join :: M\ (M\ a) \to M\ a$$
$$join\ mm = mm \gg\!= id$$
$$fmap :: (a \to b) \to M\ a \to M\ b$$
$$fmap\ f\ m = m \gg\!= return \circ f$$
$$(\gg\!=) :: M\ a \to (a \to M\ b) \to M\ b$$
$$m \gg\!= f = join\ (fmap\ f\ m)$$

Monad laws

Additionally, the following ***laws*** must be satisfied:

$$\begin{aligned} return\ x \gg\!= f &= f\ x \\ m \gg\!= return &= m \\ (m \gg\!= f) \gg\!= g &= m \gg\!= (\lambda x \to f\ x \gg\!= g) \end{aligned}$$

I.e., $return$ is the right and left identity for $(\gg\!=)$, and $(\gg\!=)$ is associative.

Exercise 3: The Identity Monad

The ***Identity Monad*** can be understood as representing ***effect-free*** computations:

$$\mathbf{type}\ I\ a = a$$

1. Provide suitable definitions of $return$ and $(\gg\!=)$.
2. Verify that the monad laws hold for your definitions.

Exercise 3: Solution

$$return :: a \to I\ a$$
$$return = id$$
$$(\gg\!=) :: I\ a \to (a \to I\ b) \to I\ b$$
$$m \gg\!= f = f\ m$$

(Or: $(\gg\!=) = flip\ (\$)$)

Simple calculations verify the laws, e.g.:

$$\begin{aligned} return\ x \gg\!= f &= id\ x \gg\!= f \\ &= x \gg\!= f \\ &= f\ x \end{aligned}$$

Reading

- Philip Wadler. The Essence of Functional Programming. *Proceedings of the 19th ACM Symposium on Principles of Programming Languages (POPL'92)*, 1992.
- Nick Benton, John Hughes, Eugenio Moggi. Monads and Effects. In *International Summer School on Applied Semantics 2000*, Caminha, Portugal, 2000.
- *All About Monads.* `http://www.haskell.org/all_about_monads`