

Functional Programming

IO Monad

Jevgeni Kabanov

Department of Computer Science
University of Tartu

Overview

- 1 Functional IO
- 2 Monadic IO
- 3 Higher Order IO

Outline

- 1 Functional IO
- 2 Monadic IO
- 3 Higher Order IO

Lazy Pure Haskell

Haskell is a *pure* and *lazy* language

Pure

- Purity means that every function result is uniquely determined by its parameters
- Compiler is free to inline any function as it pleases
- Runtime can cache any function call

Lazy

- Laziness means that values are computed on need
- If they are not needed, they are not computed
- Runtime is free to precompute them or reorder computations

Lazy Pure Haskell

Haskell is a *pure* and *lazy* language

Pure

- Purity means that every function result is uniquely determined by its parameters
- Compiler is free to inline any function as it pleases
- Runtime can cache any function call

Lazy

- Laziness means that values are computed on need
- If they are not needed, they are not computed
- Runtime is free to precompute them or reorder computations

Lazy Pure Haskell

Haskell is a *pure* and *lazy* language

Pure

- Purity means that every function result is uniquely determined by its parameters
- Compiler is free to inline any function as it pleases
- Runtime can cache any function call

Lazy

- Laziness means that values are computed on need
- If they are not needed, they are not computed
- Runtime is free to precompute them or reorder computations

Lazy Pure Haskell

Haskell is a *pure* and *lazy* language

Pure

- Purity means that every function result is uniquely determined by its parameters
- Compiler is free to inline any function as it pleases
- Runtime can cache any function call

Lazy

- Laziness means that values are computed on need
- If they are not needed, they are not computed
- Runtime is free to precompute them or reorder computations

Lazy Pure Haskell

Haskell is a *pure* and *lazy* language

Pure

- Purity means that every function result is uniquely determined by its parameters
- Compiler is free to inline any function as it pleases
- Runtime can cache any function call

Lazy

- Laziness means that values are computed on need
- If they are not needed, they are not computed
- Runtime is free to precompute them or reorder computations

Lazy Pure Haskell

Haskell is a *pure* and *lazy* language

Pure

- Purity means that every function result is uniquely determined by its parameters
- Compiler is free to inline any function as it pleases
- Runtime can cache any function call

Lazy

- Laziness means that values are computed on need
- If they are not needed, they are not computed
- Runtime is free to precompute them or reorder computations

Lazy Pure Haskell

Haskell is a *pure* and *lazy* language

Pure

- Purity means that every function result is uniquely determined by its parameters
- Compiler is free to inline any function as it pleases
- Runtime can cache any function call

Lazy

- Laziness means that values are computed on need
- If they are not needed, they are not computed
- Runtime is free to precompute them or reorder computations

Pure IO

The following function should read a character from standard input

getChar :: Char

Let's see an example:

get2chars = [*getChar*, *getChar*]

Questions

- How many characters will be read?
- In what order will they be read?

Pure IO

The following function should read a character from standard input

getChar :: *Char*

Let's see an example:

get2chars = [*getChar*, *getChar*]

Questions

- How many characters will be read?
- In what order will they be read?

Pure IO

The following function should read a character from standard input

getChar :: *Char*

Let's see an example:

get2chars = [*getChar*, *getChar*]

Questions

- How many characters will be read?
- In what order will they be read?

Pure IO

The following function should read a character from standard input

getChar :: *Char*

Let's see an example:

get2chars = [*getChar*, *getChar*]

Questions

- How many characters will be read?
- In what order will they be read?

Pure IO

Problem 1

Let's try to solve these problems:

$$\text{getChar} :: \text{Int} \rightarrow \text{Char}$$
$$\text{get2chars} = [\text{getChar } 1, \text{getChar } 2]$$

Now the values are distinct, yet the order is still undefined

Problem 2

Also we now need to add the same *Int* argument to *get2chars* for the same reason:

$$\text{get4chars} :: \text{Int} \rightarrow [\text{Char}]$$
$$\text{get4chars} = \text{get2chars } 1 \text{ ++ get2chars } 2$$

Pure IO

Problem 1

Let's try to solve these problems:

$$\text{getChar} :: \text{Int} \rightarrow \text{Char}$$
$$\text{get2chars} = [\text{getChar } 1, \text{getChar } 2]$$

Now the values are distinct, yet the order is still undefined

Problem 2

Also we now need to add the same *Int* argument to *get2chars* for the same reason:

$$\text{get4chars} :: \text{Int} \rightarrow [\text{Char}]$$
$$\text{get4chars} = \text{get2chars } 1 \text{ ++ get2chars } 2$$

Pure IO

Problem 1

Let's try to solve these problems:

$$\text{getChar} :: \text{Int} \rightarrow \text{Char}$$
$$\text{get2chars} = [\text{getChar } 1, \text{getChar } 2]$$

Now the values are distinct, yet the order is still undefined

Problem 2

Also we now need to add the same *Int* argument to *get2chars* for the same reason:

$$\text{get4chars} :: \text{Int} \rightarrow [\text{Char}]$$
$$\text{get4chars} = \text{get2chars } 1 \text{ ++ get2chars } 2$$

Lazy IO

Problem 3

getChar :: *Int* → *Char*

get2chars = [*getChar* 1, *getChar* 2]

- What will be called first?
- What we really need is some kind of dependency!
- The only kind of runtime dependency Haskell provides is value dependency:

getchar :: *Int* → (*Char*, *Int*)

get2chars _ = [*a*, *b*] where (*a*, *i*) = *getChar* 1
 (*b*, _) = *getChar* *i*

Lazy IO

Problem 3

getChar :: *Int* → *Char*

get2chars = [*getChar* 1, *getChar* 2]

- What will be called first?
- What we really need is some kind of dependency!
- The only kind of runtime dependency Haskell provides is value dependency:

getchar :: *Int* → (*Char*, *Int*)

get2chars _ = [*a*, *b*] where (*a*, *i*) = *getChar* 1
 (*b*, _) = *getChar* *i*

Lazy IO

Problem 3

getChar :: *Int* → *Char*

get2chars = [*getChar* 1, *getChar* 2]

- What will be called first?
- What we really need is some kind of dependency!
- The only kind of runtime dependency Haskell provides is value dependency:

getchar :: *Int* → (*Char*, *Int*)

get2chars _ = [*a*, *b*] **where** (*a*, *i*) = *getChar* 1
 (*b*, _) = *getChar* *i*

Lazy IO

Problem 4

The problem now is *get2chars* – the parameter is neither used nor unique and compiler might still reorder or omit it.

$$\text{get4chars} = [\text{get2chars } 1, \text{get2chars } 2]$$

What we need is a unique parameter associated with *every* IO action:

```
get2chars :: Int → (String, Int)
get2chars i0 = ([a, b], i2) where (a, i1) = getChar i0
                                (b, i2) = getChar i1
get4chars i0 = (a ++ b, i2) where (a, i1) = get2chars i0
                                (b, i2) = get2chars i1
```

Lazy IO

Problem 4

The problem now is *get2chars* – the parameter is neither used nor unique and compiler might still reorder or omit it.

$$\text{get4chars} = [\text{get2chars } 1, \text{get2chars } 2]$$

What we need is a unique parameter associated with *every* IO action:

$$\text{get2chars} :: \text{Int} \rightarrow (\text{String}, \text{Int})$$
$$\text{get2chars } i0 = ([a, b], i2) \text{ where } (a, i1) = \text{getChar } i0$$
$$(b, i2) = \text{getChar } i1$$
$$\text{get4chars } i0 = (a ++ b, i2) \text{ where } (a, i1) = \text{get2chars } i0$$
$$(b, i2) = \text{get2chars } i1$$

Lazy IO

Problem 5

However there is still one problem left, can you see it?:

*get2chars i0 = ([a, b], i2) where (a, i1) = getChar i0
(b, i2) = getChar i0*

A simple mistake and our carefully built up IO is ruined!

NB!

What we need is a way for compiler to ensure that the value is unique!

Lazy IO

Problem 5

However there is still one problem left, can you see it?:

*get2chars i0 = ([a, b], i2) where (a, i1) = getChar i0
(b, i2) = getChar i0*

A simple mistake and our carefully built up IO is ruined!

NB!

What we need is a way for compiler to ensure that the value is unique!

Lazy IO

Problem 5

However there is still one problem left, can you see it?:

*get2chars i0 = ([a, b], i2) where (a, i1) = getChar i0
(b, i2) = getChar i0*

A simple mistake and our carefully built up IO is ruined!

NB!

What we need is a way for compiler to ensure that the value is unique!

Outline

- 1 Functional IO
- 2 Monadic IO
- 3 Higher Order IO

Some History

When Haskell was created three main ways for doing pure, lazy IO were known.

- 1 Input-output streams
- 2 Continuations
- 3 Threaded unique value

Of course Haskellers chose the fourth way!

Clean

Interestingly enough a Haskell-like language Clean chose to introduce *uniqueness typing* that made the third way possible directly

Some History

When Haskell was created three main ways for doing pure, lazy IO were known.

- 1 Input-output streams
- 2 Continuations
- 3 Threaded unique value

Of course Haskellers chose the fourth way!

Clean

Interestingly enough a Haskell-like language Clean chose to introduce *uniqueness typing* that made the third way possible directly

Some History

When Haskell was created three main ways for doing pure, lazy IO were known.

- 1 Input-output streams
- 2 Continuations
- 3 Threaded unique value

Of course Haskellers chose the fourth way!

Clean

Interestingly enough a Haskell-like language Clean chose to introduce *uniqueness typing* that made the third way possible directly

Some History

When Haskell was created three main ways for doing pure, lazy IO were known.

- 1 Input-output streams
- 2 Continuations
- 3 Threaded unique value

Of course Haskellers chose the fourth way!

Clean

Interestingly enough a Haskell-like language Clean chose to introduce *uniqueness typing* that made the third way possible directly

Some History

When Haskell was created three main ways for doing pure, lazy IO were known.

- 1 Input-output streams
- 2 Continuations
- 3 Threaded unique value

Of course Haskellers chose the fourth way!

Clean

Interestingly enough a Haskell-like language Clean chose to introduce *uniqueness typing* that made the third way possible directly

Some History

When Haskell was created three main ways for doing pure, lazy IO were known.

- 1 Input-output streams
- 2 Continuations
- 3 Threaded unique value

Of course Haskellers chose the fourth way!

Clean

Interestingly enough a Haskell-like language Clean chose to introduce *uniqueness typing* that made the third way possible directly

Real World IO

We continue the previous example by interpreting the threaded value as the token of the world state:

$$getChar :: RealWorld \rightarrow (Char, RealWorld)$$

In that case the program is of type

$$main :: RealWorld \rightarrow ((), RealWorld)$$

We can introduce a type synonym and rewrite it as:

$$\text{type } IO\ a = RealWorld \rightarrow (a, RealWorld)$$
$$main :: IO\ ()$$
$$getChar :: IO\ Char$$

Real World IO

We continue the previous example by interpreting the threaded value as the token of the world state:

$$getChar :: RealWorld \rightarrow (Char, RealWorld)$$

In that case the program is of type

$$main :: RealWorld \rightarrow ((), RealWorld)$$

We can introduce a type synonym and rewrite it as:

$$\text{type } IO\ a = RealWorld \rightarrow (a, RealWorld)$$
$$main :: IO\ ()$$
$$getChar :: IO\ Char$$

Hello, Monad!

Let's see a program reading two characters from input:

```
main world0 = let (a, world1) = getChar world0  
              (b, world2) = getChar world1  
              in ([a, b], world2)
```

We could hide the token from the programmer by threading it automatically:

```
(>>=) :: IO a → (a → IO b) → IO b  
(action1 >>= action2) world0 =  
  let (a, world1) = action1 world0  
      (b, world2) = action2 a world1  
  in (b, world2)  
  
main =  
  getChar >>= λa → getChar >>= λb → (λw → ([a, b], w))
```

Hello, Monad!

Let's see a program reading two characters from input:

```
main world0 = let (a, world1) = getChar world0
               (b, world2) = getChar world1
               in ([a, b], world2)
```

We could hide the token from the programmer by threading it automatically:

```
(>>=) :: IO a → (a → IO b) → IO b
(action1 >>= action2) world0 =
  let (a, world1) = action1 world0
      (b, world2) = action2 a world1
  in (b, world2)

main =
  getChar >>= λa → getChar >>= λb → (λw → ([a, b], w))
```

Doing It with Class

Now we can make *IO a* a monad!

```
instance Monad (IO a) where
  (>>=) :: IO a → (a → IO b) → IO b
  (action1 >>= action2) world0 =
    let (a, world1) = action1 world0
      (b, world2) = action2 a world1
    in (b, world2)
  return :: a → IO a
  return a = (λworld → (a, world))
```

We can now use the *do*-notation so the previous example becomes

```
main = do
  a ← getChar
  b ← getChar
  return [a, b]
```

Doing It with Class

Now we can make $IO\ a$ a monad!

```
instance Monad (IO a) where
  (>>=) :: IO a → (a → IO b) → IO b
  (action1 >>= action2) world0 =
    let (a, world1) = action1 world0
        (b, world2) = action2 a world1
    in (b, world2)
  return :: a → IO a
  return a = (λworld → (a, world))
```

We can now use the *do*-notation so the previous example becomes

```
main = do
  a ← getChar
  b ← getChar
  return [a, b]
```

Interlude

Reflections

- It is logical to make *IO a* abstract, so that programmer would be protected from making mistakes.
- Then a value of *IO a* is nothing more than an *action* or a *computation* to be executed.
- However this also means that all IO functions must be predefined with priority access to IO monad.

IO functions

IO module necessarily contains functions for working with:

- Files and file systems
- Mutable variables
- Random numbers
- System clock

Which makes it pretty big!

Interlude

Reflections

- It is logical to make *IO a* abstract, so that programmer would be protected from making mistakes.
- Then a value of *IO a* is nothing more than an *action* or a *computation* to be executed.
- However this also means that all IO functions must be predefined with priority access to IO monad.

IO functions

IO module necessarily contains functions for working with:

- Files and file systems
- Mutable variables
- Random numbers
- System clock

Which makes it pretty big!

Interlude

Reflections

- It is logical to make *IO a* abstract, so that programmer would be protected from making mistakes.
- Then a value of *IO a* is nothing more than an *action* or a *computation* to be executed.
- However this also means that all IO functions must be predefined with priority access to IO monad.

IO functions

IO module necessarily contains functions for working with:

- Files and file systems
- Mutable variables
- Random numbers
- System clock

Which makes it pretty big!

Interlude

Reflections

- It is logical to make *IO a* abstract, so that programmer would be protected from making mistakes.
- Then a value of *IO a* is nothing more than an *action* or a *computation* to be executed.
- However this also means that all IO functions must be predefined with priority access to IO monad.

IO functions

IO module necessarily contains functions for working with:

- Files and file systems
- Mutable variables
- Random numbers
- System clock

Which makes it pretty big!

Outline

- 1 Functional IO
- 2 Monadic IO
- 3 Higher Order IO**

Console Operations

Definition

Most of console functions are self-explanatory:

```
putChar :: Char → IO ()  
putStr :: String → IO ()  
putStrLn :: String → IO ()  
print :: Show a ⇒ a → IO ()  
readLn :: Read a ⇒ IO a  
  
getChar :: IO Char  
getLine :: IO String  
getContents :: IO String
```

Console Operations

Example

```
main = do  
  putStr "What is your name?"  
  a  $\leftarrow$  readLn  
  putStr "How old are you?"  
  b  $\leftarrow$  readLn  
  print (a, b)
```

IO Actions as Values

Example

Let's define a list of IO actions:

```
ioActions :: [IO ()]  
ioActions =  
  [(print "Hello!"),  
   (putStr "just kidding"),  
   (getChar >> return ())]
```

Remember, that $IO\ a = RealWorld \rightarrow (a, RealWorld)$, so these are usual functional values.

IO Actions as Values

Example

However these IO actions can also be used directly:

```
main = do  
  head ioActions  
  ioActions !! 1  
  last ioActions
```

Output:

```
"Hello!"  
just kidding (Wait for input)
```

IO Action Lists

Definition

Sequencing a list of actions can be reused (*foldr* is just a reminder):

$$\text{foldr} \quad :: (a \rightarrow b \rightarrow b) \rightarrow b \rightarrow [a] \rightarrow b$$
$$\text{foldr } f \ z \ [] = \quad \quad \quad z$$
$$\text{foldr } f \ z \ (x : xs) = f \ x \ (\text{foldr } f \ z \ xs)$$
$$\text{sequence_} :: \text{Monad } m \Rightarrow [m \ a] \rightarrow m \ ()$$
$$\text{sequence_} = \text{foldr } (>>) \ (\text{return } ())$$

A lot more interesting monad combinators are available and we will examine them later.

IO Action Lists

Definition

The previous example then becomes:

```
ioActions :: [IO ()]  
ioActions =  
  [(print "Hello!"),  
   (putStr "just kidding"),  
   (getChar >> return ())]  
main = do sequence _ ioActions
```

Output:

```
"Hello!"  
just kidding (Wait for input)
```

File Operations

Definition

Files are opened as *Handles* and can be used accordingly:

```
openFile :: String → IOMode → IO Handle  
hSeek :: Handle → SeekMode → Integer → IO ()  
hGetChar :: Handle → IO Char  
hPutChar :: Handle → Char → IO ()  
hClose :: Handle → IO ()  
data SeekMode  
    = AbsoluteSeek  
    | RelativeSeek  
    | SeekFromEnd
```

File Operations

Example

Let's define a function reading one character from a given handle and position:

```
readi h i = do  
  hSeek h i AbsoluteSeek  
  hGetChar h
```

Next we want to open the file and just read characters from it:

```
readfilei :: String → IO (Integer → IO Char)  
readfilei name = do  
  h ← openFile name ReadMode  
  return (readi h)
```

File Operations

Example

Let's define a function reading one character from a given handle and position:

```
readi h i = do  
  hSeek h i AbsoluteSeek  
  hGetChar h
```

Next we want to open the file and just read characters from it:

```
readfilei :: String → IO (Integer → IO Char)  
readfilei name = do  
  h ← openFile name ReadMode  
  return (readi h)
```

File Operations

Example

Since IO functions can be higher order we can apply *readi* repeatedly:

```
main = do
  myfile ← readfilei "test"
  a ← myfile 0
  b ← myfile 1
  print (a, b)
```

IORef

Definition

IORef a allows to create mutable variables:

```
data IORef a  
newIORef :: a → IO (IORef a)  
readIORef :: IORef a → IO a  
writeIORef :: IORef a → a → IO ()  
modifyIORef :: IORef a → (a → a) → IO ()
```

IORef

Example

```
main = do  
  varA  $\leftarrow$  newIORef 0  
  a0  $\leftarrow$  readIORef varA  
  writeIORef varA 1  
  a1  $\leftarrow$  readIORef varA  
  print (a0, a1)
```

Haskell Objects

Figure

Let's try to define a *Figure* object:

```
data Figure = Figure{  
    draw :: IO (), move :: Displacement → IO ()}  
type Displacement = (Int, Int)
```

We will define only one figure – a circle:

```
circle :: Point → Radius → IO Figure  
type Point = (Int, Int)    -- point coordinates  
type Radius = Int         -- circle radius in points
```

Haskell Objects

Drawing

We can test the objects when we define them with this code:

```
main = do
  figures ← sequence [circle (10,10) 5,
    rectangle (10,10) (20,20)]
  drawAll figures
  mapM_ ( $\lambda fig \rightarrow move\ fig\ (10,10)$ ) figures
  drawAll figures

drawAll :: [Figure] → IO ()
drawAll figures = do
  putStrLn "Drawing figures:"
  mapM_ draw figures
```

Circle

```
circle center radius = do
  centerVar ← newIORef center
  let drawF = do
    center ← readIORef centerVar
    putStrLn ("  Circle at " ++ show center
      ++ " with radius " ++ show radius)
  let moveF (addX, addY) = do
    (x, y) ← readIORef centerVar
    writeIORef centerVar (x + addX, y + addY)
  return $ Figure{draw = drawF, move = moveF}
```

Rectangle

```
rectangle from to = do
  fromVar ← newIORef from
  toVar ← newIORef to
  let drawF = do
    from ← readIORef fromVar
    to ← readIORef toVar
    putStrLn ("  Rectangle " ++ show from
              ++ "-" ++ show to)
  let moveF (addX, addY) = do
    (fromX, fromY) ← readIORef fromVar
    (toX, toY) ← readIORef toVar
    writeIORef fromVar (fromX + addX, fromY + addY)
    writeIORef toVar (toX + addX, toY + addY)
  return $ Figure{draw = drawF, move = moveF }
```

Higher Order IO

Lazy IO

How lazy is the IO monad?

```
list1, list2 :: IO [Int]
list1 = return ◦ repeat $ 0
list2 = sequence ◦ repeat ◦ return $ 0
main = do
  lst ← list1
  putStrLn ◦ show ◦ take 10 $ lst
  readLn
  lst2 ← list2
  putStrLn ◦ show ◦ take 10 $ lst2
```

Higher Order IO

Lazy IO

How lazy is the IO monad?

```
list1, list2 :: IO [Int]
list1 = return ◦ repeat $ 0
list2 = sequence ◦ repeat ◦ return $ 0
main = do
  lst ← list1
  putStrLn ◦ show ◦ take 10 $ lst
  readLn
  lst2 ← list2
  putStrLn ◦ show ◦ take 10 $ lst2
```

The first list will show the first 10 elements, while the second one will go infinite.

Conclusions

Advantages

- IO in Haskell is pure and lazy
- The order of operations that end up executed is guaranteed by the compiler
- IO actions are treated in Haskell first-class, they can be passed around, combined and curried

Disadvantages

- Every new function that needs to do side-effects needs access to IO monad implementation
- IO actions tend to contaminate pure code, lifting it to IO monad because some particular parts need to do IO
- Since IO operations are the fastest, a large portion of a program may end up under IO monad

Conclusions

Advantages

- IO in Haskell is pure and lazy
- The order of operations that end up executed is guaranteed by the compiler
- IO actions are treated in Haskell first-class, they can be passed around, combined and curried

Disadvantages

- Every new function that needs to do side-effects needs access to IO monad implementation
- IO actions tend to contaminate pure code, lifting it to IO monad because some particular parts need to do IO
- Since IO operations are the fastest, a large portion of a program may end up under IO monad

Conclusions

Advantages

- IO in Haskell is pure and lazy
- The order of operations that end up executed is guaranteed by the compiler
- IO actions are treated in Haskell first-class, they can be passed around, combined and curried

Disadvantages

- Every new function that needs to do side-effects needs access to IO monad implementation
- IO actions tend to contaminate pure code, lifting it to IO monad because some particular parts need to do IO
- Since IO operations are the fastest, a large portion of a program may end up under IO monad

Conclusions

Advantages

- IO in Haskell is pure and lazy
- The order of operations that end up executed is guaranteed by the compiler
- IO actions are treated in Haskell first-class, they can be passed around, combined and curried

Disadvantages

- Every new function that needs to do side-effects needs access to IO monad implementation
- IO actions tend to contaminate pure code, lifting it to IO monad because some particular parts need to do IO
- Since IO operations are the fastest, a large portion of a program may end up under IO monad

Conclusions

Advantages

- IO in Haskell is pure and lazy
- The order of operations that end up executed is guaranteed by the compiler
- IO actions are treated in Haskell first-class, they can be passed around, combined and curried

Disadvantages

- Every new function that needs to do side-effects needs access to IO monad implementation
- IO actions tend to contaminate pure code, lifting it to IO monad because some particular parts need to do IO
- Since IO operations are the fastest, a large portion of a program may end up under IO monad

Conclusions

Advantages

- IO in Haskell is pure and lazy
- The order of operations that end up executed is guaranteed by the compiler
- IO actions are treated in Haskell first-class, they can be passed around, combined and curried

Disadvantages

- Every new function that needs to do side-effects needs access to IO monad implementation
- IO actions tend to contaminate pure code, lifting it to IO monad because some particular parts need to do IO
- Since IO operations are the fastest, a large portion of a program may end up under IO monad