## Other Collections

Principles of Functional Programming

#### Other Sequences

We have seen that lists are *linear*: Access to the first element is much faster than access to the middle or end of a list.

The Scala library also defines an alternative sequence implementation, Vector.

This one has more evenly balanced access patterns than List.

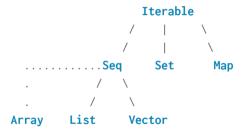
### Operations on Vectors

Vectors are created analogously to lists:

```
val nums = Vector(1, 2, 3, -88)
  val people = Vector("Bob", "James", "Peter")
They support the same operations as lists, with the exception of ::
Instead of x :: xs, there is
   x +: xs Create a new vector with leading element x, followed
             by all elements of xs.
   xs: + x Create a new vector with trailing element x, preceded
             by all elements of xs.
(Note that the : always points to the sequence.)
```

## Collection Hierarchy

A common base class of List and Vector is Seq, the class of all *sequences*. Seq itself is a subclass of Iterable.



### Arrays and Strings

Arrays and Strings support the same operations as Seq and can implicitly be converted to sequences where needed.

(They cannot be subclasses of Seq because they come from Java)

```
val xs: Array[Int] = Array(1, 2, 3)
xs.map(x => 2 * x)

val ys: String = "Hello world!"
ys.filter(_.isUpper)
```

#### Ranges

Another simple kind of sequence is the *range*.

It represents a sequence of evenly spaced integers.

Three operators:

to (inclusive), until (exclusive), by (to determine step value):

```
val r: Range = 1 until 5
val s: Range = 1 to 5
1 to 10 by 3
6 to 1 by -2
```

A Range is represented as a single object with three fields: lower bound, upper bound, step value.

## Some more Sequence Operations:

xs.exists(p)	true if there is an element x of xs such that $p(x)$ holds,
	false otherwise.
xs.forall(p)	true if $p(x)$ holds for all elements $x$ of $xs$ , false otherwise.
xs.zip(ys)	A sequence of pairs drawn from corresponding elements of sequences xs and ys.
xs.unzip	Splits a sequence of pairs $xs$ into two sequences consisting of the first, respectively second halves of all pairs.
xs.flatMap(f)	Applies collection-valued function f to all elements of xs and concatenates the results
xs.sum	The sum of all elements of this numeric collection.
xs.product	The product of all elements of this numeric collection
xs.max	The maximum of all elements of this collection (an Ordering must exist)
xs.min	The minimum of all elements of this collection

## Example: Combinations

To list all combinations of numbers x and y where x is drawn from 1..M and y is drawn from 1..N:

```
(1 to M).flatMap(x =>
```

### Example: Combinations

To list all combinations of numbers x and y where x is drawn from 1..M and y is drawn from 1..N:

```
(1 to M).flatMap(x \Rightarrow (1 to N).map(y \Rightarrow (x, y)))
```

To compute the scalar product of two vectors:

```
def scalarProduct(xs: Vector[Double], ys: Vector[Double]): Double =
    xs.zip(ys).map((x, y) => x * y).sum
```

To compute the scalar product of two vectors:

```
def scalarProduct(xs: Vector[Double], ys: Vector[Double]): Double =
    xs.zip(ys).map((x, y) => x * y).sum
```

Note that there is some automatic decomposition going on here.

Each pair of elements from xs and ys is split into its halves which are then passed as the x and y parameters to the lambda.

If we wanted to be more explicit, we could also write scalar product like this:

```
def scalarProduct(xs: Vector[Double], ys: Vector[Double]): Double =
    xs.zip(ys).map(xy => xy._1 * xy._2).sum
```

On the other hand, if we wanted to be more even more concise, we could also write it like this:

```
def scalarProduct(xs: Vector[Double], ys: Vector[Double]): Double =
    xs.zip(ys).map(_ * _).sum
```

A number n is *prime* if the only divisors of n are 1 and n itself.

What is a high-level way to write a test for primality of numbers? For once, value conciseness over efficiency.

```
def isPrime(n: Int): Boolean = ???
```

A number n is *prime* if the only divisors of n are 1 and n itself.

What is a high-level way to write a test for primality of numbers? For once, value conciseness over efficiency.

```
def isPrime(n: Int): Boolean =
   (2 to n - 1).forall(d => n % d == 0)
```

# Combinatorial Search and For-Expressions

Principles of Functional Programming

## Handling Nested Sequences

We can extend the usage of higher order functions on sequences to many calculations which are usually expressed using nested loops.

**Example**: Given a positive integer n, find all pairs of positive integers i and j, with  $1 \le j \le i \le n$  such that i + j is prime.

For example, if n = 7, the sought pairs are

## Algorithm

A natural way to do this is to:

- ▶ Generate the sequence of all pairs of integers (i, j) such that 1 <= j < i < n.</p>
- Filter the pairs for which i + j is prime.

One natural way to generate the sequence of pairs is to:

- Generate all the integers i between 1 and n (excluded).
- ► For each integer i, generate the list of pairs (i, 1), ..., (i, i-1).

This can be achieved by combining until and map:

```
(1 until n).map(i =>
  (1 until i).map(j => (i, j)))
```

#### Generate Pairs

The previous step gave a sequence of sequences, let's call it xss.

We can combine all the sub-sequences using foldRight with ++:

```
xss.foldRight(\textcolor{red}{\textbf{Seq[Int]}())(\_ ++ \_)}
```

Or, equivalently, we use the built-in method flatten

```
xss.flatten
```

This gives:

```
((1 until n).map(i =>
  (1 until i).map(j => (i, j)))).flatten
```

## Generate Pairs (2)

Here's a useful law:

```
xs.flatMap(f) = xs.map(f).flatten
```

Hence, the above expression can be simplified to

```
(1 until n).flatMap(i =>
      (1 until i).map(j => (i, j)))
```

## Assembling the pieces

By reassembling the pieces, we obtain the following expression:

```
(1 until n)
  .flatMap(i => (1 until i).map(j => (i, j)))
  .filter((x, y) => isPrime(x + y))
```

This works, but makes most people's head hurt.

Is there a simpler way?

### For-Expressions

Higher-order functions such as map, flatMap or filter provide powerful constructs for manipulating lists.

But sometimes the level of abstraction required by these function make the program difficult to understand.

In this case, Scala's for expression notation can help.

### For-Expression Example

Let persons be a list of elements of class Person, with fields name and age.

```
case class Person(name: String, age: Int)
```

To obtain the names of persons over 20 years old, you can write:

```
for p <- persons if p.age > 20 yield p.name
```

which is equivalent to:

```
persons
.filter(p => p.age > 20)
.map(p => p.name)
```

The for-expression is similar to loops in imperative languages, except that it builds a list of the results of all iterations.

## Syntax of For

A for-expression is of the form

for s yield e

where s is a sequence of *generators* and *filters*, and e is an expression whose value is returned by an iteration.

- ▶ A *generator* is of the form p <- e, where p is a pattern and e an expression whose value is a collection.
- A *filter* is of the form if f where f is a boolean expression.
- ► The sequence must start with a generator.
- ► If there are several generators in the sequence, the last generators vary faster than the first.

#### Use of For

Here are two examples which were previously solved with higher-order functions:

Given a positive integer n, find all the pairs of positive integers (i, j) such that  $1 \le j \le i \le n$ , and i + j is prime.

```
for
    i <- 1 until n
    j <- 1 until i
    if isPrime(i + j)
yield (i, j)</pre>
```

Write a version of scalarProduct (see last session) that makes use of a for:

```
def scalarProduct(xs: List[Double], ys: List[Double]) : Double =
```

Write a version of scalarProduct (see last session) that makes use of a for:

```
def scalarProduct(xs: List[Double], ys: List[Double]) : Double =
    (for (x, y) <- xs.zip(ys) yield x * y).sum</pre>
```

Write a version of scalarProduct (see last session) that makes use of a for:

```
def scalarProduct(xs: List[Double], ys: List[Double]) : Double =
    (for (x, y) <- xs.zip(ys) yield x * y).sum

Question: What will the following produce?
    (for x <- xs; y <- ys yield x * y).sum</pre>
```

Write a version of scalarProduct (see last session) that makes use of a for:

```
def scalarProduct(xs: List[Double], ys: List[Double]) : Double =
    (for (x, y) <- xs.zip(ys) yield x * y).sum

Question: What will the following produce?
    (for x <- xs; y <- ys yield x * y).sum</pre>
```

*Answer:* It would multiply *every* element of xs with *every* element of ys and sum up the results.

# Combinatorial Search Example

Principles of Functional Programming

#### Sets

Sets are another basic abstraction in the Scala collections.

A set is written analogously to a sequence:

```
val fruit = Set("apple", "banana", "pear")
val s = (1 to 6).toSet
```

Most operations on sequences are also available on sets:

```
s.map(_ + 2)
fruit.filter(_.startsWith("app"))
s.nonEmpty
```

(see Scaladoc for scala. Set for a list of all supported operations)

### Sets vs Sequences

The principal differences between sets and sequences are:

- 1. Sets are unordered; the elements of a set do not have a predefined order in which they appear in the set
- 2. sets do not have duplicate elements:

```
s.map(_ / 2) // Set(2, 0, 3, 1)
```

3. The fundamental operation on sets is contains:

```
s.contains(5) // true
```

#### Example: N-Queens

The eight queens problem is to place eight queens on a chessboard so that no queen is threatened by another.

In other words, there can't be two queens in the same row, column, or diagonal.

We now develop a solution for a chessboard of any size, not just 8.

One way to solve the problem is to place a queen on each row.

Once we have placed k-1 queens, one must place the kth queen in a column where it's not "in check" with any other queen on the board.

## Algorithm

We can solve this problem with a recursive algorithm:

- ► Suppose that we have already generated all the solutions consisting of placing k-1 queens on a board of size n.
- ► Each solution is represented by a list (of length k-1) containing the numbers of columns (between 0 and n-1).
- ► The column number of the queen in the k-1th row comes first in the list, followed by the column number of the queen in row k-2, etc.
- ► The solution set is thus represented as a set of lists, with one element for each solution.
- Now, to place the kth queen, we generate all possible extensions of each solution preceded by a new queen:

#### **Implementation**

```
def queens(n: Int) =
 def placeOueens(k: Int): Set[List[Int]] =
    if k == 0 then Set(List())
    else
      for
        queens <- placeQueens(k - 1)</pre>
        col <- 0 until n
        if isSafe(col, queens)
      yield col :: queens
 placeQueens(n)
```

#### Write a function

```
def isSafe(col: Int, queens: List[Int]): Boolean
```

which tests if a queen placed in an indicated column col is secure amongst the other placed queens.

It is assumed that the new queen is placed in the next available row after the other placed queens (in other words: in row queens.length).

```
def isSafe(col: Int, queens: List[Int]): Boolean =
 !checks(col, 1, queens)
```

where the checks predicate takes in an additional second parameter delta the distance in rows between the first row of queens and the row where the current queen is placed. checks is defined as follows:

# Maps

Principles of Functional Programming

### Map

Another fundamental collection type is the *map*.

A map of type Map[Key, Value] is a data structure that associates keys of type Key with values of type Value.

### Examples:

```
val romanNumerals = Map("I" -> 1, "V" -> 5, "X" -> 10)
val capitalOfCountry = Map("US" -> "Washington", "Switzerland" -> "Bern")
```

### Maps are Iterables

Class Map[Key, Value] extends the collection type Iterable[(Key, Value)].

Therefore, maps support the same collection operations as other iterables do. Example:

Note that maps extend iterables of key/value pairs.

In fact, the syntax key -> value is just an alternative way to write the pair (key, value). (-> implemented as an extension method in Predef).

### Maps are Functions

Class Map[Key, Value] also extends the function type Key => Value, so maps can be used everywhere functions can.

In particular, maps can be applied to key arguments:

```
capitalOfCountry("US") // "Washington"
```

### Querying Map

Applying a map to a non-existing key gives an error:

```
capitalOfCountry("Andorra")
// java.util.NoSuchElementException: key not found: Andorra
```

To query a map without knowing beforehand whether it contains a given key, you can use the get operation:

```
capitalOfCountry.get("US") // Some("Washington")
capitalOfCountry.get("Andorra") // None
```

The result of a get operation is an Option value.

### The Option Type

The Option type is defined as:

```
trait Option[+A]

case class Some[+A](value: A) extends Option[A]
object None extends Option[Nothing]
```

The expression map.get(key) returns

- ▶ None if map does not contain the given key,
- ► Some(x) if map associates the given key with the value x.

### **Decomposing Option**

Since options are defined as case classes, they can be decomposed using pattern matching:

```
def showCapital(country: String) = capitalOfCountry.get(country) match
  case Some(capital) => capital
  case None => "missing data"

showCapital("US") // "Washington"
showCapital("Andorra") // "missing data"
```

Options also support quite a few operations of the other collections.

I invite you to try them out!

### **Updating Maps**

Functional updates of a map are done with the + and ++ operations:

```
m + (k -> v) The map that takes key 'k' to value 'v'
and is otherwise equal to 'm'
m ++ kvs The map 'm' updated via '+' with all key/value
pairs in 'kvs'
```

These operations are purely functional. For instance,

### Sorted and GroupBy

Two useful operations known from SQL queries are groupBy and orderBy.

orderBy on a collection can be expressed using sortWith and sorted.

```
val fruit = List("apple", "pear", "orange", "pineapple")
fruit.sortWith(_.length < _.length) // List("pear", "apple", "orange", "pineap
fruit.sorted // List("apple", "orange", "pear", "pineap</pre>
```

groupBy is available on Scala collections. It partitions a collection into a map of collections according to a discriminator function f.

#### Example:

### Map Example

A polynomial can be seen as a map from exponents to coefficients.

For instance,  $x^3 - 2x + 5$  can be represented with the map.

$$Map(0 \rightarrow 5, 1 \rightarrow -2, 3 \rightarrow 1)$$

Based on this observation, let's design a class Polynom that represents polynomials as maps.

### **Default Values**

So far, maps were *partial functions*: Applying a map to a key value in map(key) could lead to an exception, if the key was not stored in the map.

There is an operation withDefaultValue that turns a map into a total function:

### Variable Length Argument Lists

It's quite inconvenient to have to write

```
Polynom(Map(1 -> 2.0, 3 -> 4.0, 5 -> 6.2))
```

Can one do without the Map(...)?

Problem: The number of key -> value pairs passed to Map can vary.

### Variable Length Argument Lists

It's quite inconvenient to have to write

```
Polynom(Map(1 -> 2.0, 3 -> 4.0, 5 -> 6.2))
```

Can one do without the Map(...)?

Problem: The number of key -> value pairs passed to Map can vary.

We can accommodate this pattern using a repeated parameter.

```
def Polynom(bindings: (Int, Double)*) =
   Polynom(bindings.toMap.withDefaultValue(0))
Polynom(1 -> 2.0, 3 -> 4.0, 5 -> 6.2)
```

Inside the Polynom function, bindings is seen as a Seq[(Int, Double)].

### Final Implementation of Polynom

```
class Polynom(nonZeroTerms: Map[Int. Double]):
  def this(bindings: (Int, Double)*) = this(bindings.toMap)
  def terms = nonZeroTerms.withDefaultValue(0.0)
  def + (other: Polynom) =
   Polynom(terms ++ other.terms.map((exp, coeff) => (exp, terms(exp) + coeff)))
  override def toString =
   val termStrings =
     for (exp, coeff) <- terms.toList.sorted.reverse</pre>
      vield
        val exponent = if exp == 0 then "" else s"x^$exp"
        s"$coeff$exponent"
    if terms.isEmpty then "0" else termStrings.mkString(" + ")
```

The + operation on Polynom used map concatenation with ++. Design another version of + in terms of foldLeft:

```
def + (other: Polynom) =
   Polynom(other.terms.foldLeft(???)(addTerm))

def addTerm(terms: Map[Int, Double], term: (Int, Double)) =
   ???
```

```
O The version using ++
O The version using foldLeft
```

The + operation on Polynom used map concatenation with ++. Design another version of + in terms of foldLeft:

```
def + (other: Polynom) =
   Polynom(other.terms.foldLeft(???)(addTerm))

def addTerm(terms: Map[Int, Double], term: (Int, Double)) =
   ???
```

```
O The version using ++
O The version using foldLeft
```

The + operation on Polynom used map concatenation with ++. Design another version of + in terms of foldLeft:

```
def + (other: Polynom) =
   Polynom(other.terms.foldLeft(terms)(addTerm))

def addTerm(terms: Map[Int, Double], term: (Int, Double)) =
   val (exp, coeff) = term
   terms + (exp, coeff + terms(exp))
```

```
O The version using ++
X The version using foldLeft
```

The + operation on Polynom used map concatenation with ++. Design another version of + in terms of foldLeft:

```
def + (other: Polynom) =
   Polynom(other.terms.foldLeft(terms)(addTerm))

def addTerm(terms: Map[Int, Double], term: (Int, Double)) =
   val (exp, coeff) = term
   terms + (exp, coeff + terms(exp))
```

```
O The version using ++
X The version using foldLeft
```

# Putting the Pieces Together

Principles of Functional Programming

### Task

Once upon a time, before smartphones, phone keys had mnemonics assigned to them.

```
val mnemonics = Map(
    '2' -> "ABC", '3' -> "DEF", '4' -> "GHI", '5' -> "JKL",
    '6' -> "MNO", '7' -> "PQRS", '8' -> "TUV", '9' -> "WXYZ")
```

Assume you are given a dictionary words as a list of words.

Design a method encode such that

```
encode(phoneNumber)
```

produces all phrases of words that can serve as mnemonics for the phone number

**Example**: The phone number "7225247386" should have the mnemonic Scala is fun as one element of the set of solution phrases.

#### Outline

```
class Coder(words: List[String]):
  val mnemonics = Map(...)
  /** Maps a letter to the digit it represents */
  private val charCode: Map[Char, Char] = ???
  /** Maps a word to the digit string it can represent */
  private def wordCode(word: String): String = ???
  /** Maps a digit string to all words in the dictionary that represent it */
  private val wordsForNum: Map[String. List[String]] = ???
  /** All ways to encode a number as a list of words */
  def encode(number: String): Set[List[String]] = ???
```

```
class Coder(words: List[String]):
    val mnemonics = Map(...)

/** Maps a letter to the digit it represents */
    private val charCode: Map[Char, Char] =
```

```
class Coder(words: List[String]):
    val mnemonics = Map(...)

/** Maps a letter to the digit it represents */
private val charCode: Map[Char, Char] =
    for
        (digit, str) <- mnemonics
        ltr <- str
        yield ltr -> digit
```

```
class Coder(words: List[String]):
    val mnemonics = Map(...)

/** Maps a letter to the digit it represents */
private val charCode: Map[Char, Char] =
    for (digit, str) <- mnemonics; ltr <- str yield ltr -> digit

/** Maps a word to the digit string it can represent */
private def wordCode(word: String): String =
```

```
class Coder(words: List[String]):
    val mnemonics = Map(...)

/** Maps a letter to the digit it represents */
private val charCode: Map[Char, Char] =
    for (digit, str) <- mnemonics; ltr <- str yield ltr -> digit

/** Maps a word to the digit string it can represent */
private def wordCode(word: String): String = word.toUpperCase.map(charCode)
```

```
class Coder(words: List[String]):
  val mnemonics = Map(...)
  /** Maps a letter to the digit it represents */
  private val charCode: Map[Char, Char] =
    for (digit, str) <- mnemonics: ltr <- str vield ltr -> digit
  /** Maps a word to the digit string it can represent */
  private def wordCode(word: String): String = word.toUpperCase.map(charCode)
  /** Maps a digit string to all words in the dictionary that represent it */
  private val wordsForNum: Map[String, List[String]] =
```

```
class Coder(words: List[String]):
  val mnemonics = Map(...)
  /** Maps a letter to the digit it represents */
  private val charCode: Map[Char, Char] =
    for (digit, str) <- mnemonics: ltr <- str vield ltr -> digit
  /** Maps a word to the digit string it can represent */
  private def wordCode(word: String): String = word.toUpperCase.map(charCode)
  /** Maps a digit string to all words in the dictionary that represent it */
  private val wordsForNum: Map[String, List[String]] =
    words.groupBv(wordCode).withDefaultValue(Nil)
```

```
/** All ways to encode a number as a list of words */
def encode(number: String): Set[List[String]] =
```

Idea: use divide and conquer

```
/** All ways to encode a number as a list of words */
def encode(number: String): Set[List[String]] =
   if number.isEmpty then ???
   else ???
```

```
/** All ways to encode a number as a list of words */
def encode(number: String): Set[List[String]] =
   if number.isEmpty then Set(Nil)
   else ???
```

```
/** All ways to encode a number as a list of words */
def encode(number: String): Set[List[String]] =
   if number.isEmpty then Set(Nil)
   else
     for
       splitPoint <- (1 to number.length).toSet
       word <- ???
     rest <- ???
   yield word :: rest</pre>
```

```
/** All ways to encode a number as a list of words */
def encode(number: String): Set[List[String]] =
   if number.isEmpty then Set(Nil)
   else
     for
       splitPoint <- (1 to number.length).toSet
       word <- wordsForNum(number.take(splitPoint))
       rest <- ???
     yield word :: rest</pre>
```

```
/** All ways to encode a number as a list of words */
def encode(number: String): Set[List[String]] =
   if number.isEmpty then Set(Nil)
   else
     for
       splitPoint <- (1 to number.length).toSet
       word <- wordsForNum(number.take(splitPoint))
       rest <- encode(number.drop(splitPoint))
   yield word :: rest</pre>
```

### Testing It

#### A test program:

```
@main def code(number: String) =
  val coder = Coder(List(
    "Scala", "Python", "Ruby", "C",
    "rocks", "socks", "sucks", "works", "pack"))
  coder.encode(number).map(_.mkString(" "))

A sample run:
> scala code "7225276257"
HashSet(Scala rocks, pack C rocks, pack C socks, Scala socks)
```

### Background

This example was taken from:

Lutz Prechelt: An Empirical Comparison of Seven Programming Languages. IEEE Computer 33(10): 23-29 (2000)

Tested with Tcl, Python, Perl, Rexx, Java, C++, C.

Code size medians:

- ▶ 100 loc for scripting languages
- ▶ 200-300 loc for the others

#### **Benefits**

#### Scala's immutable collections are:

- easy to use: few steps to do the job.
- concise: one word replaces a whole loop.
- safe: type checker is really good at catching errors.
- fast: collection ops are tuned, can be parallelized.
- universal: one vocabulary to work on all kinds of collections.

This makes them an attractive tool for software development