
What Does It Mean To Be Reactive?

Crushing opinions with the power of Mathematics

By

EDDY BERTOLUZZO



Faculty of Electrical Engineering, Mathematics and Computer Science
DELFT UNIVERSITY OF TECHNOLOGY

A dissertation submitted to Delft University of Technology in
accordance with the requirements of the MASTER OF SCIENCE
degree in the Faculty of Electrical Engineering, Mathematics
and Computer Science.

OCTOBER 2016

ABSTRACT

In the last few years the *** of Reactive Programming has gained much traction in the developer's community, especially with the general trend for programming languages to embrace functional programming and the shifting of today's applications towards an asynchronous approach to modeling data. The goal of this work is to shed some light to the sea of more or less informed opinions regarding what the meaning of Reactive Programming is thorough the use of Mathematics...

DEDICATION AND ACKNOWLEDGEMENTS

Here goes the dedication.

TABLE OF CONTENTS

	Page
1 Introduction	1
1.1 Iterables	2
1.2 Into the rabbit hole: Deriving the Observable	3
1.3 Observables are Continuations	5
1.4 Out of the rabbit hole: Towards a usable API	7
1.4.1 Adding Termination and Error handling	7
1.4.2 Adding Subscriptions	9
1.5 The Reactive Contract	10
1.5.1 Subsection	10
A Appendix A	13
A.1 Duality	13
A.2 Continuation Passing Style	13
A.3 Product and Sum Types	13
A.4 Currying and Uncurrying	13

INTRODUCTION

Science Bitch Quote

The following write up is meant as a summary of what my thesis will expand upon. When thinking about effects and how they combine with different types of computations, we can come up with the following diagram that nicely summarizes them.

	One	Many
Synch	$f :: () \rightarrow a$	Iterable a
Asynch	$\text{callback} :: a \rightarrow ()$???

Find duality in the first column to justify the use in the second? Does what I wrote up there work?

When our goal is to retrieve a single value of type `a` synchronously, we will make use of a function whose return type is `a` itself. In an asynchronous setting, on the other hand, our function will return a `Future a`, i.e. a computation that at a certain point in the future will result in a value of type `a`.

When we are dealing with multiple values, we will typically return a collection which will contain our results. An `Iterable a` is an abstraction over collections which generalizes the concept of a data structure containing multiple values which are synchronously retrievable.

The purpose of this thesis is to explore the design space left empty in the above diagram and derive a solution to the problem of dealing with multiple asynchronous values in our programs.

1.1 Iterables

We begin by introducing the concept of `Iterable`, made famous by the Gang of Four pattern of the same name and widely adopted by most recent programming languages.

An `Iterable` is a programming structure which enables the user to traverse a collection of data, abstracting over the underlying implementation.

The interface and semantics of `Iterables` are fixed across programming languages.

Add C# version

```
1 class Iterable a where
2     getIterator :: () -> IO (Iterator a)
3
4 -- Java-like Iterator
5 class Iterator a where
6     next      :: () -> IO a
7     hasNext  :: () -> Bool
8
9 -- C#-like Iterator
10 class Iterator a where
11     current   :: () -> a
12     moveNext  :: () -> IO Bool
```

Note that different programming languages might expose slightly different interfaces for the `Iterator` construct. These differences affect the way the API is used by programmes, but do not affect the essence of the `Iterator` pattern and therefore are not relevant. For this reason and without loss of generality, we will make use of the Java-like definition of `Iterator` in the reminder of the discussion.

We will now make use of a few mathematical concepts in order to explore the design space around collections of asynchronous values: duality, continuations and CPS, products and coproducts and lastly currying and uncurrying. See Appendix A for a quick introduction to these concepts.

Let's start by stripping down the `Iterable` interface of all the unnecessary API/imperative design features so that we can derive it's essence and easily reason about it's properties.

The first step is simplifying the `Iterator` interface: when we take a close look, we can easily see that an `Iterator` is nothing but a function which returns either a value - when `hasNext` returns

true and next is called - or nothing - when hasNext returns false. Additionally, and Iterator can also throw an exception in the case that next is called when the underlying collection has been exhausted, something that was not explicit in the C# definition.

Merging these consideration with the notion of coproducts and the maybe types, we obtain the following definition of Iterable.

```
1 class Iterable a where
2   getIterator :: () -> IO (Iterator a)
3 class Iterator a where
4   moveNext :: () -> IO (Either Error (Maybe a))
```

Note how the moveNext function still produces an `IO` computation.

Let's now go one step further and simplify our definition even more, with the purpose of understanding the real essence of an Iterable. In order to achieve this, let's forget for a moment about errors and no values returned from an iterator, i.e. let's ignore operational concerns such as exceptions and termination.

```
1 type Iterable a = () -> IO (() -> IO a)
2 type Iterator a = () -> IO a
```

An Iterable is simply a function which, when invoked, returns an Iterator and an Iterator is itself a function which returns a computation that will eventually result in a value of type `a`. In some way, we can think of an Iterator as a getter function, and of an Iterable as a getter of a getter.

1.2 Into the rabbit hole: Deriving the Observable

As we noticed before, the way we deal with a single value in an synchronous environment is dual to the asynchronous method, flipping the arrows brings us from a function resulting in a value of type `a` to one accepting an `a`.

It seems reasonable now to apply the same concept of duality to the Iterable in order to find a possible solution to our original problem, dealing with a collection of asynchronous values.

```
1 type Iterable a = () -> IO (() -> IO a)
2 type Iterator a = () -> IO a
3
4 type Observable a = (a -> IO ()) -> IO ()
5 type Observer a = a -> IO ()
```

It is straightforward to see how duality plays out in the derivation of the new types.

This dualisation leaves us with an interface which is interesting under many point of views. First of all, to link back to the observation that an Iterable is a getter of a getter, we can observe that the Observable plays exactly the opposite role, that is, a setter of a setter, the Observer.

This brings us to another interesting observation, the Iterable embodies the idea of pull collections, the Iterable will give us a new Iterator whenever we ask for it and the Iterator, in turn, will provide us with the next value whenever we decide to pull one from it. Dually, the Observable interface embodies the idea of push collections: we push a callback, the Observer, inside the Observable, which, in turn, will push a value into the Observer whenever one becomes available.

From here we can take two roads, one is to analyze further the type that we obtained and see what it's properties are and what we can understand from it; the second road is to un-simplify its definition and augment it so as to arrive to a usable API for handling asynchronous collections. We will start with the first.

For those who are familiar with functional programming it will not be hard to see how the Observable type resembles a function written in continuation passing style.

```
1 cont      :: (a -> r) -> r
2 observable :: (a -> IO ()) -> IO ()
```

We can easily observe how an Observable is nothing more than a CPS function where the result type `r` is instantiated to `IO ()`.

This observation opens up a great deal of mathematics properties and laws that we can prove about observables.

First of all, let's convince ourselves that a CPS function, and therefore an Observable, truly embodies a push based model of computation. The very definition of continuation tells us that a continuation represents the "rest of the computation"; when looking at it from an observable prospective, the continuation, i.e. the observer, is the function that specifies what needs to happen

to a value, whenever this becomes available, that is, whenever the observable pushes it to the observer. Since the continuation can be called multiple times, it is easy to see how this structure lets us deal with multiple values that might come in at different times in the future.

Note how an observable is "paused" until it receives an observer to which it can push the values it produces. This comes precisely from the definition of CPS function, i.e. a suspended computation which, given another function as argument, the continuation, will produce the final result.

Therefore an observable is a suspended computation that will start producing values once we pass in an observer, the continuation that will specify what happens to a value once it is computed. The observable will call the observer every time it produces a value.

1.3 Observables are Continuations

Talk about the interface that we want to get, even a minimal one now and explain how this relates to the cont monad, e.g. `subscribe = ContT`. Introduce the `ContT` monad and show the minimal reference implementation.

Our goal is now to start from this theoretical explanation of a push based collection and build a usable API. In the previous section we observed how and `Observable` is a particular case of the continuation monad where the result type is a computation which has `unit` as a result type. The Haskell language provides a monad construct for expressing continuations as well as a monad transformer in order to stack continuations on top of other monads. A monad transformer is exactly what we need in order to express our continuations resulting in `IO`.

```
1 type Observable a = ContT () IO a
2 type Observer a = a -> IO ()
```

Our API will also need to provide functions to create as well as start the `Observable` stream.

```
1 -- Simply wraps the function :: (a -> IO ()) -> IO ()
2 -- inside the Observable datatype
3 observable :: (Observer a -> IO ()) -> Observable a
4 observable = ContT
5
6 -- Runs the Observable by providing the continuation - i.e. the Observer -
7 -- that will handle the asynchronous data.
8 subscribe :: Observable a -> Observer a -> IO ()
9 subscribe = runContT
```

We can already notice how This should come as no surprise after our discussion regarding the connection between Observables and the Continuation monad. Conversely, these equivalences should justify even more the use of this model of computation for reactive programming.

At this point we have all the necessary tools to create and run an Observable

```
1 let obs = observable $ \obr ->
2     do obr 1
3         obr 2
4         obr 3
5
6 subscribe obs print
7
8 output>
9 1
10 2
11 3
```

The example above is a toy example, let's try with a more realistic one such that we can show that our basic implementation of rx based on continuations works just as well as a full blown one in terms of handling asynchronous data.

Put the actual demo that listens to keyboard presses. In the example, press 3 times and see that all the events are handled and none are lost.

```
13 obs = observable $ \obr -> do
14     passiveMotionCallback $= Just (\p -> obr p)
15
16 main :: IO ()
17 main = do
18     (_progName, _args) <- getArgsAndInitialize
19     _window <- createWindow "Hello World"
20     subscribe obs print
21     mainLoop
```

Elaborate better this part. Ask Erik for input.

It is worth noting that rx in itself is not at all async in handling data unless we use schedulers, although it does handle async data. This is a common misconception, even if you have a single

thread that doesn't mean you cannot handle async data, actually you are async because the control flow isn't linear. Naturally, the thing is that certain queries (the ones that don't use schedulers) will simply block your single thread and prevent other things from happening.

This demo shows exactly this, even though the processing of the data is synchronous, the data itself, being mouse movements, is inherently async.

At this point in the discussion we have a working implementation of a push based collection purely derived from the underlying theory of duality and continuations. The next step in augmenting out library is to note that continuations are monads and, being the Observable an instance of the continuation monad, it is itself a monad. This observation comes with great benefits, we can get mathematical laws - the monad laws - proven for free for our structure and, from a more practical point of view, we get functions defined for free for the Observable: `fmap` (from Functor), `flatMap` and `return`.

Say more on the monad laws, even if they are trivially proved.

We will soon see how the implementation of these functions will change and move from the standard one the moment we start moving towards a more operational implementation for real world use.

1.4 Out of the rabbit hole: Towards a usable API

Start from the minimal reference implementation, add exceptions and termination, explain how `flatMap` changes and lastly add subscription and talk about cancellation and the contract.

Up until now we have analyzed the essence of the Observable, leaving out the operational concerns that would come up when our goal is to design a usable API. Now we will slowly and step by step re-introduce these concerns in order to go from a theoretical definition of Observables to a more operational and therefore usable one.

Remember how, at the beginning of our discussion, we greatly simplified the Iterable interface in order to derive a type that represents the essence of an Iterable. We later applied the duality principle from category theory in order to derive the Observable. Now, we want to walk the simplification path backwards and, step by step, re-introduce all that we simplified before in order to arrive to a usable API.

1.4.1 Adding Termination and Error handling

The first thing we want to re-introduce is handling exceptions and termination of a stream. Where an Iterable can return a value, terminate or throw an exception when we ask for a value, and

Observable, being its dual, can produce one or more values, terminate or throw an exception when it is subscribed to.

A more appropriate type for our interface is then the following.

```
4 type Observable a = ContT () IO (Either Error (Maybe a))
5 type Observer a = Either Error (Maybe a) -> IO ()
```

Now, this code is not exactly the definition of readable; let's apply some good design skills to make it more pleasant to the eye without changing its meaning.

```
4 -- datatype representing Either Error (Maybe a)
5 data Event a =
6     OnNext a
7     | OnError SomeException
8     | OnCompleted
9
10 type Observable a = ContT () IO (Event a)
11 type Observer a = Event a -> IO ()
```

Although this might not look like a big change, it greatly influences the design of our API. We are, in fact, changing our instantiation of the continuation monad to an input type that is not `a` anymore, but `Event a`. On the other hand, our type variable for `Observable` is still `a`. This is not an issue per se, but it has one big consequence: the `flatmap` function that we inherit from the continuation monad is not the one that we want to expose from our API anymore. The types differ like so.

```
13 -- bind inherited from the Continuation monad
14 (>>=) :: Observable a -> (Event a -> Observable b) -> Observable b
15
16 -- bind that we would like to expose from our API
17 flatmap :: Observable a -> (a -> Observable b) -> Observable b
```

This has many implications, first of all, we are gonna need to implement `flatmap` by ourselves.. see todo below...

ASK ERIK: Observable being a type synonym, it still is a monad in the cont sense (i.e. with a instantiated to `Event a`). Do we need to instantiate monad with our `flatmap` definition and prove the laws there as well?


```

19  -- Monad Laws
20  -- return a >>= k  =  k a
21  -- m >>= return  =  m
22  -- m >>= (x -> k x >>= h)  =  (m >>= k) >>= h

```

We could have implemented Observer as a datatype containing 3 functions, but that would complicate the explanation of Event a and the flatmap implementation problem, so I chose to go with the Event a

Modify keyboard press example from before to handle errors and termination. Point to later discussion regarding the rx contract, since now we can detect termination and errors but there is no guarantee that nothing will come after we receive them, i.e. that we abide the contract.

1.4.2 Adding Subscriptions

With error handling and termination, our implementation starts getting more and more usable. We now want to add a mechanism that will allow us to stop an observable stream from the outside (as opposed to waiting for an onCompleted) whenever we don't require it's data anymore.

In order to achieve this we will use a new datatype, Subscription. The idea is that the subscribe function will return a Subscription to the user who will later be able to call unsubscribe on it and stop the stream associated to it from producing any more values.

Talk about the best effort in canceling work and eventual consistency with the contract.

A Subscription in itself is nothing more than an action to perform once we stop an observable stream.

```

1  type Subscription = IO ()

```

Our goal now is to implement a cancellation mechanism that would stop the Observable. This will be achieved by calling a function unsubscribe, which will prevent, from that moment on, any events to be signaled to any subscribed observer.

This is achieved by wrapping the user supplied observer to the subscribe function with an internal one which adds this functionality and forwards all accepted events to the supplied one.

the code for safe observer.

```

1  type Subscription = IO ()

```

As a design decision, when unsubscribe is called on an observable subscription, the observable sequence will make a best effort attempt to stop all outstanding work. This means that any

queued work that has not been started will not start. Any work that is already in progress might still complete as it is not always safe to abort work that is in progress. Results from this work will not be signaled to any previously subscribed observer instances.

Motivation for children subscriptions: some operators will create inner observables and therefore will need to unsubscribe from them when the outer observable is unsubscribed from.

1.5 The Reactive Contract

Begins a section.

1.5.1 Subsection

Begins a subsection.

TODO LIST

■ Science Bitch Quote	1
■ Find duality in the first column to justify the use in the second? Does what I wrote up there work?	1
■ Add C# version	2
■ Talk about the interface that we want to get, even a minimal one now and explain how this relates to the cont monad, e.g. subscribe = ContT. Introduce the ContT monad and show the minimal reference implementation.	5
■ Put the actual demo that listens to keyboard presses. In the example, press 3 times and see that all the events are handled and none are lost.	6
■ Elaborate better this part. Ask Erik for input.	6
■ Say more on the monad laws, even if they are trivially proved.	7
■ Start from the minimal reference implementation, add exceptions and termination, explain how flatmap changes and lastly add subscription and talk about cancellation and the contract.	7
■ ASK ERIK: Observable being a type synonym, it still is a monad in the cont sense (i.e. with a instantiated to Event a). Do we need to instantiate monad with our flatmap definition and prove the laws there as well?	8
■ We could have implemented Observer as a datatype containing 3 functions, but that would complicate the explanation of Event a and the flatmap implementation problem, so I chose to go with the Event a	9

■ Modify keyboard press example from before to handle errors and termination. Point to later discussion regarding the rx contract, since now we can detect termination and errors but there is no guarantee that nothing will come after we receive them, i.e. that we abide the contract.	9
■ Talk about the best effort in canceling work and eventual consistency with the contract.	9
■ the code for safe observer.	9
■ Motivation for children subscriptions: some operators will create inner observables and therefore will need to unsubscribe from them when the outer observable is unsubscribed from.	10

**APPENDIX A**

Begins an appendix

A.1 Duality**A.2 Continuation Passing Style****A.3 Product and Sum Types****A.4 Currying and Uncurrying**

