

Introduction to Scala

Kenny Zhuo Ming Lu

School of Information Technology, Nanyang Polytechnic

June 22, 2016

Learning outcomes

By the end of this lecture, you will have some basic understanding on

- Object Oriented Programming in Scala
- Functional Programming in Scala
 - Algebraic data type and pattern matching
 - Higher order functions
 - Partial functions
 - Implicit functions
 - Type classes
 - Monads
- Data parallelism in Scala
- Concurrent programming in Scala with threads and actors

Scala is ...

A general purpose programming language that is

- Object Oriented and Functional
- with distributed programming builtin
- with a rich type system and a huge set of libraries and eco-systems
- Widely used in academic research projects (e.g. Spark was from UC Berkeley) and industries (e.g. Scalding from Twitter, Kakfa from Linkedin, Coursera Walmart, ...)

Hello World in a script

Let's say we have a Scala file Script.scala

```
println("Hello world!")
```

To execute it

\$ scala Script.scala Hello world!

Hello World in a console application

Compiling Scala program into an application

```
object Main extends App
{
    println("Hello world!")
}
```

To execute it

\$ scalac Main.scala
\$ scala Main
Hello world!

object is like a singleton class.

It is very similar to Java

```
class Person(n:String,i:String) {
  private val name:String = n
  private val id:String = i
  def getName():String = name
  def getId():String = id
}
```

- Person is the name of the class, Person(n:String,i:String) is a constructor.
- private sets the accessibility scope of the members
- val introduces a value, (immutable variable)
- def introduces a method defintion.

Class Inheritence via the extends keyword

- var introduces a mutable variable
- Most of the type annotations are optional and will be inferred by the type system.

Traits are like Java interfaces and abstract classes

```
trait NightOwl {
   def stayUpLate():Unit
}
class Student(n:String, i:String, g:Double)
      extends Person(n,i) with NightOwl {
   private var gpa = g
   def getGPA() = gpa
   def setGPA(g:Double) = {
      gpa = g
   }
   override def stayUpLate():Unit = {
      println("woohoo")
```

Running our first Scala program

Scala comes with an interpretor (AKA REPL)

```
scala> :load OOP.scala
Loading OOP.scala...
defined class Person
defined trait NightOwl
defined class Student
defined class Staff
scala> val tom = new Student("Tom", "X1235", 4.0)
tom: Student = Student@601c1dfc
scala> val jerry = new Staff("Jerry", "T0001", 500000.0)
jerry: Staff = Staff@650fbe32
scala> tom.stayUpLate
woohoo
```

Functional Programming in Scala

OOP is useful, but often less interesting, let's consider the FP features in Scala.



Image taken from https://www.toptal.com/scala/why-should-i-learn-scala

Algebraic Data Type

In Scala, this is how we define a data type

```
sealed trait Exp
case class Val(v:Int) extends Exp
case class Plus(e1:Exp, e2:Exp) extends Exp

def simp(e:Exp):Exp = e match {
   case Val(v) => e
   case Plus(Val(0), e2) => e2
   case Plus(e1,e2) => Plus(simp(e1), simp(e2))
}
```

- sealed requires all extensions to the trait Exp must be declared in the same module.
- case makes a class "pattern-matchable"

Algebraic Data Type

In Java, to achieve the above we have to use abstract class and method overriding. (For brevity, setter and getter methods are omitted.)

```
public abstract class Exp { public abstract Exp simp(); }
public class Val extends Exp {
 private int v;
 public Val(int v) { this.v = v; }
 public Exp simp() { return this; }
public class Plus extends Exp {
 private Exp e1; private Exp e2;
 public Plus(Exp e1, Exp e2) {
   this.e1 = e1: this.e2 = e2:
 public Exp simp() {
    if (e1 instanceof Val) {
     Val v1 = (Val)e1;
      if (v1.getV() == 0) { return e2; }
   return new Plus(e1.simp(),e2.simp());
```

List

Algebraic data type and pattern matching are useful in case of handling recursive/nested data structures, e.g. a linked list.

```
sealed trait List[A]
case class Nil[A]() extends List[A]
case class Cons[A](x:A, xs:List[A]) extends List[A]
val x = Cons(1, Cons(2, Nil()))
def length[A](1:List[A]):Int = 1 match {
   case Nil() => 0
   case Cons(\_,xs) \Rightarrow 1 + length(xs)
```

Observation

- A is a type variable (AKA generics)
- Parametric polymorphism
- Nil() takes no argument, it's always (almost) singleton.

Mixing subtying with parametric polymorphism

```
sealed trait List[+A]
case object Nil extends List[Nothing]
case class Cons[A](x:A, xs:List[A]) extends List[A]

val x = Cons(1,Cons(2,Nil))

def length[A](l:List[A]):Int = 1 match {
   case Nil => 0
   case Cons(_,xs) => 1 + length(xs)
}
```

- case object introduces a singleton value constructor.
- + indicates that A is a covariant type parameter, hence List[Nothing] <: List[A] for any A.
- Nothing is the bottom type in the sub-typing hierachy.

List is a builtin data type comes with Scala library

```
val x = List(1,2)

def length[A](1:List[A]):Int = 1 match {
   case Nil => 0
   case (_::xs) => 1 + length(xs)
}
```

which is equivalent to the following Haskell program

```
x = [1,2]
length :: [a] -> Int
length l = case l of
   [] -> 0
   (_:xs) -> 1 + length xs
```

List Comprehension

- for ... yield denotes a sequence comprehension.
- C.f. set builder notation in math,

$$\{\mathit{nat} + 1 | \mathit{nat} \in \mathit{nats} \land \mathit{nat} \bmod 2 = 0\}$$

For loop vs comprehension

Note that for ... yield denotes a sequence comprehension.

While for introduces a for loop.

```
for (i <- 1 to 10) { println(i) }
```

prints the numbers 1 to 10.

```
for (nat <- nats if nat % 2 == 0) { println(nat) }</pre>
```

prints the numbers 2 and 4.

List operations

Length

```
val 11 = List(1,2,3)
val len = 11.length // 3
```

Concatenation

```
val 12 = List(-1,-2,-3)
val 13 = 11 ++ 12 // List(1,2,3,-1,-2,3)
```

Reverse

```
val 14 = 11.reverse // List(3,2,1)
```

List operations

val max = 11.max // 3

min/max // recall 11 = List(1,2,3) val min = 11.min // 1

```
sorting
// recall 13 = List(1,2,3,-1,-2,3)
val s13 = 13.sortWith(_<_) // List(-3, -2, -1, 1, 2, 3)
// we will look into _<_ in a short while.</pre>
```

sub-list

```
val sl3a = sl3.take(3) // List(-3,-2,1)
val sl3b = sl3.drop(3) // List(1,2,3)
```

Map and Fold

1.map(f) returns a new list by applying the unary function f to
each element in 1

```
val 1 = List(1,2,3,4)
val 12 = 1.map(x => x * 2) // List(2,4,6,8)
```

Note that $x \Rightarrow x * 2$ is a lambda-expression, which can be shortened as $_{-} * 2$. The . is optional.

which is equivalent to the following Haskell program

$$1 = [1,2,3,4]$$

$$12 = map (\x -> x*2) 1$$

Map and Fold

1.foldLeft(b)(f) agreggate all the elements in 1 with the binary function f and the base value b.

```
// recall 1 = List(1,2,3,4)
val 13 = 1.foldLeft(0)((x,y) => x + y)
```

In essence the above is computing

$$(((0 + 1) + 2) + 3) + 4$$

1.foldRight(b)(f) agreggate all the elements in 1 with the binary function f and the base value b with right associativity.

In essence the above is computing

$$0 + (1 + (2 + (3 + 4)))$$

Higher order function

The defintion of map.

```
sealed trait List[+A] {
  def map[B](f:A=>B):List[B] = this match {
    case Nil => Nil
    case x::xs => f(x)::xs.map(f)
  }
}
case object Nil extends List[Nothing]
case class Cons[A](x:A, xs:List[A]) extends List[A]
```

Function f:A=>B is used as a formal argument. Equivalently, in Haskell, we have.

```
map :: (a -> b) -> [a] -> [b]
map f [] = []
map f (x:xs) = (f x):(map f xs)
```

Higher order function

Function composition

```
val 11 = List(1,2,3)
val 12 = 11.map (((x:Int) => x*2) compose (y=>y+1))
// List(4, 6, 8)

(f compose g)(x) is equivalent to f(g(x))
val 13 = 11.map (((x:Int) => x*2) andThen (y=>y+1))
// List(3, 5, 7)

(f andThen g)(x) is equivalent to g(f(x))
```

Higher order function

In Scala, functions are instances/objects of trait Function

The defintions of compose and andThen. object Function { // this is like a package trait Function1[+A][-B] { def apply(x:A):B {...} // apply() gives us // the syntactic sugar form like f(x) def compose[C](that:C=>A):C=>B = (c:C) => this(that(c)) def andThen[C](that:B=>C):A=>C = (a:A) => that(this(a))

Both compose and andThen return functions as results.

Functions vs methods

In Scala there is a clear distinction between functions and methods.

Methods are introduced by def keyword, e.g.

```
scala> def f(x:Int) = x
f: (x: Int)Int
```

 Functions are implementations of the FunctionX traits. e.g. lambda expressions are functions.

```
scala> val g = (x:Int) => x
g: Int => Int = <function1>
```

Similarity: both can be used as argument and results

```
scala> List(1,2).map (f) == List(1,2).map (g)
res2: Boolean = true
```

Functions vs methods

In Scala there is a clear distinction between functions and methods.

To convert a method to a function.

```
scala> f _
res1: Int => Int = <function1>
```

Before we take a break...

Let's play some quiz games. Open the web browser in your mobile, tablet or laptop.

- visit http://kahoot.it
- key in the game pin, then
- key in your name, please use real name to claim the prize.





Before we take a break...



Partial functions

There are situations in which we need to deal with partial functions, naturally, we can use the Option type in Scala (similar to the Maybe type in Haskell).

```
def div(x:Int)(y:Int):Option[Int] = {
    if (y == 0) None else Some(x/y)
}
val xs = List(0,1,2) map (x \Rightarrow div(2)(x))
yields List(None, Some(2), Some(1))
val ys = xs filter (x => x.nonEmpty)
yields List(Some(2), Some(1))
val zs = ys map ( x \Rightarrow x match { case Some(y) \Rightarrow y } )
yields List(2,1)
```

Partial functions

Alternatively, we may use the Scala bultin PartialFunction

```
def div(x:Int): PartialFunction[Int,Int] = {
   case (y:Int) if y != 0 => x/y
}
val xs = List(0,1,2) collect ( div(2) )
```

yields List(2, 1).

Note that we have to use the collect method instead of map

```
val ys = List(0,1,2) map ( div(2) )
```

Will result in a scala.MatchError

Partial functions

orElse allows us to compose partial functions which are complimenting on their domains.

```
def div(x:Int): PartialFunction[Int,Int] = {
    case (y:Int) if y != 0 => x/y
}

val mkZero: PartialFunction[Int,Int] = {
  case (y:Int) if y == 0 => 0
}

val ys = List(0,1,2) map ( div(2) orElse mkZero )
```

yields in a List(0, 2, 1)

Implicit functions

Implicit function and implicit parameter allow us to define default behaviors.

```
implicit def logger(mesg:String):Unit = {
    println("[Log] " + mesg)
}
def fib(x:Int)(implicit log:String=>Unit):Int = {
  x match {
    case 0 \Rightarrow 0
    case 1 \Rightarrow 1
    case _ => {
      log("computing fib(" + x +")")
      fib(x-1) + fib(x-2)
```

Implicit functions

```
scala> fib(5)
[Log] computing fib(5)
[Log] computing fib(4)
[Log] computing fib(3)
[Log] computing fib(2)
[Log] computing fib(2)
[Log] computing fib(3)
[Log] computing fib(2)
res7: Int = 5
scala> fib(5)( x => Unit)
res8: Int = 5
```

Implicit functions

Scala resolves implicit parameters as follows,

- Search in current scope
 - Implicits defined in current scope
 - Implicits defined in Explicit imports
 - Implicits defined in wildcard imports
- Search for associated types in
 - Companion objects of a type
 - Implicit scope of an argument's type
 - Outer objects for nested types
 - Other dimensions

Type classes

Type classes was introduced in Haskell to allow mixing parametric polymorphism and ad-hoc polymorphism.

Haskell Type Class Example

```
class Eq a where
  (==) :: a -> a -> Bool
instance Eq Int where
  (==) x y = eqInt x y
instance Eq a => Eq [a] where
  (==) [] = True
  (==) (x:xs) (y:ys) = (x == y) && (xs == ys)
  (==) _ = False
check :: (Eq a) => [a] -> [a] -> Bool
check xs ys = xs == ys
```

In Scala, type classes are "encoded" using traits and implicit definitions, i.e. implicit val, implicit def and implicit object

```
trait Eq[A] { def eq(x:A,y:A):Boolean }
implicit val eqInt = new Eq[Int] {
    def eq(x:Int,y:Int):Boolean = x == y
}
implicit def eqList[A](implicit ev:Eq[A]) =
 new Eq[List[A]] {
    def eq(xs:List[A],ys:List[A]):Boolean =
      (xs,ys) match {
        case (Nil, Nil) => true
        case (x::xs2, y::ys2) \Rightarrow ev.eq(x,y) && eq(xs2,ys2)
        case (\_, \_) \Rightarrow false
     }
```

By resolving implicit parameter the compiler builds the "evidence" ev AKA the dictionary.

```
def check[A](x:List[A], y:List[A])
  (implicit ev:Eq[List[A]]) = ev.eq(x,y)
```

```
scala> check(List(1,2,3), List(1,2,3))
res6: Boolean = true
```

The above builds the evidence ev = eqList(eqInt)

To mimick type classes' effect in OOP languages like Java, we would have to use interface and function overriding with lots of cares. (We omit the getter and setter methods for brevity.)

```
public interface Eq<T> { public boolean eq(T that); }
public class Int implements Eq<Int> {
 private int val;
 public Int(int x) { val = x; }
 public boolean eq(Int that) {
   return (val == that.getVal());
public class List<T extends Eq<T>> implements Eq<List<T>> {
 private Node<T> pHead;
 public List() { pHead = null;}
 public void insert(T v) {
    Node<T> n = new Node<T>(v):
   n.setNext(pHead);
   pHead = n;
 public boolean eq(List<T> that) {
   return pHead.eq(that.getPHead());
```

To mimick type classes' effect in OOP languages like Java, we would have to use interface and function overriding with lots of cares. (We omit the getter and setter methods for brevity.)

```
public class Node<T extends Eq<T>> implements Eq<Node<T>> {
 private T val:
 private Node<T> next;
 public Node(T v) {
   val = v;
   next = null;
 public boolean eq(Node<T> that) {
    if (!val.eq(that.getVal())) { return false; }
    else {
      if ((next == null) && (that.getNext() == null)) { return true; }
      else if ((next == null) && (that.getNext() != null)) { return false; }
      else if ((next != null) && (that.getNext() == null)) { return false; }
      else { return next.eq(that.getNext()); }
```

To mimick type classes' effect in OOP languages like Java, we would have to use interface and function overriding with lots of cares.

Finally we can implement the check function in Java as follows,

```
public static <T extends Eq<T>> boolean check(List<T> 11, List<T> 12) {
    return 11.eq(12);
}
```

For real world applications implemented in Scala, we do not define all the type classes from the scratch. We use a library project called scalaz. https://github.com/scalaz/scalaz which provides many useful type classes.

A vague analogy

- Algebraic data types define abstractions of run-time data (or values), i.e. what are the possible values.
- Monads define abstractions of computations, i.e. what the possible computations.

Let's model safe arithmetic computation

```
sealed trait SafeVal[+T]
case class Val[T](x:T) extends SafeVal[T]
case class Err(msg:String) extends SafeVal[Nothing]

type SafeInt = SafeVal[Int]
```

```
def add(x:SafeInt, y:SafeInt):SafeInt = x match {
  case Err(msg) => Err(msg)
  case Val(i) => y match {
    case Err(msg) => Err(msg)
    case Val(j) => Val(i + j)
def div(x:SafeInt, y:SafeInt):SafeInt = x match {
  case Err(msg) => Err(msg)
  case Val(i) => y match {
    case Err(msg) => Err(msg)
    case Val(j) => {
        if (j == 0) {Err("div by 0")} else {Val (i/j)}
```

We spot many similar code structure, let's rewrite them in terms of some higher order combinators pnt and bnd.

```
def pnt[A](x: => A):SafeVal[A] = Val(x)
def bnd[A,B](sa: SafeVal[A])
  (f: A => SafeVal[B]): SafeVal[B] = sa match {
  case Err(msg) => Err(msg)
  case Val(a) => f(a)
}
```

- Note that x:=> A defines a lazy formal parameter x of type A.
- pnt "lifts" / "injects" a value into the SafeVal object.
- bnd "composes" the previous SafeVal with the subsequent computation.

```
def add(x:SafeInt, y:SafeInt):SafeInt = {
  bnd(x)(x1 \Rightarrow \{
    bnd(y)(y1 \Rightarrow pnt(x1+y1))
  })
def div(x:SafeInt, y:SafeInt):SafeInt = {
  bnd(x)(x1 \Rightarrow \{
    bnd(y)(y1 \Rightarrow {
       if (y1 == 0) { Err("div by 0") }
       else { pnt(x1 / y1) }
    })
```

If we use type classes Monad and MonadPlus provided by scalaz, we can make it even more concise.

```
implicit object safeValMonadPlus
  extends MonadPlus[SafeVal] {
  override def point[A](a: => A):SafeVal[A] = Val(a)
  override def bind[A, B](sa: SafeVal[A])
  (f: A => SafeVal[B]): SafeVal[B] = sa match {
    case Err(msg) => Err(msg)
    case Val(a) \Rightarrow f(a)
```

```
def add(x:SafeInt,y:SafeInt):SafeInt = {
  for (x1 < -x
  ; y1 <- y
  ) yield (x1 + y1)
def div(x:SafeInt,y:SafeInt):SafeInt = {
  for ( x1 < -x
  ; y1 <- y
  ; if y1 != 0
  ) yield (x1 / y1)
```

```
Roughly speaking, for (x \leftarrow e1; e2) is desugared to bind(e1)(x => e2)
```

There many types of commonly used monads

- List monad list traversal and aggregate computation.
- State monad computation with explicit state passing
- Reader monad input scanning computation
- Monadic parser combinator LL-style top-down parsing computation

Monads are composable via monad trasformer.

So you still want to see how to mimick monads in Java?

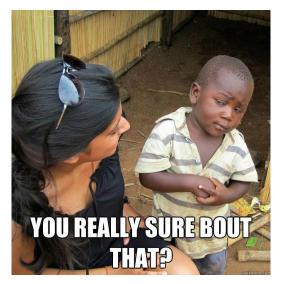


Image taken from http://quickmeme.com

Kenny Zhuo Ming Lu

Ok, if you are really interested, look here. http://github.com/luzhuomi/parsecj/tree/master/ Parsec/src/parsec

Data Parallelism

Scala has built-in support of data parallelism. For instance,

```
object Fib extends App {
  def fib(x:Int):Int = x match {
    case 0 => 0
    case 1 => 1
    case n => fib(n-1) + fib(n-2)
  }
  println((30 to 40).toList.map(fib))
}
```

Data Parallelism

.par converts a sequentials data collection into a parallel counter
part. .seq has the opposite effect.

```
object ParFib extends App {
  def fib(x:Int):Int = x match {
    case 0 => 0
    case 1 => 1
    case n => fib(n-1) + fib(n-2)
  }
  println((30 to 40).toList.par.map(fib))
}
```

Concurrent Programming

- Multi-threading
- Actors
- Async computation via Future

Concurrent Programming

Multi-threading in Scala is similar to Java

```
object ThreadFib extends App {
  def fib(x:Int):Int = x match {
    case 0 \Rightarrow 0
    case 1 \Rightarrow 1
    case n \Rightarrow fib(n-1) + fib(n-2)
  }
  for (i <- 30 to 40) {
    new Thread(new Runnable {
      def run() {
         println(fib(i))
    }).start
```

Actor model is a concurrent programming model.

- Actors are primitive objects with specific roles and functionalities
- Actors (objects) are coordinating with each other via message passing

Let's consider a ping pong simulation using Scala actors

- There are two actors Ping and Pong.
- Ping keeps an internal counter
 - when it receives a start message, it increments the counter and send a ping message to Pong
 - when it receives a pong message, it increments the counter.
 - if the counter > 99, it sends a stop message to pong then terminates.
 - otherwise, sends a ping message to Pong.
- Pong
 - when it receives a stop message, it terminates.
 - when it receives a ping message, it sends a pong message to Ping.

```
case object PingMessage
case object PongMessage
case object StartMessage
case object StopMessage
```

```
class Ping(pong: ActorRef) extends Actor {
 var count = 0
 def incrementAndPrint = {count += 1; println("ping")
 def receive = {
    case StartMessage =>
      incrementAndPrint
     pong! PingMessage
    case PongMessage =>
      incrementAndPrint
      if (count > 99) {
        sender! StopMessage
        println("ping stopped")
        context.stop(self)
      else { sender ! PingMessage }
```

```
class Pong extends Actor {
 def receive = {
    case PingMessage =>
     println(" pong")
     sender! PongMessage
    case StopMessage =>
     println("pong stopped")
     context.stop(self)
```

```
object PingPongTest extends App {
  val system = ActorSystem("PingPongSystem")
  val pong = system.actorOf(Props[Pong],
  name = "pong")
  val ping = system.actorOf(Props(new Ping(pong)),
  name = "ping")
  ping ! StartMessage
}
```

The ping pong example is pretty contrived, let's consider a less contrived example in which we try to compute fib(n) in parallel. The idea is as follows,

- if n is less than 10, we compute fib(n) locally
- if n is greater than or equal to 10, we create two fib actors.
 The first actor computes fib(n-1) and the second actor computes fib(n-2). The two actors can run in parallel if there is enough threads. When both actors revert with some results, we sum up the results.

Let's define the messages for communication as follows,

```
case class FibRequest(n:Int)
case class FibResult(n:Int)
case object Stop
```

```
class FibActor extends Actor {
 def receive = {
    case Stop => context.stop(self)
    case FibRequest(n) if n < 10 => sender ! FibResult(fib(n))
    case FibRequest(n) if n \ge 10 = 10
      val child1 = context.actorOf(Props[FibActor])
      val child2 = context.actorOf(Props[FibActor])
      val future1:Future[Any] = child1 ? FibRequest(n-1) // ''ask'' op
      val future2:Future[Any] = child2 ? FibRequest(n-2)
      val dest = sender
      val future1and2 = for ( // Future[T] is a monad
       FibResult(v1) <- future1;
        FibResult(v2) <- future2
      ) yield v1+v2
      future1and2 onComplete {
        case Success(v) =>
        child1 ! Stop; child2 ! Stop
        dest ! FibResult(v) // can't use sender
        case Failure(e) => println("failed!" + e.toString)
```

```
def main(args:Array[String]):Unit = {
  val system = ActorSystem("FibSystem")
  val fibActor = system.actorOf(Props[FibActor])
 val n = if (args.length > 0) args(0).toInt else 10
  val f = fibActor ? FibRequest(n)
  f onComplete {
    case Success(v) => { println(v) ; fibActor ! Stop}
    case Failure(e) => println("failed!" + e.toString)
```

In Summary

We've covered

- OOP in Scala
- Functional Programming in Scala which includes,
 - Algebraic data type and pattern matching
 - Higher order functions
 - Partial functions
 - Implicit functions
 - Type classes
 - Monads
- Parallel programming in Scala
- Concurrent programming in Scala with
 - threadings
 - actors

Latex source and example codes can be found here https://github.com/luzhuomi/learning-scala