



**INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA**

**Área Departamental de Engenharia de Electrónica e Telecomunicações e de  
Computadores**

## **Inferring an Iterator from a bulk traversing function**

**DIOGO RAFAEL ESTEVES POEIRA**

Licenciado em Engenharia Informática e de Computadores

Dissertação para obtenção do Grau de Mestre  
em Engenharia Informática e de Computadores

Orientador : Doutor Fernando Miguel Gamboa de Carvalho

**Junho, 2018**





**INSTITUTO SUPERIOR DE ENGENHARIA DE LISBOA**

**Área Departamental de Engenharia de Electrónica e Telecomunicações e de  
Computadores**

## **Inferring an Iterator from a bulk traversing function**

**DIOGO RAFAEL ESTEVES POEIRA**

Licenciado em Engenharia Informática e de Computadores

Dissertação para obtenção do Grau de Mestre  
em Engenharia Informática e de Computadores

Orientador : Doutor Fernando Miguel Gamboa de Carvalho

**Junho, 2018**



# Contents

<b>Contents</b>	<b>v</b>
<b>Listings</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Traversing sequences . . . . .	3
1.3 Use Case . . . . .	5
<b>2 Background and State of the Art</b>	<b>7</b>
2.1 Spliterator . . . . .	7
2.2 Guava . . . . .	8
2.3 StreamEx . . . . .	10
2.4 JOOL and Cyclops . . . . .	12
2.5 Vavr . . . . .	12
<b>3 Solution</b>	<b>13</b>
3.1 Traverser . . . . .	13
3.2 Yield . . . . .	14
3.3 Query . . . . .	15
3.4 Advancer . . . . .	15

3.4.1	Approach . . . . .	16
3.4.2	SerializedLambda . . . . .	18
3.4.3	Visitors . . . . .	18
3.4.4	State Machine . . . . .	19
<b>Bibliography</b>		<b>23</b>
<b>Bibliography</b>		<b>23</b>

# Listings

1.1	Fruit . . . . .	2
1.2	Basket . . . . .	2
1.3	Collections approach . . . . .	2
1.4	Stream approach . . . . .	3
1.5	Splititerator example . . . . .	4
1.6	Jayield filterOdd . . . . .	5
2.1	Guava's approach . . . . .	8
2.2	Guava's approach . . . . .	8
2.3	StreamEx's approach . . . . .	11
3.1	Traverser . . . . .	13
3.2	Traverser map to name example . . . . .	13
3.3	Traverser as sequence generator . . . . .	14
3.4	Yield . . . . .	14
3.5	Yield example . . . . .	15
3.6	Query then method . . . . .	15
3.7	Advancer . . . . .	15
3.8	Advancer example . . . . .	16
3.9	Map Traverser . . . . .	17
3.10	Map Advancer . . . . .	17
3.11	Generated Advancer . . . . .	19

3.12 Duplicate Traverser function . . . . .	19
3.13 Duplicate Advancer translation with if based State Machine . . . .	20
3.14 Duplicate Advancer translation with switch based State Machine .	21





# Introduction

Data structures traversal can be achieved with several approaches and currently in Java it can be done in three different ways:

- Iterating elements with an Iterator object
- Traversing elements individually (i.e. `tryAdvance()`)
- Traversing elements in bulk (i.e. `forEachRemaining()`)

In `java.util.stream` the implementation of new query operations requires the definition of both individually and bulk traversal algorithms. This incurs in undesired verbosity. In this project we propose a solution - Jayield - which infers the individual element traversal definition from a given bulk traversal method, and thus allows the implementation of new query operations with a single method (e.g. lambda expression).

## 1.1 Motivation

Java 8 brings with it the `java.util.stream` API (henceforth named just Stream) which allows us to perform queries fluently on a sequences of data, by providing methods that accept behaviour as parameters, allowing the programmers to express what their intentions are when calling the method.

Let us consider the following two classes.

```
1 public class Fruit {
2     String name;
3     String color;
4
5     public Fruit(String name, String color) {
6         this.name = name;
7         this.color = color;
8     }
9 }
```

Listing 1.1: Fruit

```
1 import static java.util.Arrays.asList;
2
3 public class Basket {
4
5     Fruit orange = new Fruit("Orange", "orange");
6     Fruit lemon = new Fruit("Lemon", "yellow");
7     ...
8     Fruit mango = new Fruit("Mango", "orange");
9
10    public final List<Fruit> fruits = asList(orange, lemon, mango, ...);
11
12 }
```

Listing 1.2: Basket

Listing 1.1 specifies the definition of a Fruit and listing 1.2 makes use of said definition and specifies the fruits inside a basket. Now, let's say we want to know the names of all the fruits inside the basket that have the color orange. In Java 7 and previous versions we would do something like this:

```
1 public static List<String> getOrangeFruitList(Basket basket) {
2     Iterator<Fruit> iterator = basket.fruits.iterator();
3     List<String> result = new ArrayList<Fruit>();
4     while (iterator.hasNext()) {
5         Fruit current = iterator.next();
6         if (current.color.equals("orange")) {
7             result.add(current.name);
8         }
9     }
10    return result;
11 }
```

Listing 1.3: Collections approach

As we can see in, we would attain an iterator and as we iterate over the fruits on the basket we check their color, if it matches our desired color we add it's name to the list and when we are done, we return the new list. Now lets look at how we can do this in Java 8 making use of Streams.

---

```
1 public static Stream<String> getOrangeFruitStream(Basket basket) {  
2     return basket.fruits.stream()  
3         .filter(current -> current.color.equals("orange"))  
4         .map(current -> current.name);  
5  
6 }
```

---

Listing 1.4: Stream approach

This approach lets us define our intentions, first we say we are only interested in fruits with the color orange, and that for each fruit that matches that description we are only interested in it's name. This approach lets us specify our intentions clearly which in turn makes the code easier to read and understand.

However, the end use API is not the only difference between these two approaches. The collections approach is eager, meaning, it immediately goes through all the fruits in the basket to select the ones we are interested, taking note of the names as we go through them. The stream approach is lazy, as we said before, all we are doing is specifying our intention, no iteration is made on the sequence until it is consumed by a terminal operation, such as a `forEach()`.

## 1.2 Traversing sequences

Going back to our initial example, the basket of fruits, we now have the names of all the orange fruits in our basket. Suppose that we want to take the names of the fruits in our basket alternately, skipping the even items and keeping the odd items. To that end we would like to have a `filterOdd()` operation chained in the pipeline of filter-map of listing 1.4. However, the Stream API that Java 8 provides does not support this operation, so we have to implement it ourselves.

Java 8 Provides different ways of creating a new Stream. Since we want to preserve the laziness characteristic of streams we should implement the new Stream through a traversal protocol. In this case, Java 8 introduces a new iteration approach specified by the `Splititerator` interface which we can implement as well as providing the `AbstractSplititerator` class that we can extend.

In listing 1.5 we show an example of the implementation of a filterOdd operation (line 2) which in turn requires an auxiliary class implementing spliterator (i.e. OddFilter, lines 3 and 5). In this case we extend the AbstractSpliterator instead of implementing the Spliterator interface to reduce the number of methods we're required to implement.

---

```

1 public class Example {
2     static <T> Stream <T> filterOdd (Stream <T> source ) {
3         return StreamSupport.stream(new OddFilter<>(source.spliterator()),
4             false);
5     }
6     static class OddFilter <T> extends AbstractSpliterator <T> {
7         final Consumer <T> doNothing = item -> {};
8         final Spliterator <T> source ;
9         boolean isOdd = false;
10        public Odd (Spliterator <T> source) {
11            super (odd(source.estimateSize()), source.characteristics());
12            this.source = source;
13        }
14        @Override
15        public boolean tryAdvance(Consumer <? super T> action ) {
16            if (!source.tryAdvance(doNothing)) return false ;
17            return source.tryAdvance(action);
18        }
19        @Override
20        public void forEachRemaining(Consumer<? super T> action) {
21            source.forEachRemaining(item -> {
22                if(isOdd){
23                    action.accept(item);
24                }
25                isOdd = !isOdd;
26            });
27        }
28        private static long odd( long l) {
29            return l == Long.MAX_VALUE ? l : (l +1)/2;
30        }
31        public static void main(String[] args) {
32            filterOdd(getOrangeFruitStream(new Basket()))
33                .forEach(System.out::println);
34        }
35    }

```

---

Listing 1.5: Spliterator example

This approach is verbose, and only benefits from bulk traversing because we have overridden `forEachRemaining()`, as the `Splititerator` documentation states, the `forEachRemaining()` [8] method's default implementation should be overridden whenever possible as it just calls `tryAdvance()` until it returns false, in other words, it will traverse element by element instead of the whole sequence. We will be discussing this approach further in Chapter 2.

To overcome the shortcomings found in the Stream API we will implement a solution that allows the user to define a single way to traverse a sequence and we will generate its counterpart. In this case the `filterOdd()` operation can be expressed in a lambda expression chained with the sequence of items as shown in listing 1.6.

---

```
1 getOrangeFruitQuery(new Basket())
2     .then(source -> {
3         boolean[] isOdd = {false};
4         return yield -> source.traverse(item -> {
5             if(isOdd[0]) {
6                 yield.ret(item);
7             }
8             isOdd[0] = !isOdd[0];
9         });
10    });
```

---

Listing 1.6: Jayield filterOdd

The programmer just needs to write one method to be able to traverse a sequence as the individually advance will be inferred from the previous definition.

## 1.3 Use Case

The use case that will be analysed, developed and tested will be the generation of an Advancer from either a Query or a Traverser.

The difficulty with this use case is the fact that a Traverser is used to bulk traverse an entire sequence, meaning it is not (or should not be) prepared to verify if it can advance at each element of the sequence. To do this we will copy the Traverser's code and instrument the required parts to generate a new Advancer. To test if the generated Advancer is working properly we will verify that each call to `tryAdvance` that succeeds, yields one element and only one element.





## Background and State of the Art

In order to further understand the context of this project we will be further describing the Splititerator interface, some alternatives to the Java 8 Stream API as well as discussing their advantages and disadvantages.

### 2.1 Splititerator

So what was the reasoning behind the Splititerator approach, why not just an Iterator? One of the reasons was to achieve parallelism in the iteration of the sequence. From a use case point of view, most of the time, we do not wish to parallelize work, most of our data isn't big enough to justify the parallelization.

Another reason was the implementation of short-circuiting operations, as we have two methods in Splititerator that we can override to traverse a sequence, `tryAdvance()` and `forEachRemaining()`, `tryAdvance()` is stipulated to individually traverse a sequence and therefore the one that should be used when short-circuiting a traversal. However, there are other ways to stop the traversing of a sequence, namely we can simply raise an exception when bulk traversing and the traversal will stop. Granted, this is somewhat of a polemic approach, but there is no hard evidence that states that it isn't a valid one, in fact, Python uses this approach, and nowadays there are lightweight Exceptions that can be used to this effect.

Lastly, and from our point of view, the most valid reason, operations that combine two sequences. Operations such as Zip that require that an element from each sequence be taken to combine it into a single element, cannot be achieved by a bulk traversing function. Oddly enough Java's Stream API does not implement the Zip operation.

## 2.2 Guava

Guava [3] was for some years the library most programmers defaulted to when looking for Functional programming utilities in Java, until Java 8 arrived. To achieve this functionality the Guava library provides their own sequential type, `FluentIterable`, which provides a way to chain operations together like `Stream` does.

In listing 2.1 we show how the previous example, where we obtain the names of all the orange fruits, could be achieved using the Guava library and lambda functions.

---

```
1 import com.google.common.collect.FluentIterable;
2
3 import static com.google.common.collect.FluentIterable.from;
4
5 public class GuavaExample {
6     public FluentIterable<String> getOrangeFruits(Basket basket) {
7         return from(basket.fruits)
8             .filter(fruit -> fruit.color.equals("orange"))
9             .transform(fruit -> fruit.name);
10    }
11 }
```

---

Listing 2.1: Guava's approach

Just like the `Stream` approach, here we can specify our intentions without traversing the sequence. When it comes to extending `FluentIterables`'s operations even further however, the programmer has to implement an `Iterator`, like with `Streams` one would have to implement a `Spliterator`. So if presented with the same outcome as before, an implementation of `filterOdd` would look something like what is shown in listing 2.2 which results in a quite verbose approach and not what we are aiming to achieve.

---

```
1 import com.google.common.collect.FluentIterable;
2
```



```
3 import java.util.Iterator;
4 import java.util.NoSuchElementException;
5
6 import static com.google.common.collect.FluentIterable.from;
7
8 public class FilterOdd<T> implements Iterator<T> {
9     public static <T> FluentIterable<T> filterOdd(FluentIterable<T>
10         source) {
11         return from(() -> new FilterOdd<>(source.iterator()));
12     }
13
14     private final Iterator<T> source;
15     private T current;
16
17     public FilterOdd(Iterator<T> source) {
18         this.source = source;
19     }
20
21     @Override
22     public boolean hasNext() {
23         if(current == null && source.hasNext()) {
24             source.next(); // ignore even
25             if(source.hasNext()) {
26                 current = source.next(); // get odd
27             }
28         }
29         return current != null;
30     }
31
32     @Override
33     public T next() {
34         if(hasNext()) {
35             T aux = current;
36             current = null;
37             return aux;
38         } else {
39             throw new NoSuchElementException();
40         }
41     }
42 }
```

---

Listing 2.2: Guava's approach

## 2.3 StreamEx

StreamEx [1] is the only library, of those discussed in this document, that provides function based extensibility, instead of requiring the implementation of a specific type. For instance, looking back at our examples, to implement a filterOdd operation we would only need to define what is needed for the tryAdvance method described in listing 1.5 which iterates over two items whenever the downstream request one item. Listing 2.3 shows an example.

---

```
1 import one.util.streamex.StreamEx;
2
3 import java.util.Spliterator;
4 import java.util.function.Consumer;
5
6 import static one.util.streamex.StreamEx.of;
7 import static one.util.streamex.StreamEx.produce;
8
9 public class StreamExExample {
10
11     public static <T> StreamEx<T> filterOdd(StreamEx<T> source) {
12         final Consumer<T> doNothing = item -> {};
13         final Spliterator<T> iterator = source.spliterator();
14         return produce(action -> {
15             if (!iterator.tryAdvance(doNothing))
16                 return false;
17             return iterator.tryAdvance(action);
18         });
19     }
20
21     public StreamEx<String> getOrangeFruits(Basket basket) {
22         return of(basket.fruits)
23             .filter(fruit -> fruit.color.equals("orange"))
24             .map(fruit -> fruit.name);
25     }
26 }
```

---

Listing 2.3: StreamEx's approach

This approach is less verbose than the one implemented by Guava and Java 8's Streams. However StreamEx's implementation still has part of the drawbacks presented by Java 8's Stream API:

- Extensibility of the API breaks the fluent API. Even though the extension of the API requires less code, it is not possible to chain together with an

already existing StreamEx instance, which breaks the fluent API provided by this library.

- The flatMap implementation breaks any short-circuiting operations.

## 2.4 JOOL and Cyclops

JOOL [5] and Cyclops [6] are equivalent to Stream. Both these libraries provide their own sequential types, Seq for JOOL and ReactiveSeq for Cyclops. They both expand the catalogue of operations when comparing with Stream, but when it comes to extending either API the programmer still needs to implement a Split-erator just like the one presented in listing 1.5.

## 2.5 Vavr

VAVR [4] provides its own sequential type, Seq, and like Guava, to extend its API the programmer needs to implement an Iterator which makes impossible for fluent API extensions. This library provides the widest set of functional data structures and operations from the libraries we studied and is one of most popular libraries on Github, second only to Guava. It provides an implementation of flatMap that does not break a short-circuiting operation on an infinite sequence.



# 3

## Solution

In this chapter we will present the Jayield solution, going through it's different components and understanding how they work and interact with each other.

### 3.1 Traverser

Traverser is an interface that is used to define ways of bulk traversing a sequence. To that end it provides the method `traverse` which receives a `Yield` function as argument to consume elements as they are traversed. Listing 3.1 shows the definition of the `Traverser` interface.

---

```
1 public interface Traverser<T> extends Serializable {  
2     void traverse(Yield<? super T> yield);  
3 }
```

---

Listing 3.1: Traverser

Looking at some examples on how to use this interface, if we have an upstream sequence of `Fruits`, listing 3.2 shows of how to transform such a sequence into a sequence of `Fruit`'s names.

---

```
1 public Traverser<String> mapToName(Traverser<Fruit> upstream) {  
2     return yield -> upstream.traverse(fruit -> yield.ret(fruit.name));  
3 }
```

---

Listing 3.2: Traverser map to name example

We can also define new sequences with the Traverser interface, lets say we want to generate a sequence of Fruits, listing 3.3 shows an example of how to do it.

---

```
1 public Traverser<Fruit> getFruitSequence() {
2     return yield -> {
3         yield.ret(Basket.orange);
4         yield.ret(Basket.blueberry);
5         ...
6         yield.ret(Basket.lemon);
7         yield.ret(Basket.mango);
8         yield.ret(Basket.strawberry);
9     };
10 }
```

---

Listing 3.3: Traverser as sequence generator

## 3.2 Yield

Yield is the interface that represents the element generator in our solution. It's interface is equivalent to a Java Consumer, however, it has a few key differences. Listing 3.4 presents the definition of the Yield interface.

---

```
1 public interface Yield<T> {
2     static final TraversableFinishError finishTraversal = new
        TraversableFinishError();
3
4     static public void bye() { throw finishTraversal; }
5
6     public void ret(T item);
7 }
```

---

Listing 3.4: Yield

As we can see Yield has two methods:

- `bye()`, which is used to short circuit the traversal of a sequence.
- `ret(T item)`, which is the method used to define what to do with each item that a sequence returns when traversing it.

Lets say we were traversing a sequence of Fruit names and we would like to write down their names, listing 3.5 shows what the Yield lambda to achieve this looks like.

---

```
1 name -> System.out.println(name)
```

---

Listing 3.5: Yield example

### 3.3 Query

Looking now at the Query class in Jayield, it represents a sequence of elements supporting sequential operations. This class implements all the utility methods with the same name, semantics and behaviour as those provided by Java's Stream. This class has an Traverser field that represents the current sequence. It also provides a way for the user to define new operations through the then method (listing 3.6).

---

```
1 <R> Query<R> then(Function<Query<T>, Traverser<R>> next)
```

---

Listing 3.6: Query then method

This method receives a Function as parameter that defines the way to transform the upstream sequence (represented by a Query) into the new traversal function, a Traverser. This function is similar to the method presented in the previous section in listing 3.2, with that in listing 3.2 we receive a Traverser as a parameter and the Function parameter in next receives a Query. An example of how to use this method can be looked at back in Chapter 1's listing 1.6.

### 3.4 Advancer

The Advancer is the interface that represents our individual traverser and the main focus of the use case of we are implementing.

The implementation for the Advancer (i.e. the individual traverser) will only be generated if requested to the Advancer interface as it is not always necessary to traverse. To do so we can call one of the two from methods provided in Advancer as shown in the 3.7 listing.

---

```
1 public interface Advancer<T> extends Serializable {
2
3     static <U> Iterator<U> iterator(Query<U> source) {...}
4
5     static <U> Advancer<U> from(Query<U> source) {...}
```

---

---

```

6
7  static <U> Advancer<U> from(Traverser<U> source) {...}
8
9  static <U> Iterator<U> iterator(Traverser<U> source) {...}
10
11 default Iterator<T> iterator() {...}
12
13 default <U> Advancer<U> then(Function<Advancer<T>,Advancer<U>> next)
    {...}
14
15 boolean tryAdvance(Yield<T> yield);
16
17 }

```

---

Listing 3.7: Advancer

This interface provides a way for the programmer to individually traverse a sequence, either through the tryAdvance method or through the use of an iterator. For instance, if we wanted to print the name of the first Fruit in a sequence we could take the approach shown in listing 3.8 which will print the first fruit's name, and only the first, if it exists.

---

```

1 void printFirstName(Query<Fruit> fruits) {
2     Advancer.from(fruits)
3         .tryAdvance(fruit -> System.out.println(fruit.name));
4 }

```

---

Listing 3.8: Advancer example

So, the question now is, how do we generate an Advancer from a Query or from a Traverser?

### 3.4.1 Approach

To achieve the expected result a pseudo-solution was thought of, assuming one of the following is true:

- A call to source.traverse(...) is made
- A call to yield.ret(...) is made

Assuming any of these are true, we would generate an Advancer doing the following:



- Instead of calling `source.traverse(...)` we would need to call `source.tryAdvance(...)`
- When calling `yield.ret(...)` we would register that an element has been found that meets the criteria.
- The call to `source.tryAdvance(...)` is made like so: `while(noElementFound && source.tryAdvance(...))`

So, lets say we have the following definition for a map Traverser:

---

```

1 <T, R> Traverser<R> map(Query<T> source, Function<T, R> mapper) {
2   return yield -> {
3     source.traverse(e -> yield.ret(mapper.apply(e)));
4   };
5 }
```

---

Listing 3.9: Map Traverser

Given the bulk traversal definition of map in listing 3.9, then the resulting map which implements the individual traversal is shown in listing 3.10

---

```

1 <T, R> Advancer<R> map(Query<T> source, Function<T, R> mapper) {
2   boolean[] noElementFound = new boolean[] {true};
3   return yield -> {
4     noElementFound[0] = true;
5     while(noElementFound[0] && source.tryAdvance(e -> {
6       yield.ret(mapper.apply(e));
7       noElementFound[0] = false;
8     }));
9     return noElementFound[0] == false;
10  };
11 }
```

---

Listing 3.10: Map Advancer

With that said, there are corner cases that are not so easily mapped as is the case of the `flatMap` operation which traverses over multiple sequences, these cases need a different approach.

Nevertheless, following the suggested approach we will need to generate code at runtime, to do this we will use ASM, a library for java bytecode manipulation at runtime. In order to read the bytecode via ASM we first need to give it (via the `ClassReader`) the actual bytecode of the Traverser. The problem is, if the Traverser was generated using a lambda function, that the lambda's class is generated at runtime via the `LambdaMetaFactory` class and therefore the normal method of

getting the bytecode does not work, aside from that we also needed to get any fields the Traverser's class may have, in the case of the lambda, they would be captured arguments.

### 3.4.2 SerializedLambda

You may have noticed in listing 3.1 that the Traverser interface extends from Serializable, the fact is, we actually need the Traverser to extend Serializable in order to read its code. This solution was found in an answer to a question in Stackoverflow[2], which makes use of the SerializedLambda[7].

The reason for this is that, if the lambda we are trying to read implements an interface that extends Serializable, the LambdaMetaFactory class will generate a private method called "writeReplace" which provides an instance of the class SerializedLambda. This instance can be used to retrieve both the actual method that was generated for this lambda as well as get the captured arguments and being able to reach the lambda's method allow us to instrument it.

### 3.4.3 Visitors

Now having access to the lambda's code, it was time to instrument it in order to transform the Traverser into an Advancer. To do so, we started by using a custom ClassVisitor that copies the original declaring class of the lambda as well as making some changes along the way, such as:

- The name of the class would now be the name of the lambda's method.
- Each call to a method on the declared class would now be a call to the generated class, due to calls to private methods.
- The return type of the lambda's method would now be boolean instead of void, to be coherent with the Advancer Interface.
- The lambda's method would now call tryAdvance instead of traverse and return the result of tryAdvance as well.

To obtain an Advancer out of the generated class all we would need to do now would be:

- Retrieve the original lambda's captured arguments

- Retrieve the generated method from the new class
- Apply the approach defined above.

Which would result in something like this:

---

```

1 final BoolBox elementFound = new BoolBox();
2 Method generatedTryAdvance = getTryAdvance(...);
3 Advancer generated = y -> {
4     elementFound.reset();
5     Yield<R> wrapper = wr -> {
6         y.ret(wr);
7         elementFound.set();
8     };
9     arguments[arguments.length - 1] = wrapper;
10    while(elementFound.isFalse() && (Boolean) generatedTryAdvance.invoke(
        newClass, arguments));
11    return elementFound.isTrue();
12 };

```

---

Listing 3.11: Generated Advancer

This solution makes use of a wrapper over the yield, received as parameter, to register when an element is found, this wrapper is then provided as the last argument in an argument array for the generated method, with the first parameters being those captured by the lambda.

### 3.4.4 State Machine

This approach however, as said before, has flaws and the above described solution does not work for every case. One of the first issues identified was the case functions that would generate more elements than the original sequence, that being either, a duplicate function or a generator of values. Lets look at listing 3.12 as an example of an traverser that duplicates values.

---

```

1 public <T> Traverser<T> dup(Traverser<T> upstream) {
2     return yield -> upstream.traverse(item -> {
3         yield.ret(item);
4         yield.ret(item);
5     });
6 }

```

---

Listing 3.12: Duplicate Traverser function

For this Traverser, the solution described above would not work as it would not stop when the first element was found nor would it return to the same point when the next element was requested. To solve this issue two things were necessary, a state machine and external iteration.

The first approach to this issue was based on declaring the end of a state, and the beginning of the next, every time a `yield.ret(...)` was made. This would translate to something like what is shown in listing 3.13.

---

```

1 yield -> {
2   this.hasElement.reset();
3   YieldWrapper wrapper = new YieldWrapper(yield, this.hasElement);
4   Query source = this.source;
5   Yield instrumentedYield = (item) -> {
6     if(this.state == 0){
7       this.state++;
8       wrapper.ret(item);
9       return;
10    } else if (this.state == 1) {
11      this.state++;
12      wrapper.ret(item);
13      return;
14    } else if (this.state == 2) {
15      this.state = 0;
16      this.validValue = false;
17      return;
18    }
19  };
20
21  while(this.hasElement.isFalse() && (this.iterator.hasNext() || this.
    validValue)) {
22    if (!this.validValue && iterator.hasNext()) {
23      this.current = iterator.next();
24      this.validValue = true;
25    }
26    instrumentedYield.ret(this.current);
27  }
28
29  return this.hasElement.isTrue() ? true : this.hasElement.isTrue();
30 }

```

---

Listing 3.13: Duplicate Advancer translation with if based State Machine

As you can see there are a few key differences:

- the instrumented yield operation now isolates `yield.ret()` calls into different

states

- the `validValue` field, which lets us know if we need to get a new value or if the old one still needs to be used in another state.

This solution had its problems as well, for instance, if a `yield` return was made inside an `if` condition it would have to take that into account which would lead to even more complicated translations.

At this moment we took a step back to re-evaluate and it was then that we started to look into Jon Skeet's *C# in Depth*[9], chapter 6. The approach described here, makes use of a `switch` case as well as labels, with the labels being the delimiters between the end of a state and the beginning of a new one. This approach requires some pre-processing of the Traverser code, to know how many states there will be in the `switch` as in Java to declare a `switch` we need to know in advance how many states there will be, this approach was also applied to value generating lambdas, as the same principles apply. Listing 3.14 shows an example of the translation of the `dup` operation.

---

```

1 yield -> {
2     this.hasElement.reset();
3     YieldWrapper wrapper = new YieldWrapper(yield, this.hasElement);
4     Query source = this.source;
5     Yield instrumentedYield = (item) -> {
6         switch(this.state) {
7             case 0:
8                 wrapper.ret(item);
9                 this.state = 1;
10                return;
11            case 1:
12                wrapper.ret(item);
13                this.state = 2;
14                return;
15            case 2:
16                this.state = 0;
17                this.validValue = false;
18                return;
19            default:
20        }
21    };
22
23    while(this.hasElement.isFalse() && (this.iterator.hasNext() || this.
24        validValue)) {
25        if (!this.validValue && iterator.hasNext()) {

```

---

### 3. SOLUTION

---

```
25         this.current = iterator.next();
26         this.validValue = true;
27     }
28     instrumentedYield.ret(this.current);
29 }
30
31 return this.hasElement.isTrue() ? true : this.hasElement.isTrue();
32 }
```

---

Listing 3.14: Duplicate Advancer translation with switch based State Machine

# Bibliography

- [1] amaembo. Streamex. URL <https://github.com/amaembo/streamex>. (p. 10)
- [2] bennyl. How to read lambda expression bytecode using asm, March 2017. URL <https://stackoverflow.com/a/35223119>. (p. 18)
- [3] Kevin Bourrillion and Google Jared Levy. Guava, 2009. URL <https://github.com/google/guava>. (p. 8)
- [4] Daniel Dietrich. Cyclops, 2014. URL <https://github.com/vavr-io/vavr>. (p. 12)
- [5] Lukas Eder. jool, 2014. URL <https://github.com/jOOQ/jOOL>. (p. 12)
- [6] John McClean. Cyclops, 2014. URL <https://github.com/aol/cyclops-react/>. (p. 12)
- [7] Oracle. Serialized lambda, . URL <https://docs.oracle.com/javase/8/docs/api/java/lang/invoke/SerializedLambda.html>. (p. 18)
- [8] Oracle. Spliterator, . URL <https://docs.oracle.com/javase/8/docs/api/java/util/Spliterator.html#forEachRemaining-java.util.function.Consumer->. (p. 5)
- [9] Jon Skeet. C# in depth, 2013. URL <http://csharpindepth.com/Articles/Chapter6/IteratorBlockImplementation.aspx>. (p. 21)