Principles of Programming Languages
Course Notes

Elaine Li

January 22, 2026

Preface

This document contains the lecture notes for the NYU undergraduate course CSCI-UA.0480-055, “Principles of Programming Languages”, in Spring 2026. The document will be extended throughout the semester. So please stay tuned!

Course Summary

Computing professionals have to learn new programming languages all the time. This course teaches the fundamental principles of programming languages that enable you to learn new languages quickly and help you decide which one is best suited for a given task.

We will explore new ways of viewing computation and programs, and new ways of approaching algorithmic problems, making you better programmers overall. The topics covered in this course include:

- recursion and induction
- algebraic data types and pattern matching
- higher-order functions
- continuations and tail recursion
- programming language syntax and semantics
- type systems
- monads
- objects and classes

We will explore this material by building interpreters for programming languages of increasing complexity. The course will thus be accompanied by extensive programming assignments. We will use the programming language Scala for these assignments, which you will also learn in this course.

Contents

List of Figures

Chapter 1

Scala Basics

1.1 Getting Started

In the following, we assume that you have installed sbt and Intellij Idea with the Scala plugin. If you have not yet done so, please do it now. You can find a link to the installation instructions on the course web site on Brightscape.

1.1.1 Compiling and Running Scala Applications

Compiling and running Scala applications works similar to Java. For example, you can open a text editor and type in the following Scala code:

```
package greeter

object Hello extends App:
  def main(args: Array[String])
    println("Hello␣World!")
```

This code creates an object `Hello` in the package `greeter`. The object `Hello` contains a method called `main`, which means that `Hello` can serve as the entry point of a Scala application that calls `main` upon start, providing the command line arguments via the parameter `args`. When this application is started, the object `main` prints the message `"Hello␣World!"` on standard output.

The above code is roughly equivalent to the following Java code:

```
package greeter;

public class Hello {
  public static void main(String[] args) {
    System.out.println("Hello␣World!");
  }
}
```

You can safe the Scala code, say, in a file called `Hello.scala`, and then compile it with the Scala compiler. To do this, open a command prompt, go to the directory where you saved the file, and type `scalac Hello.scala`. This will create a file `Hello.class`, which contains the compiled byte code of the object `Hello`.

To execute the program, type `scala greeter.Hello` in your terminal. This will start the Scala runtime environment, which will execute the byte code in `Hello.class` using the Java virtual machine. You should see the message `"Hello␣World!"` printed in your terminal.

If you are using the Intelij Idea IDE, you can import the `in-class-code` project following the instructions provided in the `README.md` file of the repository. The repository contains a file `ScalaGreeter.scala`, which you can find in the package `popl.class03`. To compile and run the application, right-click on the file and select “Run 'ScalaGreeter' ”. This should print `"Hello Scala"` in the output view at the bottom of the window.

1.1.2 The Scala REPL and Worksheets in the IDE

If you want to experiment with the Scala language, it is quite cumbersome to write an extra application for every small code snippet that you would like to execute. To make life easier, Scala provides a useful tool called a read-eval-print loop, or *REPL* for short. The Scala REPL is essentially a command line calculator on steroids. It allows you to type arbitrary Scala code in a terminal. The code is then evaluated and the result of the evaluation is printed in the terminal.

To start the Scala REPL, open a terminal and execute `scala`. This will start the REPL and a message similar to the following should appear:

```
Welcome to Scala 3.3.0 (18.0.2-ea, Java OpenJDK 64-Bit Server VM).
Type in expressions for evaluation. Or try :help.

scala>
```

Now, you can type a Scala expression. For example typing 3 + 4 yields

```
scala> 3 + 4
val res0: Int = 7
```

You can exit the REPL by typing `:quit` or by pressing `Ctrl-d` (respectively, `Cmd-d` on OS X).

If you only installed sbt and Intelij Idea, you can start the REPL by typing `sbt console` in a terminal. Intelij Idea provides a feature similar to the REPL, called *Worksheets*. You can use this feature as follows:

1. Start Intelij Idea. (In the following, I assume you have previously imported the `in-class-code` project with the `popl` package.)
2. Right-click the `popl` package in the package explorer and choose New/Scala Worksheet. Name the worksheet “MyWorksheet” and press Finish.

3. An empty Scala source file called `MyWorksheet.sc` will open in the editor window.
4. You can now type Scala expressions in the editor window. Each time you click the "Evaluate Worksheet" button at the top of the editor view (the button with the green arrow icon), the expressions in the file are evaluated and the result of the evaluation appear in a separate view next to the editor.
5. If you click the 'Show worksheet settings' button (the button with the wrench icon), you can set the worksheet to Interactive Mode, which will cause all expressions to be evaluated automatically as you type.

1.2 Scala Crash Course

In the following, we assume that you have started the Scala REPL. Though, (almost) all of these steps can also be done in a worksheet.

1.2.1 Expressions, Values, and Types

After you type an expression in the REPL, such as `3 + 4`, and hit enter:

```
scala> 3 + 4
```

The interpreter will print:

```
val res0: Int = 7
```

This line includes:

- the keyword `val`, indicating that you have defined a new value resulting from evaluating the expression.
- an automatically generated name `res0` that is *bound* to that new value,
- a colon `:`, followed by the type `Int` of the expression,
- an equals sign `=`,
- the value `7` resulting from evaluating the expression.

The type `Int` names the class `Int` in the package `scala`. Packages in Scala partition the global name space and provide mechanisms for information hiding, similar to Java packages. Values of class `Int` correspond to values of Java's primitive type `int` (Scala makes no difference between primitive and object types). More generally, all of Java's primitive types have corresponding classes in the `scala` package.

We can reuse the automatically generated name `res0` to refer to the computed value in subsequent expressions (this only works in the REPL but not in a worksheet):

```
scala> res0 * res0
val res1: Int = 9
```

Java's ternary conditional operator `? :` has an equivalent in Scala, which looks as follows:

```
scala> if res1 > 10 then res0 - 5 else res0 + 5
val res2: Int = 2
```

In addition to the `? :` operator, Java also has if-then-else statements. Scala, on the other hand, is a functional language and makes no difference between expressions and statements: every programming construct is an expression that evaluates to some value. In particular, we can use if-then-else expressions where we would normally use if-then-else statements in Java.

```
scala> if res1 > 2 then println("Large!")
       else println("Not␣so␣large!")
Large!
```

Note that the result value is not automatically bound to a name in this case. The if-then-else expression still evaluates to the value `()`, which is of type `Unit`. This type indicates that the sole purpose of evaluating the expression is the side-effect of the evaluation (here, printing a message). In other words, in Scala, statements are expressions of type `Unit`. Thus, the type `Unit` is similar to the type void in Java (which however, has no values). The value `()` is the only value of type `Unit`.

1.2.2 Names

We can use the **`val`** keyword to give a user-defined name to a value, so that we can subsequently refer to it in other expressions:

```
scala> val x = 3
val x: Int = 3
scala> x * x
val res0: Int = 9
```

Note that Scala automatically infers that `x` has type `Int`. Sometimes, automated type inference fails, in which case you have to provide the type yourself. This can be done by annotating the declared name with its type:

```
scala> val x: Int = 3
val x: Int = 3
```

A **`val`** is similar to a `final` variable in Java. That is, you cannot reassign it another value:

```
-- [E052] Type Error: ----------------------------------------------------------
1 |x = 5
  |^^^^^
  |Reassignment to val x
```

Scala also has an equivalent to standard Java variables, which can be reassigned. These are declared with the **var** keyword

```
scala> var y = 5
var y: Int = 5
scala> y = 3
y: Int = 5
```

The type of a variable is the type inferred from its initialization expression. It is fixed throughout the lifetime of the variable. Attempting to reassign a value of incompatible type results in a type error:

```
-- [E007] Type Mismatch Error: -----------------------------------------------
1 |y = "Hello"
  |    ^^^^^^^
  |    Found:    ("Hello" : String)
  |    Required: Int
```

For the time being, we will pretend that variables do not exist. Repeat after me: **val**s are gooood! **var**s are baaaad!

1.2.3 Functions

Here is how you write functions in Scala:

```
scala> def max(x: Int, y: Int): Int =
         if x > y then x else y
def max(x: Int, y: Int): Int
```

Function definitions start with **def**, followed by the function's name, in this case `max`. After the name comes a comma separated list of parameters enclosed by parenthesis, here `x` and `y`. Note that the types of parameters must be provided explicitly since the Scala compiler does not infer parameter types. The type annotation after the parameter list gives the result type of the function. The result type is followed by the equality symbol, indicating that the function returns a value, and the body of the function which computes that value. The expression in the body that defines the result value is enclosed in curly braces.

If the defined function is not recursive, as is the case for `max`, the result type can be omitted because it is automatically inferred by the compiler. However, it is often helpful to provide the result type anyway to document the signature of the function.

Once you have defined a function, you can call it using its name:

```
scala> max(6, 3)
val res3: Int = 3
```

Naturally, you can use values and functions that are defined outside of a function's body in the function's body:

```
scala> val pi = 3.14159
val pi: Double = 3.14159

scala> def circ(r: Double) = 2 * pi * r
def circ(x: Double): Double
```

You can also nest value and function definitions:

```
scala> def area(r: Double) =
         val pi = 3.14159
         def square(x: Double) = x * x
         pi * square(r)

def area(Double): Double
```

Note that the scope of the body of a function definition is determined automatically by the indentation level. For instance, the following code does not compile because the last line is no longer interpreted to be part of the body of the function `area`:

```
scala> def area(r: Double) =
         val pi = 3.14159
         def square(x: Double) = x * x
       pi * square(r)
-- [E006] Not Found Error: ---------------------------------------------
4 |pi * square(r)
  |^^
  |Not found: pi
```

If you have a longer function definition, it can be helpful to mark the end of the function explicitly using an **end** marker, followed by the name of the function:

```
def area(r: Double) =
  val pi = 3.14159
  def square(x: Double) = x * x
  pi * square(r)
end area
```

Recursive functions can be written as expected. For example, the following function `fac` computes the factorial numbers:

```
scala> def fac(n: Int): Int = if n <= 0 then 1 else n*fac(n-1)
def fac(n: Int): Int

scala> fac(5)
val res4: Int = 120
```

1.2.4 Scopes

You can use curly braces `{ ... }` to create block scopes. Scala's scoping rules are almost identical to Java's:

```
val a = 5
// only a in scope
{
  val b = 4
  // b and a in scope

  def f(x: Int) =
    // f, x, b, and a in scope
    a * x + b

  // f, b, and a in scope
}
// only a in scope
```

There are two difference to Java, though. First, the scope of a name extends the entire block in which it is defined. Using a name before its definition leads to an error:

```
val a = 3
{
  val b = a // Refers to 'a' defined on the next line.
  val a = 4 // Does not compile.
}
```

However, unlike Java, Scala allows you to redefine names in nested scopes, thereby shadowing definitions in outer scopes.

```
val a = 3
{
  val a = 4 // Shadows outer definition of a.
  a + a     // Yields 8
}
```

As in Java, you cannot redefine a name in the same scope:

```
val a = 3
val a = 4 // Does not compile.
```

1.2.5 Tuples

Scala provides ways to create new compound data types without requiring you to define simplistic data-heavy classes. One of the most useful of these constructs are *tuples*. A tuple combines a fixed number of items together so that they can be passed around as a whole. The individual items can have different types. For example, here is a tuple holding an `Int` and a `String`:

```
scala> val p = (1, "banana")
val p: (Int, String) = (1, "banana")
```

and here is a tuple holding three items: two `String`s and a `Double` value:

```
scala> val q = ("apple", "pear", 1.0)
val q: (String, String, Double) = (apple, pear, 1.0)
```

To access the items of a tuple, you can use method `_1` to access the first item, method `_2` to access the second, and so on:

```
scala> p._1
val res5: Int = 1

scala> p._2
val res6: String = banana
```

Additionally, you can assign each element of the tuple to its own **`val`**:

```
scala> val (fst, snd) = p
val fst: Int = 1
val snd: String = banana
```

Be aware that tuples are not automatically decomposed when you pass them to functions:

```
def f(x: Int, s: String) = x

f(p._1, p._2) // Works.
f(p) // Does not compile.

def g(p: (Int, String)) = p._1

g(p) // Works.
g((1, "banana")) // Works.
g(1, "banana") // Works.
```

1.3 Recursion

Recursion will be our main device for expressing unbounded computations. In the following, we study how recursive functions are evaluated. We will further see that there is a close connection between certain recursive functions and loops in imperative programs.

1.3.1 Evaluating Recursive Functions

Consider the following function which computes the sum of the integer values in the interval given by the parameters `a` and `b`.

```
def sum(a: Int, b: Int): Int =
  if a < b then a + sum(a + 1, b) else 0
```

How are calls to such functions evaluated? Conceptually, we can think of the evaluation of a Scala expression as a process that rewrites expressions into simpler expressions. This rewriting process terminates when we obtain an expression that cannot be further simplified, e.g., an integer number. Expressions that cannot be simplified further are called *values*. Concretely, if we have a function call such as `sum(1 + 1, 0 + 2)`, we proceed as follows to compute a value using rewriting:

- First, we rewrite the call expression by rewriting the arguments of the call until they are reduced to values. In our example, this step yields the simplified call expression `sum(2, 2)`.
- Next, we replace the entire call expression by the body of the function. At the same time, we replace the formal parameters occurring in the function body (i.e., the occurrences of `a` and `b` in the example) by the actual arguments of the call. In our example, this step yields the expression

  ```
  if 2 < 2 then 2 + sum(2 + 1, 2) else 0
  ```

- Finally, we continue rewriting the function body recursively in the same manner until we obtain a value that cannot be simplified further. In our example, this process eventually terminates, producing the result value `0`.

Here is how we compute the value of `sum(1, 4)` using rewriting:

```
sum(1, 4)
-> if 1 < 4 then 1 + sum(1 + 1, 4) else 0
-> if true then 1 + sum(1 + 1, 4) else 0
-> 1 + sum(1 + 1, 4)
-> 1 + sum(2, 4)
-> 1 + (if 2 < 4 then 2 + sum(2 + 1, 4) else 0)
-> 1 + (if true then 2 + sum(2 + 1, 4) else 0)
-> 1 + (2 + sum(2 + 1, 4))
-> 1 + (2 + sum(3, 4))
-> 1 + (2 + (if 3 < 4 then 3 + sum(3 + 1, 4) else 0))
-> 1 + (2 + (if true then 3 + sum(3 + 1, 4) else 0))
-> 1 + (2 + (3 + sum(3 + 1, 4)))
-> 1 + (2 + (3 + sum(4, 4)))
-> 1 + (2 + (3 + (if 4 < 4 then 4 + sum(4 + 1, 4) else 0)))
-> 1 + (2 + (3 + (if false then 4 + sum(4 + 1, 4) else 0)))
-> 1 + (2 + (3 + 0))
```

```
-> 1 + (2 + 3)
-> 1 + 5
-> 6
```

We refer to this sequence of rewriting steps as an *execution trace*.

Termination. Does the rewriting process always terminate and produce a finite execution trace? Consider the following recursive function:

```
def loop(x: Int): Int = loop(x)
```

If we evaluate, e.g., the call `loop(0)`, we obtain an infinite rewriting sequence:

```
loop(0) -> loop(0) -> loop(0) -> ...
```

In order to guarantee termination of a recursive function, we have to make sure that each recursive call makes progress according to some progress measure. For example, in the recursive call to the function `sum` in our example above, the difference `b - a` between the arguments decreases with every recursive call. This means that `b - a` will eventually reach 0 or become negative. At this point, we take the `else` branch in the body of `sum` and the recursion stops. For our non-terminating function `loop`, it is impossible to find such a progress measure.

1.3.2 Tail Recursion

If we apply the function `sum` to larger intervals we observe the following:

```
scala> sum(1, 1000000)
java.lang.StackOverflowError
...
```

The problem is that a call to a function requires the Scala runtime environment to allocate stack space that stores the arguments of the call and any intermediate results obtained during the evaluation of the function body in memory. For the function `sum`, the intermediate results of the evaluation must be kept on the stack until the final recursive call returns. We can see this nicely in the execution trace for the call `sum(1, 4)`. The length of the expression that we still need to simplify grows with each recursive call:

```
sum(1, 4)
-> ...
-> 1 + sum(2, 4)
-> ...
-> 1 + (2 + sum(2 + 1, 4))
-> ...
-> 1 + (2 + (3 + sum(2 + 1, 4)))
-> ...
-> 6
```

Only when the final call to `sum` has returned, can we simplify the entire expression to a value.

During execution of a Scala expression, the arguments of functions that have been called, but have not yet returned, are maintained on the *call stack*. The stack space that is needed for evaluating a call `sum(a, b)` grows linearly with the recursion depth, which is given by the size of the interval `b - a`. Since the Scala runtime environment only reserves a relatively small amount of memory for the call stack, a call to `sum` for large interval sizes runs out of stack space. This is signaled by a `StackOverflowError` exception.

Can we implement the function `sum` so that it only requires constant space? To this end, consider the following *imperative* implementation of `sum`, which uses a **`while`** loop and mutable variables to perform the summation:

```
def sumImp(a: Int, b: Int): Int =
  var acc = 0
  var i = a
  while i < b do
    acc = i + acc
    i = i + 1
  acc
```

This implementation requires only constant space, since it involves only a single function call. Moreover, the execution of a single loop iteration for the summation does not allocate memory that persists across iterations. The intermediate results are stored in the variables `i` and `acc`, which are reused in each iteration. Unfortunately, this implementation uses mutable variables. Mutable variables make it more difficult to reason about the correctness of the code. However, we can turn the imperative **`while`** loop into a recursive function by hoisting the loop counter `i` and accumulator `acc` to function parameters:

```
def loop(acc: Int, i: Int, b: Int): Int =
  if i < b then loop(i + acc, i + 1, b) else acc

def sumTail(a: Int, b: Int): Int =
  loop(0, a, b)
```

Note how the function `loop` closely mimics the **`while`** loop in the imperative implementation without relying on mutable variables. We simply pass the new values that we obtain for the loop counter `i` and the accumulator `acc` to the recursive call of `loop`.

The function `loop` has an important property: the recursive call to `loop` in the *then* branch of the conditional expression is the final computation that is performed before the function returns. That is, in the recursive case, the function directly returns the result of the recursive call. We refer to functions in which all recursive calls are of this form as *tail-recursive* functions. Contrast the new implementation of `sum` with our original implementation, which added `a` to the result of the recursive call and was therefore not tail-recursive. The tail recursive implementation has an interesting effect on the execution trace:

```
sumTail(1, 4)
   -> loop(0, 1, 4)
   -> if 1 < 4 then loop(1 + 0, 1 + 1, 4) else 0
   -> if true then loop(1 + 0, 1 + 1, 4) else 0
   -> loop(1, 2, 4)
   -> if 2 < 4 then loop(2 + 1, 2 + 1, 4) else 1
   -> if true then loop(2 + 1, 2 + 1, 4) else 1
   -> loop(3, 3, 4)
   -> if 3 < 4 then loop(3 + 3, 3 + 1, 4) else 3
   -> if true then loop(3 + 3, 3 + 1, 4) else 3
   -> loop(6, 3, 4)
   -> if 4 < 4 then loop(4 + 6, 4 + 1, 4) else 6
   -> if false then loop(4 + 6, 4 + 1, 4) else 6
   -> 6
```

Observe that the size of the expressions that we obtain throughout the trace does not grow with the recursion depth. This is because the tail-recursive call to `loop` is the final computation that is performed in the body of `loop`, before the function returns.

To simplify our implementation, we can move the declaration of the function `loop` inside the body of the function `sumTail`:

```
def sumTail(a: Int, b: Int): Int =
  def loop(acc: Int, i: Int): Int =
    if i < b then loop(i + acc, i + 1) else acc
  loop(0, a)
```

Note that in this version, we have dropped the third parameter `b` of the first version of the function `loop` since it is just passed to the recursive call without change. The occurrence of `b` in the body of the new nested version of `loop` now always refers to the parameter `b` of the outer function `sumTail`.

For tail-recursive functions, the stack space that is allocated for the current call can be reused by the recursive call. In particular, the memory that is needed to store the arguments of the current call can be reused to store the arguments of the recursive call. By reusing the current stack space, we effectively turn the recursive function back into an imperative loop. This optimization is referred to as *tail call elimination*. Many modern compilers, including the Scala compiler, automatically eliminate tail calls. Thus, tail-recursive functions are guaranteed to execute in constant stack space. We can test this feature by rerunning the tail-recursive version of `sum` for large interval sizes:

```
scala> sumTail(1, 1000000)
val res0: Int = 704982704
```

This time the function terminates normally without throwing an exception.

With tail call elimination we get the best of both worlds: we obtain the efficiency of an imperative implementation and the simplicity of a functional implementation. If you are unsure about how to write a tail-recursive function, it is often helpful to first write the function using a **`while`** loop and then transform the loop into a tail-recursive function, as we have done above. Once you get more used to functional programming, you will find writing tail-recursive functions as natural as writing loops.

If you are in doubt whether a recursive function that you wrote is tail-recursive, you can add the `@tailrec` annotation to the declaration of the function:

```
import scala.annotation.tailrec
...
def sumTail(a: Int, b: Int): Int =
  @tailrec def loop(acc: Int, i: Int): Int =
    if i < b then loop(i + acc, i + 1) else acc
  loop(0, a)
```

If the compiler fails to apply tail call elimination to a `@tailrec` annotated function, then it will issue a warning:

```
-- Error: ----------------------------------------------------
2 |  if a < b then a + sum(a + 1, b) else 0
  |                    ^^^^^^^^^^^^^
  |   Cannot rewrite recursive call: it is not in tail position
```

You may wonder whether non-tail-recursive functions should be avoided at all costs. This depends on the function. Often, tail-recursive functions are harder to understand than a recursive function that performs the same computation, but that is not tail-recursive. [1]. If you know that the recursion depth of the calls to your function will be small in practice, you may want to write the function without tail-recursion. In general, if you are in doubt, you should always value the clarity of your code higher than its efficiency. When you observe that your code is inefficient, you can still optimize it later.

1.4 Classes and Objects

In the previous sections, we have learned about the basic language features of Scala. In this section, we will learn how Scala programs are organized. Scala is an object-oriented language, so Scala programs are organized using *classes* and *objects*.

1.4.1 Classes, Fields, and Methods

Similar to Java, Scala allows you to define classes with *fields* and *methods*, which you can extend using inheritance, override, etc. Fortunately, Scala's syntax for classes is much more lightweight than Java's. For example, consider the following Java class which we can use to wrap pairs of integer values in a single

[1] Similarly, recursive functions are easier to understand than computations that use imperative loops.

object:

```
public class Pair {
  private int first;
  private int second;

  public Pair(int fst, int snd) {
    first = fst;
    second = snd;
  }

  public int getFirst() {
    return first;
  }

  public int getSecond() {
    return second;
  }
}
```

The class consists of:

- two fields called `first` and `second` of type `int` to store the two values;
- a constructor, which takes values to initialize the two fields;
- two "getter" methods to retrieve the two values (we follow good practice and declare all non-final fields as private so that their values cannot be modified without explicit method calls.).

Let us ignore for the moment that we can represent pairs directly in Scala using a tuple type. Here is how we can define the corresponding class in Scala:

```
class Pair(val first: Int, val second: Int)
```

There are some important differences between the Java and Scala version of the class `Pair`:

- In Scala, the class name is followed by a list of class parameters. These parameters serve two purposes:
  1. Parameters that are prefixed by a **`val`** or **`var`** keyword automatically create a field with the given name and type.
  2. The parameter list implicitly defines a constructor with a corresponding list of arguments. The values that are provided for arguments prefixed with **`val`** or **`var`** will be used to initialize the associated fields.
- The default visibility of classes, fields, and methods in Scala is `public`. Hence, we can access the values of the fields `first` and `second` directly and we do not need to define extra getter methods. Note that this makes

sense because Scala discourages mutable state. In particular, we defined the two fields as **val**s, so their values cannot be changed, once an instance of class `Pair` has been created. In Scala, you can leave out the braces around an empty class body, so **`class`** `C` is the same as **`class`** `C {}`.

We can create instances of class `Pair` and access their fields as usual:

```
scala> val p = new Pair(1,2)
val p: Pair = Pair@1458e1cc
scala> p.first
val res0: Int = 1
```

What if we do want to modify the values stored in a `Pair` object? In Java, we would do this by adding appropriate "setter" methods to the class:

```
public class Pair {
  private int first;
  private int second;

  ...

  public void setFirst(int fst) {
    first = fst;
  }
  public void setSecond(int snd) {
    second = snd;
  }
}
```

In Scala, we could follow the same route: change all **val**s into **var**s, make them private, and add getter and setter methods. However, we want to avoid using **var** declarations as much as possible. The idiomatic solution in Scala is to make a copy of the entire object and change the appropriate value:

```
class Pair(val first: Int, val second: Int):
  def setFirst(fst: Int): Pair = new Pair(fst, second)
  def setSecond(snd: Int): Pair = new Pair(first, snd)
```

1.4.2 Overriding Methods

Java allows us to override methods that are declared in super classes. Since method calls are dynamically dispatched at run-time, this feature allows us to modify the behavior of an object of the subclass when it is used in a context where an object of the super class is expected.

All Java classes extend the class `Object`. The class `Object` provides, among others, a method `toString`, which computes a textual representation of the object. In particular, the `toString` method can be used to pretty-print objects. By default, the textual representation of objects consists of the name of the object's class, followed by a unique object ID. We can modify the way objects of a specific class are printed, by overriding the `toString` method. In Java, this can be done as follows:

```
public class Pair {
  private int first;
  private int second;

  ...

  public String toString() {
    return "Pair(" + first + ",␣" + second + ")";
  }
}
```

In Scala, all classes extend the class `scala.Any` which also provides a method called `toString`. Scala's class hierarchy is further subdivided into the classes `scala.AnyVal` and `scala.AnyRef`, which are directly derived from `scala.Any`. All instances of `scala.AnyVal` are immutable, whereas instances of `scala.AnyRef` may have mutable state. That is, `scala.AnyRef` corresponds to Java's `Object` class.

If we want to override a method in a Scala class, we have to explicitly say so by using the **`override`** qualifier:

```
class Pair(val first: Int, val second: Int):
  ...
  override def toString = "Pair(" + first + ",␣" + second + ")"
```

The pretty printer in the REPL will now use the new `toString` method to print `Pair` objects:

```
scala> val p = new Pair(1,2)
val p: Pair = Pair(1, 2)
```

1.4.3 Singleton and Companion Objects

If the construction of an object involves complex initialization code, it is often useful to declare dedicate *factory* methods that perform this initialization. In Java, we would declare such methods as *static* members of the corresponding class:

```
public class Pair {
  ...
  public static Pair make(int fst, int, snd) {
    return new Pair(fst, snd);
  }
}
```

We can now call `Pair.make` to create new `Pair` instances.

Scala does not support static methods as they violate the philosophy of object-oriented programming that "everything is an object". Instead of static methods, it provides *singleton objects*. Singleton objects are declared just like classes, but using the keyword **object** instead of **class**. There exists exactly one instance of each **object**, which is automatically created from the **object** declaration when the program is started. Since no further instances of the object can be created, object declarations do not have parameter lists.

For every class `C` in a Scala program, one can declare a singleton object that is also called `C`. This object is referred to as the *companion object* of `C`. Companion objects have access to all private members of instances of `C`. Consequently, a method or field that is defined in the companion object is equivalent to a static method/field of `C` in Java:

```
class Pair(val first: Int, val second: Int):
  ...
object Pair:
  def make(fst: Int, snd: Int) = new Pair(fst, snd)
```

We can access members of companion objects just like static class members in Java:

```
scala> def p = Pair.make(3,4)
val p: Pair = Pair(3, 4)
```

1.4.4 The `apply` Method

Methods with the name `apply` are treated specially by the Scala compiler. For example, if we rename the factory method `make` in our companion object for the `Pair` class to `apply`

```
object Pair:
  def apply(fst: Int, snd: Int) = new Pair(fst, snd)
```

then we can call this method simply by referring to the `Pair` companion object, followed by the argument list of the call:

```
scala> def p = Pair(3,4)
p: Pair = Pair(3, 4)
```

This is equivalent to the following explicit call to the `apply` method:

```
scala> def p = Pair.apply(3,4)
p: Pair = Pair(3, 4)
```

The compiler automatically expands `Pair(3,4)` to `Pair.apply(3, 4)`. That is, objects with an `apply` method can be used as if they were functions[2]. This feature is particularly useful to enable concise calls to factory methods. In fact, factory methods for the data structures in the Scala standard library are typically implemented using `apply` methods provided by companion objects.

[2] If you are familiar with C++, then you will notice that this feature is similar to overloading the function call operator `()` in C++.

1.5 Algebraic Data Types

Algebraic data types (ADTs) and pattern matching are constructs that are commonly found in functional programming languages. They allow you to implement regular, non-encapsulated data structures (such as lists and trees) in a convenient fashion. We will make heavy use of this feature throughout this course.

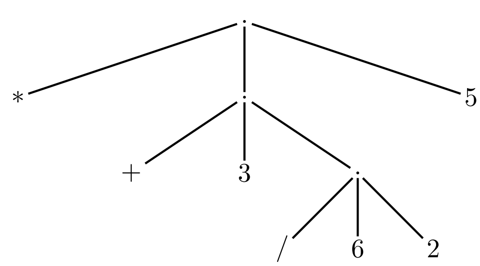
1.5.1 Enumerations

Suppose we want to implement a simple calculator program that takes arithmetic expressions such as

$$(3 + 6/2) * 5$$

as input and evaluates these expressions. This problem is quite similar to writing an interpreter for a programming language, except that the language that we are interpreting here (i.e., arithmetic expressions) is much simpler than a full-blown programming language.

One of the first questions that we have to answer is: how do we represent expressions in our program? Our representation should allow us to easily implement common tasks such as pretty printing, evaluation, and simplification of expressions. In particular, the representation should make the precedence of operators in expressions explicit. E.g., consider the expression $3 + 6/2$, then when we evaluate the expression, our representation should immediately tell us that we first have to divide 6 by 2 before we add 3. To achieve this, expressions are represented as *abstract syntax trees*, or ASTs for short. For example, the abstract syntax tree of the expression $(3 + 6/2) * 5$ can be visualized as follows:



For the AST node representing an expression $e_1\, op\, e_2$, we follow the convention of representing the operator op as the left-most child of the node, rather than putting it between the children representing the operands e_1 and e_2.

Note that the AST tells us exactly how to evaluate the expression. We start at the root. At each node that we visit, we first recurse into the second subtree to evaluate the left operand of the operation. Then we do the same for the third subtree, which represents the right operand of the operation. Finally, we combine the results obtained from the two operands according to the operator labeling the first child. We will learn more about ASTs later. For now it suffices if you have an intuitive understanding what ASTs are.

Algebraic data types allow us to represent tree-like data structures such as ASTs. In Scala, algebraic data types are constructed using *enumerations* or “enums” for short. The following enum defines the ASTs of our arithmetic expressions:

```
enum Expr:
  /* Numbers such as 1, 2, etc. */
  case Num(num: Int)
  /* Expressions composed using binary operators */
  case BinOp(op: Bop, left: Expr, right: Expr)

/* Binary operators */
enum Bop:
  case Add /* + */
  case Sub /* - */
  case Mul /* * */
  case Div /* / */
```

This code declares an enum `Expr` whose instances represent the ASTs of our expression language. The code distinguishes two types of `Expr` objects based on the root node of the represented AST. A `Num(num)` object represents an AST consisting of a single node storing the integer constant `num`. Similarly, an object `BinOp(op, left, right)` represents an AST whose root node combines two subexpressions represented by ASTs `left` and `right` using the *binary operator* `op`. Binary operators are represented by the enum `Bop` which consists of four cases representing the different kinds of arithmetic operations: `Add`, `Sub`, `Mul`, and `Div`.

We refer to the cases of an algebraic data type as its *variants*. For example, `Expr` has variants `Num` and `BinOp`.

The Scala compiler adds some convenient functionality to enums. First, it automatically generates companion objects with appropriate factory methods. These methods are particularly useful when you nest them to construct complex expressions:

```
scala> import Expr._, Bop._
scala> val e = BinOp(Add, Num(3), BinOp(Mul, Num(4), Num(5)))
val e: Expr = BinOp(Add, Num(3), BinOp(Mul, Num(4), Num(5)))
```

Note that the enums `Expr` and `Bop` itself have companion objects. In turn, these companion objects have nested companion objects for the variants of the defined ADT within them. For example, the companion object for `Expr` contains companion objects for `Num` and `BinOp`. The nested companion objects `Num` and `BinOp` then have `apply` methods for constructing objects of the corresponding variant. The **`import`** instruction in the above code snippet makes the nested companion objects directly accessible in the code. For instance, in the code `Num(4)`, `Num` refers to the companion object of variant `Num` in `Expr`. Thus, this code expands to `Expr.Num.apply(4)`.

Second, the compiler adds natural implementations of the methods `toString`, `hashCode`, and `equals` for each variant. These will print, hash, and compare a whole tree consisting of the top-level enum instance and (recursively) all its arguments. In Scala, an expression of the form `x == y` always translates into a call of the form `x.equals(y)` (just like in Java). The overriden `equals` method therefore ensures that enum instances are always compared structurally. For example, we have:

```
scala> val e1: BinOp = BinOp(Add, Num(3), Num(4))
val e1: Expr.BinOp = BinOp(Add, Num(3), Num(4))
scala> val e2 = BinOp(Add, Num(3), Num(4))
val e2: Expr.BinOp = BinOp(Add, Num(3), Num(4))
scala> e1 == e2
val res2: Boolean = true
```

In the example above, `e1` and `e2` point to two different objects in memory. However, the two ASTs represented by these objects have exactly the same structure. Hence, `e1 == e2` evaluates to **`true`**.

Next, the compiler implicitly adds a **`val`** prefix to all arguments in the parameter list of an enum object, so that they are maintained as fields:

```
scala> val n = e1.left
val n: Expr = Num(3)
```

Note that the above code only works because we explicitly declared `e1` to be of type `BinOp` earlier. Without this type annotation, the inferred type of `e1` would be `Expr`. However, `Expr` objects are not guaranteed to have a field called `left` since `Expr` also includes `Num` objects.

Finally, the compiler adds a `copy` method to your enum cases for making modified copies. This method is useful if you need to create a copy of an existing enum object `o` that is identical to `o` except for some of `o`'s attributes:

```
scala> e1.copy(op = Sub)
val res4: Expr = BinOp(Sub, Num(3), Num(4))
```

1.5.2 Pattern Matching

Suppose we want to implement an algorithm that simplifies expressions by recursively applying the following simplifications rules:

- $e + 0 \Rightarrow e$
- $e * 1 \Rightarrow e$
- $e * 0 \Rightarrow 0$

To identify whether a given expression matches one of the left-hand sides of the rules, we have to look at some of its subexpressions. E.g., to check whether an expression of the form $e_1 + e_2$ matches the left-hand side of the first rule, we have to look at the left subexpression e_1 to check whether $e_1 = 0$. Implementing this kind of pattern matching is quite tedious in many languages (including Java). Fortunately, the Scala language has inbuilt support for pattern matching that works hand-in-hand with enumerations.

Let us first reformulate the three simplification rules in terms of our enum representation of expressions:

```
BinOp(Add, e1, Num(0)) => e1
BinOp(Mul, e1, Num(1)) => e1
BinOp(Mul, e1, Num(0)) => Num(0)
```

Using pattern matching, these rules almost directly give us the implementation of the following function `simplifyTop`, which applies the rules at the top-level of the given expression `e`:

```
def simplifyTop(e: Expr) =
  e match
    case BinOp(Add, e1, Num(0)) => e1
    case BinOp(Mul, e1, Num(1)) => e1
    case BinOp(Mul, _, Num(0)) => Num(0)
    case _ => e
```

The body of `simplifyTop` is a *match expression*. A match expression consists of a *selector*, in this case `e`, followed by the keyword **`match`**, followed by a sequence of match alternatives.

Each match alternative starts with the keyword **`case`**, followed by a pattern, followed by an expression that is evaluated if the pattern matches the selector. The pattern and expression are separated by an arrow symbol **`=>`**.

A match expression is evaluated by checking whether the selector matches one of the patterns in the alternatives. The patterns are tried in the order in which they appear in the program. The first pattern that matches is selected and the expression following the arrow is evaluated. The result of the entire match expression is the result of the expression in the selected alternative.

Here is an example of a recursive function that uses pattern matching to pretty print arithmetic expressions:

```
def pretty(e: Expr): String =
  e match
    case BinOp(bop, e1, e2) =>
      val bop_str = bop match
        case Add => " + "
        case Sub => " - "
        case Mul => " * "
        case Div => " / "
      "(" + pretty(e1) + bop_str + pretty(e2) + ")"
    case Num(n) => n.toString()
```

```
scala> val e = BinOp(Add, BinOp(Mult, Num(3), Num(4)), Num(1))
val e: BinOp = BinOp(Add, BinOp(Mult, Num(3), Num(4)), Num(1))
scala> pretty(e)
val res0: String = ((3 * 4) + 1)
```

There are different types of patterns. The most important types are:

- *Constant patterns*: A constant pattern such as `0` matches values that are equal to the constant (with respect to `==`).
- *Variable patterns*: A variable pattern such as `e1` matches every value. Here, `e1` is a variable that is bound in the pattern. The variable refers to the matched value in the right-hand side of the match alternative.
- *Wildcard patterns*: A wildcard pattern `_` also matches every value, but it does not introduce a variable that refers to the matched value.
- *Constructor patterns*: A constructor pattern such as `BinOp(Add, e, Num(0))` matches all values of type `BinOp` whose first argument matches `Add`, whose second argument matches `e`, and whose third argument matches `Num(0)`. Note that the arguments to the constructor `BinOp` are themselves patterns. This allows you to write deep patterns that match complex enum values using a concise notation.

1.5.3 Binding Names in Patterns

Sometimes we want to match a subexpression against a specific pattern and also bind the matched expression to a name. This is useful when we want to reuse a matched subexpression in the right-hand side of the match alternative. For example, in the third simplification rule of `simplifyTop` we are returning `Num(0)`, which is also the second subexpression of the matched expression `e`. Instead of creating a new expression, `Num(0)` on the right-hand side of the rule, we can also directly return the second subexpression of `e`. We can do this by binding a name to that subexpression in the pattern using the operator `@` as follows:

```
def simplifyTop(e: Expr) =
  e match
    case BinOp(Add, e1, Num(0)) => e1
    case BinOp(Mul, e1, Num(1)) => e1
    case BinOp(Mul, _, e2 @ Num(0)) => e2
    case _ => e
```

Note that the pattern in the third match alternative now binds the name `e2` to the value matched by the pattern `Num(0)`. This value is then returned on the righ-hand side of the rule by referring to it using the name `e2`.

1.5.4 Pattern Guards

Suppose we want to extend our expression simplifier so that it additionally implements the following simplification rule: $e + e \Rightarrow 2 \times e$

If we directly translate the rule to a corresponding match alternative, we obtain the following implementation of `simplilfyTop`:

```
def simplifyTop(e: Expr) =
  e match
```

```
    ...
    case BinOp(Add, e1, e1) => BinOp(Mul, Num(2), e1)
    case _ => e
```

Unfortunately, the compiler will reject this function because we use the name `e1` twice within the same pattern. In general, a variable name such as `e1` may only be used once in a pattern. We can solve this problem by using a different variable name for the second subexpression, say `e2`, and then use a *pattern guard* to enforce that the subexpressions matched by `e1` and `e2` are equal:

```
def simplifyTop(e: Expr) =
  e match
    ...
    case BinOp(Add, e1, e2) if e1 == e2 =>
      BinOp(Mul, Num(2), e2)
    case _ => e
```

In general, a pattern guard can be an arbitrary Boolean expression over the names that are in the scope of the match alternative. The pattern guard is appended to the pattern of a match alternative using the keyword **`if`**.

1.5.5 Exhaustiveness Checks

Consider the following function `simplifyAll` that applies our simplification rules recursively to the given expression:

```
def simplifyAll(e: Expr): Expr =
  e match
    case BinOp(Add, e1, Num(0)) => simplifyAll(e1)
    case BinOp(Mul, e1, Num(1)) => simplifyAll(e1)
    case BinOp(Mul, _, e2 @ Num(0)) => e2
    case BinOp(Add, e1, e2) if e1 == e2 =>
      BinOp(Mul, Num(2), simplifyAll(e2))
    case BinOp(bop, e1, e2) =>
      BinOp(bop, simplifyAll(e), simplifyAll(e2))
```

Observe that in this function, the pattern alternatives are no longer exhaustive. That is, there exist values `e` that are not matched by any of the match alternatives, e.g., the value `Num(0)`. If `simplifyAll` is called with `Num(0)`, it will throw a runtime exception.

We can fix this code by adding an explicit match alternative for the `Num` constructor:

```
def simplifyAll(e: Expr): Expr =
  e match
    case BinOp(Add, e1, Num(0)) => simplifyAll(e1)
    case BinOp(Mul, e1, Num(1)) => simplifyAll(e1)
    case BinOp(Mul, _, e2 @ Num(0)) => e2
    case BinOp(Add, e1, e2) if e1 == e2 =>
```

```
      BinOp(Mul, Num(2), simplifyAll(e2))
    case BinOp(bop, e1, e2) =>
      BinOp(bop, simplifyAll(e), simplifyAll(e2))
    case Num(_) => e
```

With complex patterns like this it can be tricky to keep track of all the possible cases. Fortunately, the compiler will automatically check whether the case analysis is exhaustive. For instance, consider the following faulty implementation of `simplifyAll` where we have omitted the last case for `BinOp` from the previous implementation:

```
def simplifyAll(e: Expr): Expr =
  e match
    case BinOp(Add, e1, Num(0)) => simplifyAll(e1)
    case BinOp(Mul, e1, Num(1)) => simplifyAll(e1)
    case BinOp(Mul, _, e2 @ Num(0)) => e2
    case BinOp(Add, e1, e2) if e1 == e2 =>
      BinOp(Mul, Num(2), simplifyAll(e2))
    case Num(_) => e
```

This version does not compile:

```
-- [E029] Pattern Match Exhaustivity Warning: ----------
2 |  e match
  |  ^
  |  match may not be exhaustive.
  |
  |  It would fail on pattern case: Expr.BinOp(_, _, _)
```

Unfortunately, these exhaustiveness checks can sometimes produce spurious warnings. For example, suppose we have a function that is meant to pretty-print number expressions, but not other expressions which have not yet been reduced:

```
def prettyNumber(e: Expr): String =
  e match
    case Num(num) => num.toString()
```

Further suppose that we know that our program ensures that `prettyNumber` is never called on a `BinOp` expression. Yet, the compiler still complains about the non-exhaustive pattern matching. We can suppress this warning by declaring `e` as *unchecked*:

```
def prettyNumber(e: Expr): String =
  (e: @unchecked) match
    case Num(num) => num.toString()
```

While `@unchecked` notations are sometimes necessary to suppress spurious warnings, you should be very careful about introducing them in your code. In most cases, the compiler generated warnings indicate actual problems in your code that need your attention.

1.5.6 Option Types

Suppose we want to write a function that evaluates arithmetic expressions to `Int` values. One question is: How should we deal with undefined operations such as division by zero:

```
BinOp(Div, e, Num(0)) => ???
```

In Java, we would typically go for one of the following two solutions:

- throw an exception such as `ArithmeticException`;
- return **`null`** to indicate that the intended operation does not yield a valid result.

Both approaches have advantages and drawbacks.

Exceptions are a good solution if the undefined operation is indeed exceptional behavior that should, e.g., abort the program. In this case, we ensure that a computation that returns normally always yields a valid result. However, if the undefined operation commonly occurs in computations, we will have to catch the exception and handle it appropriately. This has two disadvantages. First, the exception mechanism is relatively expensive and should only be used in truly exceptional situations. Second, the exception handlers will clutter the code and the non-structured control flow of thrown exceptions makes it more difficult to understand what the program is doing.

If we return **`null`**, we avoid the two disadvantages of exceptions: the computation always returns normally, and there is no computational overhead such as recording the stack-trace to the point where the exception was thrown. However, **`null`** values introduce their own problems. Since **`null`** can have an arbitrary type, the type checker of the compiler will give us much weaker static correctness guarantees for our code. In particular, it will be unable to statically detect unintended accesses to the return value in cases where the return value is invalid (hello `NullPointerException`!).

In languages that support pattern matching, there is a common idiom that avoids the problem of introducing **`null`** values: option types.

The option type is an algebraic data type with two variants: `Some(v)` to indicate that a computation returned a proper result value `v`, and `None` to indicate that the intended operation was undefined and has no proper result.

In Scala, we can define an option type for `Int` values using an enum as follows:

```
enum IntOption:
  case Some(value: Int)
  case None
```

We can now use the option type similarly to null values in Java:

```
def div(x: Int, y: Int): IntOption =
  if y == 0 then None else Some(x / y)
```

Unlike in Java, where the static type checker is unable to distinguish a `null` value from a genuine result of a computation, the Scala type checker will force us to explicitly unwrap the `Int` value embedded in an `IntOption` before we can access it. Using pattern matching, we can do this conveniently. For example, suppose we want to convert the result of `div` to a double precision floating point number. By using pattern matching on the return value of `div`, we can recover from some of the cases where integer division by 0 is undefined:

```
def divToDouble(x: Int, y: Int): Double =
  div(x,y) match
    case Some(x) => x
    case None =>
      if x < 0 then Double.NegativeInfinity
      else if x > 0 then Double.PositiveInfinity
      else Double.NaN
```

Since option types are so useful, Scala already provides a generic option type, called `Option`, in its standard library. Using the predefined type `Option` we can write the function `div` like this:

```
def div(x: Int, y: Int): Option[Int] =
  if y == 0 then None else Some(x / y)
```