

Programming Elixir ≥ 1.6

Functional

- |> Concurrent
- |> Pragmatic
- |> Fun

Dave Thomas

Foreword by
José Valim,
Creator of Elixir



Programming Elixir ≥ 1.6

Functional |> Concurrent |> Pragmatic |> Fun

Dave Thomas

The Pragmatic Bookshelf

Raleigh, North Carolina



Many of the designations used by manufacturers and sellers to distinguish their products are claimed as trademarks. Where those designations appear in this book, and The Pragmatic Programmers, LLC was aware of a trademark claim, the designations have been printed in initial capital letters or in all capitals. The Pragmatic Starter Kit, The Pragmatic Programmer, Pragmatic Programming, Pragmatic Bookshelf, PragProg and the linking *g* device are trademarks of The Pragmatic Programmers, LLC.

Every precaution was taken in the preparation of this book. However, the publisher assumes no responsibility for errors or omissions, or for damages that may result from the use of information (including program listings) contained herein.

Our Pragmatic books, screencasts, and audio books can help you and your team create better software and have more fun. Visit us at <https://pragprog.com>.

The team that produced this book includes:

Publisher: Andy Hunt

VP of Operations: Janet Furlow

Managing Editor: Brian MacDonald

Supervising Editor: Jacquelyn Carter

Copy Editor: Candace Cunningham

Indexing: Potomac Indexing, LLC

Layout: Gilson Graphics

For sales, volume licensing, and support, please contact support@pragprog.com.

For international rights, please contact rights@pragprog.com.

Copyright © 2018 The Pragmatic Programmers, LLC.

All rights reserved.

No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form, or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior consent of the publisher.

Printed in the United States of America.

ISBN-13: 978-1-68050-299-2

Encoded using the finest acid-free high-entropy binary digits.

Book version: P1.0—May 2018

Contents

Foreword	ix
A Vain Attempt at a Justification, Take Two	xi
1. Take the Red Pill	1
Programming Should Be About Transforming Data	1
Installing Elixir	3
Running Elixir	4
Suggestions for Reading the Book	10
Exercises	10
Think Different(ly)	11

Part I — Conventional Programming

2. Pattern Matching	15
Assignment: I Do Not Think It Means What You Think	15
It Means.	15
More Complex Matches	16
Ignoring a Value with _ (Underscore)	18
Variables Bind Once (per Match)	18
Another Way of Looking at the Equals Sign	20
3. Immutability	21
You Already Have (Some) Immutable Data	21
Immutable Data Is Known Data	22
Performance Implications of Immutability	23
Coding with Immutable Data	24
4. Elixir Basics	25
Built-in Types	25
Value Types	26

System Types	28
Collection Types	28
Maps	31
Binaries	32
Dates and Times	33
Names, Source Files, Conventions, Operators, and So On	34
Variable Scope	36
End of the Basics	40
5. Anonymous Functions	41
Functions and Pattern Matching	42
One Function, Multiple Bodies	43
Functions Can Return Functions	45
Passing Functions as Arguments	47
Functions Are the Core	51
6. Modules and Named Functions	53
Compiling a Module	53
The Function's Body Is a Block	54
Function Calls and Pattern Matching	55
Guard Clauses	58
Default Parameters	60
Private Functions	63
The Amazing Pipe Operator: >	63
Modules	65
Module Attributes	67
Module Names: Elixir, Erlang, and Atoms	68
Calling a Function in an Erlang Library	69
Finding Libraries	69
7. Lists and Recursion	71
Heads and Tails	71
Using Head and Tail to Process a List	72
Using Head and Tail to Build a List	74
Creating a Map Function	75
Reducing a List to a Single Value	76
More Complex List Patterns	78
The List Module in Action	81
Get Friendly with Lists	82

8.	Maps, Keyword Lists, Sets, and Structs	83
	How to Choose Between Maps, Structs, and Keyword Lists	83
	Keyword Lists	84
	Maps	84
	Pattern Matching and Updating Maps	85
	Updating a Map	87
	Structs	88
	Nested Dictionary Structures	89
	Sets	95
	With Great Power Comes Great Temptation	95
9.	An Aside—What Are Types?	97
10.	Processing Collections—Enum and Stream	99
	Enum—Processing Collections	99
	Streams—Lazy Enumerables	103
	The Collectable Protocol	110
	Comprehensions	111
	Moving Past Divinity	114
11.	Strings and Binaries	117
	String Literals	117
	The Name “strings”	120
	Single-Quoted Strings—Lists of Character Codes	121
	Binaries	123
	Double-Quoted Strings Are Binaries	124
	Binaries and Pattern Matching	130
	Familiar Yet Strange	132
12.	Control Flow	133
	if and unless	133
	cond	134
	case	137
	Raising Exceptions	138
	Designing with Exceptions	138
	Doing More with Less	139
13.	Organizing a Project	141
	The Project: Fetch Issues from GitHub	141
	Step 1: Use Mix to Create Our New Project	142
	Transformation: Parse the Command Line	145
	Write Some Basic Tests	146

Refactor: Big Function Alert	149
Transformation: Fetch from GitHub	149
Step 2: Use Libraries	151
Transformation: Convert Response	156
Transformation: Sort Data	158
Transformation: Take First n Items	159
Transformation: Format the Table	160
Step 3: Make a Command-Line Executable	163
Step 4: Add Some Logging	164
Step 5: Create Project Documentation	166
Coding by Transforming Data	167
 14. Tooling	169
Debugging with IEx	169
Testing	173
Code Dependencies	186
Server Monitoring	187
Source-Code Formatting	190
Inevitably, There's More	193
 Part II — Concurrent Programming	
 15. Working with Multiple Processes	197
A Simple Process	198
Process Overhead	203
When Processes Die	206
Parallel Map—The “Hello, World” of Erlang	210
A Fibonacci Server	211
Agents—A Teaser	215
Thinking in Processes	216
 16. Nodes—The Key to Distributing Services	219
Naming Nodes	219
Naming Your Processes	223
Input, Output, PIDs, and Nodes	226
Nodes Are the Basis of Distribution	228
 17. OTP: Servers	229
Some OTP Definitions	229
An OTP Server	230
GenServer Callbacks	238

Naming a Process	240
Tidying Up the Interface	240
Making Our Server into a Component	242
18. OTP: Supervisors	247
Supervisors and Workers	247
Worker Restart Options	254
Supervisors Are the Heart of Reliability	255
19. A More Complex Example	257
Introduction to Duper	257
The Duper Application	262
But Does It Work?	272
Planning Your Elixir Application	274
Next Up	275
20. OTP: Applications	277
This Is Not Your Father's Application	277
The Application Specification File	278
Turning Our Sequence Program into an OTP Application	278
Supervision Is the Basis of Reliability	281
Releasing Your Code	282
Distillery—The Elixir Release Manager	282
OTP Is Big—Unbelievably Big	292
21. Tasks and Agents	293
Tasks	293
Agents	295
A Bigger Example	297
Agents and Tasks, or GenServer?	300

Part III — More Advanced Elixir

22. Macros and Code Evaluation	303
Implementing an if Statement	303
Macros Inject Code	304
Using the Representation as Code	307
Using Bindings to Inject Values	312
Macros Are Hygienic	313
Other Ways to Run Code Fragments	314
Macros and Operators	315

Digging Deeper	316
Digging Ridiculously Deep	316
23. Linking Modules: Behavio(u)rs and use	319
Behaviours	319
use and __using	322
Putting It Together—Tracing Method Calls	322
Use use	326
24. Protocols—Polymorphic Functions	329
Defining a Protocol	329
Implementing a Protocol	330
The Available Types	331
Protocols and Structs	332
Built-in Protocols	333
Protocols Are Polymorphism	345
25. More Cool Stuff	347
Writing Your Own Sigils	347
Multi-app Umbrella Projects	351
But Wait! There's More!	354
A1. Exceptions: raise and try, catch and throw	355
Raising an Exception	355
catch, exit, and throw	357
Defining Your Own Exceptions	358
Now Ignore This Appendix	359
A2. Type Specifications and Type Checking	361
When Specifications Are Used	361
Specifying a Type	362
Defining New Types	364
Specs for Functions and Callbacks	365
Using Dialyzer	366
Bibliography	373
Index	375

Foreword

I have always been fascinated with how changes in hardware affect how we write software.

A couple of decades ago, memory was a very limited resource. It made sense back then for our software to take hold of some piece of memory and mutate it as necessary. However, allocating this memory and cleaning up after we no longer needed it was a very error-prone task. Some memory was never freed; sometimes memory was allocated over another structure, leading to faults. At the time, garbage collection was a known technique, but we needed faster CPUs in order to use it in our daily software and free ourselves from manual memory management. That has happened—most of our languages are now garbage-collected.

Today, a similar phenomenon is happening. Our CPUs are not getting any faster. Instead, our computers get more and more cores. This means new software needs to use as many cores as it can if it is to maximize its use of the machine. This conflicts directly with how we currently write software.

In fact, mutating our memory state actually slows down our software when many cores are involved. If you have four cores trying to access and manipulate the same piece of memory, they can trip over each other. This potentially corrupts memory unless some kind of synchronization is applied.

I quickly learned that applying this synchronization is manual, error prone, and tiresome, and it hurts performance. I suddenly realized that's not how I wanted to spend time writing software in the next years of my career, and I set out to study new languages and technologies.

It was on this quest that I fell in love with the Erlang virtual machine and ecosystem.

In the Erlang VM, all code runs in tiny concurrent processes, each with its own state. Processes talk to each other via messages. And since all communication happens by message-passing, exchanging messages between different

machines on the same network is handled transparently by the VM, making it a perfect environment for building distributed software!

However, I felt there was still a gap in the Erlang ecosystem. I missed first-class support for some of the features I find necessary in my daily work—things such as metaprogramming, polymorphism, and first-class tooling. From this need, Elixir was born.

Elixir is a pragmatic approach to functional programming. It values its functional foundations and it focuses on developer productivity. Concurrency is the backbone of Elixir software. As garbage collection once freed developers from the shackles of memory management, Elixir is here to free you from antiquated concurrency mechanisms and bring you joy when writing concurrent code.

A functional programming language lets us think in terms of functions that transform data. This transformation never mutates data. Instead, each application of a function potentially creates a new, fresh version of the data. This greatly reduces the need for data-synchronization mechanisms.

Elixir also empowers developers by providing macros. Elixir code is nothing more than data, and therefore can be manipulated via macros like any other value in the language.

Finally, object-oriented programmers will find many of the mechanisms they consider essential to writing good software, such as polymorphism, in Elixir.

All this is powered by the Erlang VM, a 20-year-old virtual machine built from scratch to support robust, concurrent, and distributed software. Elixir and the Erlang VM are going to change how you write software and make you ready to tackle the upcoming years in programming.

José Valim

Creator of Elixir

Tenczynek, Poland, October 2014

A Vain Attempt at a Justification, Take Two

I'm a language nut. I love trying languages out, and I love thinking about their design and implementation. (I know; it's sad.)

I came across Ruby in 1998 because I was an avid reader of comp.lang.misc (ask your parents). I downloaded it, compiled it, and fell in love. As with any time you fall in love, it's difficult to explain why. It just worked the way I work, and it had enough depth to keep me interested.

Fast-forward 15 years. All that time I'd been looking for something new that gave me the same feeling.

I came across Elixir a while back, but for some reason never got sucked in. But a few months before starting on the first edition of this book, I was chatting with Corey Haines. I was bemoaning the fact that I wanted a way to show people functional programming concepts without the academic trappings those books seem to attract. He told me to look again at Elixir. I did, and I felt the same way I felt when I first saw Ruby.

So now I'm dangerous. I want other people to see just how great this is. I want to evangelize. So I write a book. But I don't want to write another 900-page Pickaxe book. I want this book to be short and exciting. So I'm not going into all the detail, listing all the syntax, all the library functions, all the OTP options, or....

Instead, I want to give you an idea of the power and beauty of this programming model. I want to inspire you to get involved, and then point to the online resources that will fill in the gaps.

But mostly, I want you to have fun.

Fast-forward three years. Elixir has moved on. Phoenix, its connectivity framework, introduced a whole new set of developers to the joys of a functional approach. The Nerves project makes it easy to write embedded Elixir code on Linux-based microcontrollers. The Elixir base has grown—there are international, national, and regional conferences. Job ads ask for Elixir developers.

I've been moving on, too. But I'm still using Elixir daily. I just completed my second year as an adjunct professor at Southern Methodist University, corrupting the programmers of tomorrow with the temptations of Elixir. I've written an online Elixir course.¹

And now I'm revving this book. To be honest, I don't really have to: Elixir 1.6 is not so different from 1.3 that the older book would not be useful. But my own thinking about Elixir has matured. I now do some things differently. And I'd like to share these things with you.

Acknowledgments

It seems to be a common thread—the languages I fall in love with are created by people who are both clever and extremely nice. José Valim, the creator of Elixir, takes both of these adjectives to a new level. I owe him a massive thank-you for giving me so much fun over the last 18 months. Along with him, the whole Elixir core team has done an amazing job of cranking out an entire ecosystem that feels way more mature than its years. Thank you, all.

A conversation with Corey Haines reignited my interest in Elixir—thank you, Corey, for good evenings, some interesting times in Bangalore, and the inspiration.

Bruce Tate is always an interesting sounding board, and his comments on early drafts of the book made a big difference. And I've been blessed with an incredible number of active and insightful beta readers who have made literally hundreds of suggestions for improvements. Thank you, all.

A big tip of the hat to Jessica Kerr, Anthony Eden, and Chad Fowler for letting me steal their tweets.

Kim Shrier seems to have been involved with my writing since before I started writing. Thanks, Kim, for another set of perceptive and detailed critiques.

Candace Cunningham again amazed me with her detailed copy editing: it's rare to find someone who can correct both your grammar and your code. The crew at Potomac did their customary stellar job of indexing.

Dave Thomas

dave@pragdave.me

Dallas, TX, April 2018

1. <https://codestool.coding-gnome.com>

Take the Red Pill

The Elixir programming language wraps functional programming with immutable state and an actor-based approach to concurrency in a tidy, modern syntax. And it runs on the industrial-strength, high-performance, distributed Erlang VM. But what does all that mean?

It means you can stop worrying about many of the difficult things that currently consume your time. You no longer have to think too hard about protecting your data consistency in a multithreaded environment. You worry less about scaling your applications. And, most importantly, you can think about programming in a different way.

Programming Should Be About Transforming Data

If you come from an object-oriented world, then you are used to thinking in terms of classes and their instances. A class defines behavior, and objects hold state. Developers spend time coming up with intricate hierarchies of classes that try to model their problem, much as Victorian scientists created taxonomies of butterflies.

When we code with objects, we're thinking about state. Much of our time is spent calling methods in objects and passing them other objects. Based on these calls, objects update their own state, and possibly the state of other objects. In this world, the class is king—it defines what each instance can do, and it implicitly controls the state of the data its instances hold. Our goal is data-hiding.

But that's not the real world. In the real world, we don't want to model abstract hierarchies (because in reality there aren't that many true hierarchies). We want to get things done, not maintain state.

Right now, for instance, I'm taking empty computer files and transforming them into files containing text. Soon I'll transform those files into a format

you can read. A web server somewhere will transform your request to download the book into an HTTP response containing the content.

I don't want to hide data. I want to transform it.

Combine Transformations with Pipelines

Unix users are accustomed to the philosophy of small, focused command-line tools that can be combined in arbitrary ways. Each tool takes an input, transforms it, and writes the result in a format the next tool (or a human) can use.

This philosophy is incredibly flexible and leads to fantastic reuse. The Unix utilities can be combined in ways undreamed of by their authors. And each one multiplies the potential of the others.

It's also highly reliable—each small program does one thing well, which makes it easier to test.

There's another benefit. A command pipeline can operate in parallel. If I write

```
$ grep Elixir *.pml | wc -l
```

the word-count program, `wc`, runs at the same time as the `grep` command. Because `wc` consumes `grep`'s output as it is produced, the answer is ready with virtually no delay once `grep` finishes.

Just to give you a taste of this, here's an Elixir function called `pmap`. It takes a collection and a function, and returns the list that results from applying that function to each element of the collection. But...it runs a separate process to do the conversion of each element. Don't worry about the details for now.

```
spawn/pmap1.exs
defmodule Parallel do
  def pmap(collection, func) do
    collection
    |> Enum.map(&(Task.async(&fn -> func.(&1) end)))
    |> Enum.map(&Task.await/1)
  end
end
```

We could run this function to get the squares of the numbers from 1 to 1,000.

```
result = Parallel.pmap 1..1000, &(&1 * &1)
```

And, yes, I just kicked off 1,000 background processes, and I used all the cores and processors on my machine.

The code may not make much sense, but by about halfway through the book, you'll be writing this kind of thing for yourself.

Functions Are Data Transformers

Elixir lets us solve the problem in the same way the Unix shell does. Rather than have command-line utilities, we have functions. And we can string them together as we please. The smaller—more focused—those functions, the more flexibility we have when combining them.

If we want, we can make these functions run in parallel—Elixir has a simple but powerful mechanism for passing messages between them. And these are not your father’s boring old processes or threads—we’re talking about the potential to run millions of them on a single machine and have hundreds of these machines interoperating. Bruce Tate commented on this paragraph with this thought: “Most programmers treat threads and processes as a necessary evil; Elixir developers feel they are an important simplification.” As we get deeper into the book, you’ll start to see what he means.

This idea of transformation lies at the heart of functional programming: a function transforms its inputs into its output. The trigonometric function *sin* is an example—give it $\frac{\pi}{4}$, and you’ll get back 0.7071. An HTML templating system is a function; it takes a template containing placeholders and a list of named values, and produces a completed HTML document.

But this power comes at a price. You’re going to have to unlearn a whole lot of what you *know* about programming. Many of your instincts will be wrong. And this will be frustrating, because you’re going to feel like a total n00b.

Personally, I feel that’s part of the fun. You didn’t learn, say, object-oriented programming overnight. You are unlikely to become a functional programming expert by breakfast, either.

But at some point things will click. You’ll start thinking about problems in a different way, and you’ll find yourself writing code that does amazing things with very little effort on your part. You’ll find yourself writing small chunks of code that can be used over and over, often in unexpected ways (just as *wc* and *grep* can be).

Your view of the world may even change a little as you stop thinking in terms of responsibilities and start thinking in terms of getting things done. And just about everyone can agree that will be fun.

Installing Elixir

This book assumes you’re using at least Elixir 1.6. The most up-to-date instructions for installing Elixir are available at <http://elixir-lang.org/install.html>. Go install it now.

Running Elixir

In this book, I show a terminal session like this:

```
$ echo Hello, World
Hello, World
```

The terminal prompt is the dollar sign, and the stuff you type follows. (On your system, the prompt will likely be different.) Output from the system is shown without highlighting.

iex—Interactive Elixir

To test that your Elixir installation was successful, let's start an interactive Elixir session. At your regular shell prompt, type iex.

```
$ iex
Erlang/OTP 20 [erts-9.1] [source] [64-bit] [smp:4:4] [ds:4:4:10]
          [async-threads:10] [hipe] [kernel-poll:false]h
Interactive Elixir (x.y.z) - press Ctrl+C to exit (type h() ENTER for h
elp)
iex(1)>
```

(The various version numbers you see will likely be different—I won't bother to show them on subsequent examples.)

Once you have an IEx prompt, you can enter Elixir code and you'll see the result. If you enter an expression that continues over more than one line, IEx will prompt for the additional lines with an ellipsis (...).

```
iex(1)> 3 + 4
7
iex(2)> String.reverse "madamimadam"
"madamimadam"
iex(3)> 5 *
...> 6
30
iex(4)>
```

The number in the prompt increments for each complete expression executed. I'll omit the number in most of the examples that follow.

There are several ways of exiting from IEx—none are tidy. The easiest two are typing **Ctrl-C** twice or typing **Ctrl-G** followed by **q** and **Return**. On some systems, you can also use a single **Ctrl-**.

IEx Helpers

IEx has a number of helper functions. Type `h` (followed by `Return`) to get a list:

```
iex> h
          IEx.Helpers
```

Welcome to Interactive Elixir. You are currently seeing the documentation for the module `IEx.Helpers` which provides many helpers to make Elixir's shell more joyful to work with.

This message was triggered by invoking the helper `h()`, usually referred to as `h/0` (since it expects 0 arguments).

You can use the `h/1` function to invoke the documentation for any Elixir module or function:

```
iex> h(Enum)
iex> h(Enum.map)
iex> h(Enum.reverse/1)
```

You can also use the `i/1` function to introspect any value you have in the shell:

```
iex> i("hello")
```

There are many other helpers available, here are some examples:

- `b/1` - prints callbacks info and docs for a given module
- `c/1` - compiles a file into the current directory
- `c/2` - compiles a file to the given path
- `cd/1` - changes the current directory
- `clear/0` - clears the screen
- `exports/1` - shows all exports (functions + macros) in a module
- `flush/0` - flushes all messages sent to the shell
- `h/0` - prints this help message
- `h/1` - prints help for the given module, function or macro
- `i/0` - prints information about the last value
- `i/1` - prints information about the given term
- `ls/0` - lists the contents of the current directory
- `ls/1` - lists the contents of the specified directory
- `open/1` - opens the source for the given module or function in your editor
- `pid/1` - creates a PID from a string
- `pid/3` - creates a PID with the 3 integer arguments passed
- `ref/1` - creates a Reference from a string
- `ref/4` - creates a Reference with the 4 integer arguments passed
- `pwd/0` - prints the current working directory
- `r/1` - recompiles the given module's source file
- `recompile/0` - recompiles the current project
- `runtime_info/0` - prints runtime info (versions, memory usage, stats)
- `v/0` - retrieves the last value from the history
- `v/1` - retrieves the nth value from the history

Help for all of those functions can be consulted directly from the command line using the `h/1` helper itself. Try:

```
iex> h(v/0)
```

To list all IEx helpers available, which is effectively all exports (functions and macros) in the `IEx.Helpers` module:

```
iex> exports(IEx.Helpers)
```

This module also includes helpers for debugging purposes, see `IEx.break!/4` for more information.

To learn more about IEx as a whole, type `h(IEx)`.

In the list of helper functions, the number following the slash is the number of arguments the helper expects.

Probably the most useful is `h` itself. With an argument, it gives you help on Elixir modules or individual functions in a module. This works for any modules loaded into IEx (so when we talk about projects later on, you'll see your own documentation here, too). For example, the `IO` module performs common input/output functions. For help on the module, type `h(IO)` or `h IO`:

```
iex> h IO      # or...
iex> h(IO)
```

Functions handling IO.

Many functions in this module expect an IO device as argument. An IO device must be a PID or an atom representing a process. For convenience, Elixir provides `:stdio` and `:stderr` as shortcuts to Erlang's `:standard_io` and `:standard_error`....

This book frequently uses the `puts` function in the `IO` module, which in its simplest form writes a string to the console. Let's get the documentation:

```
iex> h IO.puts
      def puts(device \\ :stdio, item)
```

Writes `item` to the given device, similar to `write/2`, but adds a newline at the end.

By default, the device is the standard output. It returns `:ok` if it succeeds.

```
## Examples
```

```
IO.puts "Hello, World!"
#=> Hello, World!

IO.puts :stderr, "error"
#=> error
```

Another informative helper is `i`, which displays information about a value:

```
iex> i 123
Term
  123
Data type
  Integer
Reference modules
  Integer
Implemented protocols
  IEx.Info, Inspect, List.Chars, String.Chars

iex> i "cat"
Term
  "cat"
Data type
  BitString
Byte size
  3
Description
  This is a string: a UTF-8 encoded binary. It's printed surrounded by
  "double quotes" because all UTF-8 encoded codepoints in it are printable.
Raw representation
  <<99, 97, 116>>
Reference modules
  String, :binary
Implemented protocols
  IEx.Info, Collectable, Inspect, List.Chars, String.Chars

iex> i %{ name: "Dave", likes: "Elixir" }
Term
  %{likes: "Elixir", name: "Dave"}
Data type
  Map
Reference modules
  Map
Implemented protocols
  IEx.Info, Collectable, Enumerable, Inspect

iex> i Map
Term
  Map
Data type
  Atom
Module bytecode
  bin/../lib/elixir/ebin/Elixir.Map.beam
Source
  lib/elixir/lib/map.ex
Version
  [234303838320399652689109978883853316190]
Compile options
  []
```

```

Description
  Use h(Map) to access its documentation.
  Call Map.module_info() to access metadata.
Raw representation
  :"Elixir.Map"
Reference modules
  Module, Atom
Implemented protocols
  IEx.Info, Inspect, List.Chars, String.Chars

```

IEx is a surprisingly powerful tool. Use it to compile and execute entire projects, log in to remote machines, and access running Elixir applications.

And, if you happen to include the occasional bug in your code (deliberately, of course), IEx has a simple debugger. We'll talk about it when we [look at tooling on page 169](#).

Customizing iex

You can customize IEx by setting options. For example, I like showing the results of evaluations in bright cyan. To find out how to do that, I used this:

```

iex> h IEx.configure
def configure(options)

Configures IEx.

The supported options are: :colors, :inspect, :default_prompt,
:alive_prompt and :history_size.

```

Colors

A keyword list that encapsulates all color settings used by the shell. See documentation for the IO.ANSI module for the list of supported colors and attributes.

The value is a keyword list. List of supported keys:

- `:enabled` - boolean value that allows for switching the coloring on and off
 - `:eval_result` - color for an expression's resulting value
 - `:eval_info` - ... various informational messages
 - `:eval_error` - ... error messages
 - `:stack_app` - ... the app in stack traces
 - `:stack_info` - ... the remaining info in stack traces
 - `:ls_directory` - ... for directory entries (ls helper)
 - `:ls_device` - ... device entries (ls helper)
- ...

I then created a file called `.iex.exs` in my home directory, containing

```
IEx.configure colors: [ eval_result: [ :cyan, :bright ] ]
```

If your IEx session looks messed up (and things such as [33m appear in the output), it's likely your console does not support ANSI escape sequences. In that case, disable colorization using

```
IEx.configure colors: [enabled: false]
```

You can put any Elixir code into .iex.exs.

Compile and Run

Once you tire of writing one-line programs in IEx, you'll want to start putting code into source files. These files will typically have the extension .ex or .exs. This is a convention—files ending in .ex are intended to be compiled into bytecodes and then run, whereas those ending in .exs are more like programs in scripting languages—they are effectively interpreted at the source level. When we come to write tests for our Elixir programs, you'll see that the application files have .ex extensions, whereas the tests have .exs because we don't need to keep compiled versions of the tests lying around.

Let's write the classic first program. Go to a working directory and create a file called hello.exs.

```
intro/hello.exs
IO.puts "Hello, World!"
```

That example shows how most of the code listings in this book are presented. The bar before the code itself shows the path and file name that contains the code. If you're reading an ebook, you'll be able to click on this to download the source file. You can also download all the code by visiting the book's page on our site and clicking on the *Source Code* link.¹



Source file names are written in lowercase with underscores. They will have the extension .ex for programs that you intend to compile into binary form, and .exs for scripts that you want to run without compiling. Our “Hello, World” example is essentially throw-away code, so we used the .exs extension for it.

1. <http://pragprog.com/titles/elixir16>

Having created our source file, let's run it. In the same directory where you created the file, run the `elixir` command:

```
$ elixir hello.exs  
Hello, World!
```

We can also compile and run it inside IEx using the `c` helper:

```
$ iex  
iex> c "hello.exs"  
Hello, World!  
[]  
iex>
```

The `c` helper compiled and executed the source file. The `[]` that follows the output is the return value of the `c` function—if the source file had contained any modules, their names would have been listed here.

The `c` helper compiled the source file as freestanding code. You can also load a file as if you'd typed each line into IEx using `import_file`. In this case, local variables set in the file are available in the IEx session.

As some folks fret over such things, the Elixir convention is to use two-column indentation and spaces (not tabs).

Suggestions for Reading the Book

This book is not a top-to-bottom reference guide to Elixir. Instead, it is intended to give you enough information to know what questions to ask and when to ask them. So approach what follows with a spirit of adventure. Try the code as you read, and don't stop there. Ask yourself questions and then try to answer them, either by coding or by searching the web.

Participate in the book's discussion forums and consider joining the Elixir mailing list.²

You're joining the Elixir community while it is still young. Things are exciting and dynamic, and there are plenty of opportunities to contribute.

Exercises

You'll find exercises sprinkled throughout this book. To become familiar with a language, you need to go beyond reading a book and following along with the examples; you need to write some code yourself. These exercises are

2. <https://elixirforum.com>

starting points so you can do some exploring of the language. Try them out, and don't be afraid to make mistakes.

Think Different(ly)

This is a book about thinking differently—about accepting that some of the things folks say about programming may not be the full story:

- Object orientation is not the only way to design code.
- Functional programming need not be complex or mathematical.
- The bases of programming are not assignments, if statements, and loops.
- Concurrency does not need locks, semaphores, monitors, and the like.
- Processes are not necessarily expensive resources.
- Metaprogramming is not just something tacked onto a language.
- Even if it is work, programming should be fun.

Of course, I'm not saying Elixir is a magic potion (well, technically it is, but you know what I mean). It isn't the *one true way* to write code. But it's different enough from the mainstream that learning it will give you more perspective and will open your mind to new ways of thinking about programming.

So let's start.

And remember to make it fun.

Part I

Conventional Programming

Elixir is great for writing highly parallel, reliable applications.

But to be a great language for parallel programming, a language first has to be great for conventional, sequential programming. In this part of the book we'll cover how to write Elixir code, and explore the idioms and conventions that make Elixir so powerful.

In this chapter, you'll see:

- How pattern matching binds values to variables
- How matching handles structured data
- How `_` (underscore) lets you ignore a match

CHAPTER 2

Pattern Matching

We started the previous chapter by saying Elixir engenders a different way of thinking about programming.

To illustrate this and to lay the foundation for a lot of Elixir programming, let's start reprogramming your brain by looking at one of the cornerstones of all programming languages—assignment.

Assignment:

I Do Not Think It Means What You Think It Means.

Let's use the interactive Elixir shell, IEx, to look at a simple piece of code. (Remember, you start IEx at the command prompt using the `ix` command. You enter Elixir code at its `iex>` prompt, and it displays the resulting values.)

```
iex> a = 1
1
iex> a + 3
4
```

Most programmers would look at this code and say, “OK, we assign 1 to a variable `a`, then on the next line we add 3 to `a`, giving us 4.”

But when it comes to Elixir, they'd be wrong. In Elixir, the equals sign is not an assignment. Instead it's like an assertion. It succeeds if Elixir can find a way of making the left-hand side equal the right-hand side. Elixir calls the `=` symbol the *match operator*.

In this case, the left-hand side is a variable and the right-hand side is an integer literal, so Elixir can make the match true by binding the variable `a` to value 1. You could argue it *is* just an assignment. But let's take it up a notch.

```
iex> a = 1
1
iex> 1 = a
1
iex> 2 = a
** (MatchError) no match of right hand side value: 1
```

Look at the second line of code, `1 = a`. It's another match, and it passes. The variable `a` already has the value `1` (it was set in the first line), so what's on the left of the equals sign is the same as what's on the right, and the match succeeds.

But the third line, `2 = a`, raises an error. You might have expected it to assign `2` to `a`, as that would make the match succeed, but Elixir will only change the value of a variable on the left side of an equals sign—on the right a variable is replaced with its value. This failing line of code is the same as `2 = 1`, which causes the error.

More Complex Matches

First, a little background syntax. Elixir lists can be created using square brackets containing a comma-separated set of values. Here are some lists:

```
[ "Humperdinck", "Buttercup", "Fezzik" ]
[ "milk", "butter", [ "iocane", 12 ] ]
```

Back to the match operator.

```
iex> list = [ 1, 2, 3 ]
[1, 2, 3]
```

To make the match true, Elixir bound the variable `list` to the list `[1, 2, 3]`.

But let's try something else:

```
iex> list = [1, 2, 3]
[1, 2, 3]
iex> [a, b, c] = list
[1, 2, 3]
iex> a
1
iex> b
2
iex> c
3
```

Elixir looks for a way to make the value of the left side the same as the value of the right side. The left side is a list containing three variables, and the right is a list of three values, so the two sides could be made the same by setting the variables to the corresponding values.

Elixir calls this process *pattern matching*. A pattern (the left side) is matched if the values (the right side) have the same structure and if each term in the pattern can be matched to the corresponding term in the values. A literal value in the pattern matches that exact value, and a variable in the pattern matches by taking on the corresponding value.

Let's look at a few more examples.

```
iex> list = [1, 2, [ 3, 4, 5 ] ]
[1, 2, [3, 4, 5]]
iex> [a, b, c] = list
[1, 2, [3, 4, 5]]
iex> a
1
iex> b
2
iex> c
[3, 4, 5]
```

The value on the right side that corresponds to the term `c` on the left side is the sublist `[3,4,5]`; that is the value given to `c` to make the match true.

Let's try a pattern containing some values and variables.

```
iex> list = [1, 2, 3]
[1, 2, 3]
iex> [a, 2, b] = list
[1, 2, 3]
iex> a
1
iex> b
3
```

The literal `2` in the pattern matched the corresponding term on the right, so the match succeeds by setting the values of `a` and `b` to `1` and `3`. But...

```
iex> list = [1, 2, 3]
[1, 2, 3]
iex> [a, 1, b] = list
** (MatchError) no match of right hand side value: [1, 2, 3]
```

Here the 1 (the second term in the list) cannot be matched against the corresponding element on the right side, so no variables are set and the match fails. This gives us a way of matching a list that meets certain criteria—in this case a length of 3, with 1 as its second element.

Your Turn

► [Exercise: PatternMatching-1](#)

Which of the following will match?

- $a = [1, 2, 3]$
- $a = 4$
- $4 = a$
- $[a, b] = [1, 2, 3]$
- $a = [[1, 2, 3]]$
- $[a] = [[1, 2, 3]]$
- $[[a]] = [[1, 2, 3]]$

Ignoring a Value with _ (Underscore)

If we didn't need to capture a value during the match, we could use the special variable `_` (an underscore). This acts like a variable but immediately discards any value given to it—in a pattern match, it is like a wildcard saying, “I'll accept any value here.” The following example matches any three-element list that has a 1 as its first element.

```
iex> [1, _, _] = [1, 2, 3]
[1, 2, 3]
iex> [1, _, _] = [1, "cat", "dog"]
[1, "cat", "dog"]
```

Variables Bind Once (per Match)

Once a variable has been bound to a value in the matching process, it keeps that value for the remainder of the match.

```
iex> [a, a] = [1, 1]
[1, 1]
iex> a
1
iex> [b, b] = [1, 2]
** (MatchError) no match of right hand side value: [1, 2]
```

The first expression in this example succeeds because `a` is initially matched with the first 1 on the right side. The value in `a` is then used in the second term to match the second 1 on the right side.

In the next expression, the first `b` matches the 1. But the second `b` corresponds to the value 2 on the right. `b` cannot have two different values, and so the match fails.

However, a variable can be bound to a new value in a subsequent match, and its current value does not participate in the new match.

```
iex> a = 1
1
iex> [1, a, 3] = [1, 2, 3]
[1, 2, 3]
iex> a
2
```

What if you instead want to force Elixir to use the existing value of the variable in the pattern? Prefix it with `^` (a caret). In Elixir, we call this the *pin operator*.

```
iex> a = 1
1
iex> a = 2
2
iex> ^a = 1
** (MatchError) no match of right hand side value: 1
```

This also works if the variable is a component of a pattern:

```
iex> a = 1
1
iex> [^a, 2, 3] = [ 1, 2, 3 ]      # use existing value of a
[1, 2, 3]
iex> a = 2
2
iex> [ ^a, 2 ] = [ 1, 2 ]
** (MatchError) no match of right hand side value: [1, 2]
```

There's one more important part of pattern matching, which we'll look at when we start [digging deeper into lists on page 71](#).

Your Turn

► [Exercise: PatternMatching-2](#)

Which of the following will match?

- `[a, b, a] = [1, 2, 3]`
- `[a, b, a] = [1, 1, 2]`
- `[a, b, a] = [1, 2, 1]`

► [Exercise: PatternMatching-3](#)

The variable `a` is bound to the value 2. Which of the following will match?

- `[a, b, a] = [1, 2, 3]`
- `[a, b, a] = [1, 1, 2]`
- `a = 1`
- `^a = 2`
- `^a = 1`
- `^a = 2 - a`

Another Way of Looking at the Equals Sign

Elixir's pattern matching is similar to Erlang's (the main difference being that Elixir allows a match to reassign to a variable that was assigned in a prior match, whereas in Erlang a variable can be assigned only once).

Joe Armstrong, Erlang's creator, compares the equals sign in Erlang to that used in algebra. When you write the equation $x = a + 1$, you are not assigning the value of $a + 1$ to x . Instead you're simply asserting that the expressions x and $a + 1$ have the same value. If you know the value of x , you can work out the value of a , and vice versa.

His point is that you had to unlearn the algebraic meaning of `=` when you first came across assignment in imperative programming languages. Now's the time to un-unlearn it.

That's why I talk about pattern matching as the first chapter in this part of the book. It is a core part of Elixir—we'll also use it in conditions, function calls, and function invocation.

But really, I wanted to get you thinking differently about programming languages and to show you that some of your existing assumptions won't work in Elixir.

And speaking of existing assumptions...the next chapter kills another sacred cow. Your current programming language is probably designed to make it easy to change data. After all, that's what programs do, right? Not Elixir. Let's talk about a language in which *all* data is immutable.

Change and decay in all around I see...

► *Henry Francis Lyte, "Abide with Me"*

CHAPTER 3

Immutability

If you listen to functional-programming aficionados, you'll hear people making a big deal about immutability—the fact that in a functional program, data cannot be altered once created.

And, indeed, Elixir enforces immutable data.

But why?

You Already Have (Some) Immutable Data

Forget about Elixir for a moment. Think about your current programming language of choice. Let's imagine you'd written this:

```
count = 99
do_something_with(count)
print(count)
```

You'd expect it to output 99. In fact, you'd be very surprised if it didn't. At your very core, you believe that 99 will always have the value 99.

Now, you could obviously bind a new value to your *variable*, but that doesn't change the fact that the value 99 is still 99.

Imagine programming in a world where you could not rely on that—where some other code, possibly running in parallel with your own, could change the value of 99. In that world, the call to `do_something_with` might kick off code that runs in the background, passing it the value 99 as an argument. And that could change the contents of the parameter it receives. Suddenly, 99 could be 100.

You'd be (rightly) upset. And, what's worse, you'd never really be able to guarantee your code produced the correct results.

Still thinking about your current language, consider this:

```
array = [ 1, 2, 3 ]
do_something_with(array)
print(array)
```

Again, you'd hope the print call would output [1,2,3]. But in most languages, `do_something_with` will receive the array as a reference. If it decides to change the second element or delete the contents entirely, the output will not be what you expect. Now it is harder to look at your code and reason about what it does.

Take this a step further—run multiple threads, all with access to the array. Who knows what state the array will be in if they all start changing it?

All this is because most compound data structures in most programming languages are mutable—you can change all or part of their content. And if pieces of your code do this in parallel, you're in a world of hurt.

By coincidence, Jessica Kerr (@jessitron) tweeted the following on the day I updated this section:

jessitron
GOTO was evil because we asked, "how did I get to this point of execution?"
Mutability leaves us with, "how did I get to this state?"
by pragdave

It's spot-on.

Immutable Data Is Known Data

Elixir sidesteps these problems. In Elixir, all values are immutable. The most complex nested list, the database record—these things behave just like the simplest integer. Their values are all immutable.

In Elixir, once a variable references a list such as [1,2,3], you know it will always reference those same values (until you rebind the variable). And this makes concurrency a lot less frightening.

But what if you *need* to add 100 to each element in [1,2,3]? Elixir does it by producing a copy of the original, containing the new values. The original remains unchanged, and your operation will not affect any other code holding a reference to that original.

This fits in nicely with the idea that programming is about transforming data. When we update [1,2,3], we don't hack it in place. Instead we transform it into something new.

Performance Implications of Immutability

It would be easy to assume that this approach to programming is inefficient. After all, you have to create a new copy of data whenever you update it, and that's going to leave lots of old values around to be garbage-collected. Let's look at these in turn.

Copying Data

Although common sense might dictate that all this copying of data is inefficient, the reverse is true. Because Elixir knows that existing data is immutable, it can reuse it, in part or as a whole, when building new structures.

Consider the following code. (It uses a new operator, [head | tail], which builds a new list with head as its first element and tail as the rest. We'll spend a whole chapter on this when we talk about lists and recursion. For now, just trust.)

```
iex> list1 = [ 3, 2, 1 ]
[3, 2, 1]
iex> list2 = [ 4 | list1 ]
[4, 3, 2, 1]
```

In most languages, list2 would be built by creating a new list containing the values 4, 3, 2, and 1. The three values in list1 would be copied into the tail of list2. And that would be necessary because list1 would be mutable.

But Elixir knows list1 will never change, so it simply constructs a new list with a head of 4 and a tail of list1.

Garbage Collection

The other performance issue with a transformational language is that you quite often end up leaving old values unused when you create new values from them. This leaves a bunch of things using up memory on the heap, so garbage collection has to reclaim them.

Most modern languages have a garbage collector, and developers have grown to be suspicious of them—they can impact performance quite badly.

But the cool thing about Elixir is that you write your code using lots and lots of processes, and each process has its own heap. The data in your application is divvied up between these processes, so each individual heap is much, much smaller than would have been the case if all the data had been in a single heap. As a result, garbage collection runs faster. If a process terminates before its heap becomes full, all its data is discarded—no garbage collection is required.

Coding with Immutable Data

Once you accept the concept, coding with immutable data is surprisingly easy. You just have to remember that any function that transforms data will return a new copy of it. Thus, we never capitalize a string. Instead, we return a capitalized copy of a string.

```
iex> name = "elixir"
"elixir"
iex> cap_name = String.capitalize name
"Elixir"
iex> name
"elixir"
```

If you're coming from an object-oriented language, you may dislike that we write `String.capitalize name` and not `name.capitalize()`. But in object-oriented languages, objects mostly have mutable state. When you make a call such as `name.capitalize()` you have no immediate indication whether you are changing the internal representation of the name, returning a capitalized copy, or both. There's plenty of scope for ambiguity.

In a functional language, we *always* transform data. We never modify it in place. The syntax reminds us of this every time we use it.

That's enough theory. It's time to start learning the language. In the next chapter we'll quickly go over the basic data types and some syntax, and in the following chapters we'll look at functions and modules.

In this chapter, you'll see:

- Five value types
- Two system types
- Four collection types
- Naming, operators, etc.
- The with expression

CHAPTER 4

Elixir Basics

In this chapter we'll look at the types that are baked into Elixir, along with a few other things you need to know to get started. This chapter is deliberately terse—you're a programmer and you know what an integer is, so I'm not going to insult you. Instead, I try to cover the Elixir-specific stuff you need to know.

Built-in Types

Elixir's built-in types are

- Value types:
 - Arbitrary-sized integers
 - Floating-point numbers
 - Atoms
 - Ranges
 - Regular expressions
- System types:
 - PIDs and ports
 - References
- Collection types:
 - Tuples
 - Lists
 - Maps
 - Binaries

Functions are a type too. They have their own chapter, following this one.

You might be surprised that this list doesn't include things such as strings and structures. Elixir has them, but they are built using the basic types from this list. However, they are important. Strings have their own chapter, and

we have a couple of chapters on lists and maps (and other dictionary-like types). The maps chapter also describes the Elixir structure facilities.

Finally, there's some debate about whether regular expressions and ranges are value types. Technically they aren't—under the hood they are just structures. But right now it's convenient to treat them as distinct types.

Value Types

The value types in Elixir represent numbers, names, ranges, and regular expressions.

Integers

Integer literals can be written as decimal (1234), hexadecimal (0xcafe), octal (0o765), and binary (0b1010).

Decimal numbers may contain underscores—these are often used to separate groups of three digits when writing large numbers, so one million could be written 1_000_000 (or perhaps 100_0000 in China and Japan).

There is no fixed limit on the size of integers—their internal representation grows to fit their magnitude.

```
factorial(10000) # => 28462596809170545189...and so on for 35640 more digits...
```

(You'll see how to write a function such as factorial in [Modules and Named Functions, on page 53](#).)

Floating-Point Numbers

Floating-point numbers are written using a decimal point. There must be at least one digit before and after the decimal point. An optional trailing exponent may be given. These are all valid floating-point literals:

```
1.0    0.2456   0.314159e1 314159.0e-5
```

Floats are IEEE 754 double precision, giving them about 16 digits of accuracy and a maximum exponent of around 10^{308} .

Atoms

Atoms are constants that represent something's name. We write them using a leading colon (:), which can be followed by an atom word or an Elixir operator. An atom word is a sequence of UTF-8 letters (including combining marks), digits, underscores, and at signs (@). It may end with an exclamation point or a question mark. You can also create atoms containing arbitrary characters

by enclosing the characters following the colon in double quotes. These are all atoms:

```
:fred :is_binary? :var@2 :<> :===: "func/3"
:"long john silver" :эликсир :mötley_crüe
```

An atom's name is its value. Two atoms with the same name will always compare as being equal, even if they were created by different applications on two computers separated by an ocean.

We'll be using atoms a lot to tag values.

Ranges

Ranges are represented as *start..end*, where *start* and *end* are integers.

Regular Expressions

Elixir has regular-expression literals, written as `~r{regexp}` or `~r{regexp}opts`. Here I show the delimiters for regular-expression literals as { and }, but they are considerably more flexible. You can choose any nonalphanumeric characters as delimiters, as described in the discussion of [sigils on page 118](#). Some people use `~/.../` for nostalgic reasons, but this is less convenient than the bracketed forms, as any forward slashes inside the pattern must be escaped.

Elixir regular expression support is provided by PCRE,¹ which basically provides a Perl 5-compatible syntax for patterns.

You can specify one or more single-character options following a regexp literal. These modify the literal's match behavior or add functionality.

Opt	Meaning
f	Force the pattern to start to match on the first line of a multiline string.
i	Make matches case insensitive.
m	If the string to be matched contains multiple lines, ^ and \$ match the start and end of these lines. \A and \z continue to match the beginning or end of the string.
s	Allow . to match any newline characters.
U	Normally modifiers like * and + are greedy, matching as much as possible. The U modifier makes them <i>ungreedy</i> , matching as little as possible.
u	Enable unicode-specific patterns like \p.
x	Enable extended mode—ignore whitespace and comments (# to end of line).

1. <http://www.pcre.org/>

You manipulate regular expressions with the `Regex` module.

```
iex> Regex.run ~r{[aeiou]}, "caterpillar"
["a"]
iex> Regex.scan ~r{[aeiou]}, "caterpillar"
[["a"], ["e"], ["i"], ["a"]]
iex> Regex.split ~r{[aeiou]}, "caterpillar"
["c", "t", "rp", "ll", "r"]
iex> Regex.replace ~r{[aeiou]}, "caterpillar", "*"
"c*t*rp*ll*r"
```

System Types

These types reflect resources in the underlying Erlang VM.

PIDs and Ports

A PID is a reference to a local or remote process, and a port is a reference to a resource (typically external to the application) that you'll be reading or writing.

The PID of the current process is available by calling `self`. A new PID is created when you spawn a new process. We'll talk about this in Part II.

References

The function `make_ref` creates a globally unique reference; no other reference will be equal to it. We don't use references in this book.

Collection Types

The types we've seen so far are common in other programming languages. Now we're getting into more exotic types, so we'll go into more detail here.

Elixir collections can hold values of any type (including other collections).

Tuples

A tuple is an ordered collection of values. As with all Elixir data structures, once created a tuple cannot be modified.

You write a tuple between braces, separating the elements with commas.

```
{ 1, 2 }      { :ok, 42, "next" }   { :error, :enoent }
```

A typical Elixir tuple has two to four elements—any more and you'll probably want to look at [maps, on page 83](#), or [structs, on page 88](#).

You can use tuples in pattern matching:

```
iex> {status, count, action} = {:ok, 42, "next"}
{:ok, 42, "next"}
iex> status
:ok
iex> count
42
iex> action
"next"
```

It is common for functions to return a tuple where the first element is the atom :ok if there were no errors. Here's an example (assuming you have a file called mix.exs in your current directory):

```
iex> {status, file} = File.open("mix.exs")
{:ok, #PID<0.39.0>}
```

Because the file was successfully opened, the tuple contains an :ok status and a PID, which is how we access the contents.

A common practice is to write matches that assume success:

```
iex> { :ok, file } = File.open("mix.exs")
{:ok, #PID<0.39.0>}
iex> { :ok, file } = File.open("non-existent-file")
** (MatchError) no match of right hand side value: {:error, :enoent}
```

The second open failed, and returned a tuple where the first element was :error. This caused the match to fail, and the error message shows that the second element contains the reason—enoent is Unix-speak for “file does not exist.”

Lists

We've already seen Elixir's list literal syntax, [1,2,3]. This might lead you to think lists are like arrays in other languages, but they are not. (In fact, tuples are the closest Elixir gets to a conventional array.) Instead, a list is effectively a linked data structure.

Definition of a List



A list may either be empty or consist of a head and a tail. The head contains a value and the tail is itself a list.

(If you've used the language Lisp, then this will all seem very familiar.)

As we'll discuss in [Chapter 7, Lists and Recursion, on page 71](#), this recursive definition of a list is the core of much Elixir programming.

Because of their implementation, lists are easy to traverse linearly, but they are expensive to access in random order. (To get to the n^{th} element, you have to scan through $n-1$ previous elements.) It is always cheap to get the head of a list and to extract the tail of a list.

Lists have one other performance characteristic. Remember that we said all Elixir data structures are immutable? That means once a list has been made, it will never be changed. So, if we want to remove the head from a list, leaving just the tail, we never have to copy the list. Instead we can return a pointer to the tail. This is the basis of all the list-traversal tricks we'll cover in [Chapter 7, Lists and Recursion, on page 71](#).

Elixir has some operators that work specifically on lists:

```
iex> [ 1, 2, 3 ] ++ [ 4, 5, 6 ]      # concatenation
[1, 2, 3, 4, 5, 6]
iex> [1, 2, 3, 4] -- [2, 4]          # difference
[1, 3]
iex> 1 in [1,2,3,4]                  # membership
true
iex> "wombat" in [1, 2, 3, 4]
false
```

Keyword Lists

Because we often need simple lists of key/value pairs, Elixir gives us a shortcut. If we write

```
[ name: "Dave", city: "Dallas", likes: "Programming" ]
```

Elixir converts it into a list of two-value tuples:

```
[ {:name, "Dave"}, {:city, "Dallas"}, {:likes, "Programming"} ]
```

Elixir allows us to leave off the square brackets if a keyword list is the last argument in a function call. Thus,

```
DB.save record, [ {:use_transaction, true}, {:logging, "HIGH"} ]
```

can be written more cleanly as

```
DB.save record, use_transaction: true, logging: "HIGH"
```

We can also leave off the brackets if a keyword list appears as the last item in any context where a list of values is expected.

```
iex> [1, fred: 1, dave: 2]
[1, {:fred, 1}, {:dave, 2}]
iex> {1, fred: 1, dave: 2}
{1, [fred: 1, dave: 2]}
```

Maps

A map is a collection of key/value pairs. A map literal looks like this:

```
%{ key => value, key => value }
```

Here are some maps:

```
iex> states = %{ "AL" => "Alabama", "WI" => "Wisconsin" }
%{"AL" => "Alabama", "WI" => "Wisconsin"}

iex> responses = %{ { :error, :enoent } => :fatal, { :error, :busy } => :retry }
%{{:error, :enoent} => :fatal, {:error, :busy} => :retry}

iex> colors = %{ :red => 0xff0000, :green => 0x00ff00, :blue => 0x0000ff }
%{blue: 255, green: 65280, red: 16711680}
```

In the first case the keys are strings, in the second they're tuples, and in the third they're atoms. Although typically all the keys in a map are the same type, that isn't required.

```
iex> %{ "one" => 1, :two => 2, {1,1,1} => 3 }
%{:two => 2, {1, 1, 1} => 3, "one" => 1}
```

If the key is an atom, you can use the same shortcut that you use with keyword lists:

```
iex> colors = %{ red: 0xff0000, green: 0x00ff00, blue: 0x0000ff }
%{blue: 255, green: 65280, red: 16711680}
```

You can also use expressions for the keys in map literals:

```
iex> name = "José Valim"
"José Valim"
iex> %{ String.downcase(name) => name }
%{"josé valim" => "José Valim"}
```

Why do we have both maps and keyword lists? Maps allow only one entry for a particular key, whereas keyword lists allow the key to be repeated. Maps are efficient (particularly as they grow), and they can be used in Elixir's pattern matching, which we discuss in later chapters.

In general, use keyword lists for things such as command-line parameters and passing around options, and use maps when you want an associative array.

Accessing a Map

You extract values from a map using the key. The square-bracket syntax works with all maps:

```
iex> states = %{"AL" => "Alabama", "WI" => "Wisconsin" }
%{"AL" => "Alabama", "WI" => "Wisconsin"}
iex> states["AL"]
"Alabama"
iex> states["TX"]
nil

iex> response_types = %{ { :error, :enoent } => :fatal,
...> { :error, :busy } => :retry }
%{{:error, :busy} => :retry, {:error, :enoent} => :fatal}
iex> response_types[{:error,:busy}]
:retry
```

If the keys are atoms, you can also use a dot notation:

```
iex> colors = %{ red: 0xff0000, green: 0x00ff00, blue: 0x0000ff }
%{blue: 255, green: 65280, red: 16711680}
iex> colors[:red]
16711680
iex> colors.green
65280
```

You'll get a `KeyError` if there's no matching key when you use the dot notation.

Binaries

Sometimes you need to access data as a sequence of bits and bytes. For example, the headers in JPEG and MP3 files contain fields where a single byte may encode two or three separate values.

Elixir supports this with the binary data type. Binary literals are enclosed between `<<` and `>>`.

The basic syntax packs successive integers into bytes:

```
iex> bin = << 1, 2 >>
<<1, 2>>
iex> byte_size bin
2
```

You can add modifiers to control the type and size of each individual field. Here's a single byte that contains three fields of widths 2, 4, and 2 bits. (The example uses some built-in libraries to show the result's binary value.)

```
iex> bin = <<3 :: size(2), 5 :: size(4), 1 :: size(2)>>
<<213>>
iex> :io.format("~.8.2b~n", :binary.bin_to_list(bin))
11010101
:ok
iex> byte_size bin
1
```

Binaries are both important and arcane. They're important because Elixir uses them to represent UTF strings. They're arcane because, at least initially, you're unlikely to use them directly.

Dates and Times

Elixir 1.3 added a calendar module and four new date- and time-related types. Initially, they were little more than data holders, but Elixir 1.5 started to add some functionality to them.

The Calendar module represents the rules used to manipulate dates. The only current implementation is Calendar.ISO, the ISO-8601 representation of the Gregorian calendar.²

The Date type holds a year, a month, a day, and a reference to the ruling calendar.

```
iex> d1 = Date.new(2018, 12, 25)
{:ok, ~D[2018-12-25]}
iex> {:ok, d1} = Date.new(2018, 12, 25)
{:ok, ~D[2018-12-25]}
iex> d2 = ~D[2018-12-25]
~D[2018-12-25]
iex> d1 == d2
true
iex> Date.day_of_week(d1)
2
iex> Date.add(d1, 7)
~D[2019-01-01]
iex> inspect d1, structs: false
"%{__struct__: Date, calendar: Calendar.ISO, day: 25, month: 12, year: 2018}"
```

(The sequences `~D[...]` and `~T[...]` are examples of Elixir's *sigils*. They are a way of constructing literal values. We'll see them again when we look at strings and binaries.)

Elixir can also represent a range of dates:

```
iex> d1 = ~D[2018-01-01]
~D[2018-01-01]
iex> d2 = ~D[2018-06-30]
~D[2018-06-30]
iex> first_half = Date.range(d1, d2)
#DateRange<~D[2018-01-01], ~D[2018-06-30]>
iex> Enum.count(first_half)
181
iex> ~D[2018-03-15] in first_half
true
```

2. <http://www.iso.org/iso/home/standards/iso8601.htm>

The `Time` type contains an hour, a minute, a second, and fractions of a second. The fraction is stored as a tuple containing microseconds and the number of significant digits. (The fact that time values track the number of significant digits in the seconds field means that `~T[12:34:56.0]` is not equal to `~T[12:34:56.00]`.)

```
iex> {:ok, t1} = Time.new(12, 34, 56)
{:ok, ~T[12:34:56]}
iex> t2 = ~T[12:34:56.78]
~T[12:34:56.78]
iex> t1 == t2
false
iex> Time.add(t1, 3600)
~T[13:34:56.000000]
iex> Time.add(t1, 3600, :millisecond)
~T[12:34:59.600000]
```

There are two date/time types: `DateTime` and `NaiveDateTime`. The naive version contains just a date and a time; `DateTime` adds the ability to associate a time zone. The `~N[...]` sigil constructs `NaiveDateTime` structs.

If you are using dates and times in your code, you might want to augment these built-in types with a third-party library, such as Lau Taarnskov’s Calendar library.³

Names, Source Files, Conventions, Operators, and So On

Elixir identifiers must start with a letter or underscore, optionally followed by letters, digits, and underscores. Here *letter* means any UTF-8 *letter character* (optionally with a *combining mark*) and *digit* means a UTF-8 *decimal-digit character*. If you’re using ASCII, this does what you’d expect. The identifiers may end with a question mark or an exclamation mark.

Here are some examples of valid variables:

```
name     josé    _age    まつもと _42  адрес!
```

And some examples of invalid variables:

```
name•    a±2   42
```

Module, record, protocol, and behavior names start with an uppercase letter and are BumpyCase. All other identifiers start with a lowercase letter or an underscore, and by convention use underscores between words. If the first character is an underscore, Elixir doesn’t report a warning if the variable is unused in a pattern match or function parameter list.

3. <https://github.com/lau/calendar>

By convention, source files use two-character indentation for nesting—and they use spaces, not tabs, to achieve this.

Comments start with a hash sign (#) and run to the end of the line.

The Elixir distribution comes with a code formatter, which can be used to convert a source file into the “approved” representation. We’ll look at this [on page 190](#). Most examples in this book follow this format (except where I think it is particularly ugly).

Truth

Elixir has three special values related to Boolean operations: true, false, and nil. nil is treated as false in Boolean contexts.

(A bit of trivia: all three of these values are aliases for atoms of the same name, so true is the same as the atom :true.)

In most contexts, any value other than false or nil is treated as true. We sometimes refer to this as *truthy* as opposed to true.

Operators

Elixir has a very rich set of operators. Here’s a subset we’ll use in this book:

Comparison operators

```
a === b      # strict equality  (so 1 === 1.0 is false)
a !== b     # strict inequality (so 1 !== 1.0 is true)
a == b      # value equality   (so 1 == 1.0 is true)
a != b      # value inequality (so 1 != 1.0 is false)
a > b       # normal comparison
a >= b      #
a < b       # :
a <= b      # :
```

The ordering comparisons in Elixir are less strict than in many languages, as you can compare values of different types. If the types are the same or are compatible (for example, 3 > 2 or 3.0 < 5), the comparison uses natural ordering. Otherwise comparison is based on type according to this rule:

number < atom < reference < function < port < pid < tuple < map < list < binary

Boolean operators

(These operators expect true or false as their first argument.)

```
a or b      # true if a is true; otherwise b
a and b     # false if a is false; otherwise b
not a       # false if a is true; true otherwise
```

Relaxed Boolean operators

These operators take arguments of any type. Any value apart from nil or false is interpreted as true.

```
a || b    # a if a is truthy; otherwise b
a && b    # b if a is truthy; otherwise a
!a        # false if a is truthy; otherwise true
```

Arithmetic operators

```
+ - * / div rem
```

Integer division yields a floating-point result. Use div(a,b) to get an integer.

rem is the *remainder operator*. It is called as a function (rem(11, 3) => 2). It differs from normal modulo operations in that the result will have the same sign as the function's first argument.

Join operators

```
binary1 <>> binary2    # concatenates two binaries (Later we'll
                        # see that binaries include strings.)
list1 ++ list2        # concatenates two lists
list1 -- list2        # removes elements of list 2 from a copy of list 1
```

The in operator

```
a in enum            # tests if a is included in enum (for example,
                      # a list, a range, or a map). For maps, a should
                      # be a {key, value} tuple.
```

Variable Scope

Elixir is lexically scoped. The basic unit of scoping is the function body. Variables defined in a function (including its parameters) are local to that function. In addition, modules define a scope for local variables, but these are accessible only at the top level of that module, and not in functions defined in the module.

Do-block Scope

Most languages let you group together multiple code statements and treat them as a single *code block*. Often languages use braces for this. Here's an example in C:

```
int line_no = 50;
/* .... */
if (line_no == 50) {
    printf("new-page\n");
    line_no = 0;
}
```

Elixir doesn't really have blocks such as these, but it does have ways of grouping expressions together. The most common of these is the do block:

```
line_no = 50
# ...
if (line_no == 50) do
  IO.puts "new-page\f"
  line_no = 0
end
IO.puts line_no
```

However, Elixir thinks this is a risky way to write code. In particular, it's easy to forget to initialize line_no outside the block, but to then rely on it having a value after the block. For that reason, you'll see a warning:

```
$ elixir back_block.ex
warning: the variable "line_no" is unsafe as it has been set inside one of:
case, cond, receive, if, and, or, &&, |||. Please explicitly return the
variable value instead. Here's an example:

  case integer do
    1 -> atom = :one
    2 -> atom = :two
  end

should be written as

atom =
  case integer do
    1 -> :one
    2 -> :two
  end

Unsafe variable found at:
t.ex:10
0
```

The with Expression

The with expression serves double duty. First, it allows you to define a local scope for variables. If you need a couple of temporary variables when calculating something, and you don't want those variables to leak out into the wider scope, use with. Second, it gives you some control over pattern-matching failures. For example, the /etc/passwd file contains lines such as

```
_installassistant:*:25:25:Install Assistant:/var/empty:/usr/bin/false
_lp:*:26:26:Printing Services:/var/spool/cups:/usr/bin/false
_postfix:*:27:27:Postfix Mail Server:/var/spool/postfix:/usr/bin/false
```

The two numbers are the user and group IDs for the given username.

The following code finds the values for the _lp user (and see the sidebar following for some notes on its layout).

```
basic-types/with-scope.exs
content = "Now is the time"

lp = with {:ok, file} = File.open("/etc/passwd"),
          content      = IO.read(file, :all), # note: same name as above
          :ok           = File.close(file),
          [_ , uid, gid] = Regex.run(~r/^_lp:.*?:(\d+):(\d+)/m, content)
do
  "Group: #{gid}, User: #{uid}"
end

IO.puts lp          #=> Group: 26, User: 26
IO.puts content    #=> Now is the time
```

A Code Formatting Comparison

The with listing is an example of code that isn't in the canonical Elixir format. If formatted using the built-in tool, it would look like the following.

```
basic-types/with-scope-fmt.exs
content = "Now is the time"

lp =
  with {:ok, file} = File.open("/etc/passwd"),
       content = IO.read(file, :all),
       :ok = File.close(file),
       [_ , uid, gid] = Regex.run(~r/^_lp:.*?:(\d+):(\d+)/m, content) do
    "Group: #{gid}, User: #{uid}"
  end

# => Group: 26, User: 26
IO.puts(lp)
# => Now is the time
IO.puts(content)
```

I'll let you be the judge of which is clearer.

The with expression lets us work with what are effectively temporary variables as we open the file, read its content, close it, and search for the line we want. The value of the with is the value of its do parameter.

The inner variable content is local to the with, and does not affect the variable in the outer scope.

with and Pattern Matching

In the previous example, the head of the with expression used = for basic pattern matches. If any of these had failed, a MatchError exception would be raised. But perhaps we'd want to handle this case in a more elegant way. That's where the

`<-` operator comes in. If you use `<-` instead of `=` in a `with` expression, it performs a match, but if it fails it returns the value that couldn't be matched.

```
iex> with [a|_] <- [1,2,3], do: a
1
iex> with [a|_] <- nil,      do: a
nil
```

We can use this to let the `with` in the previous example return `nil` if the user can't be found, rather than raising an exception.

```
basic-types/with-match.exs
result = with {:ok, file}    = File.open("/etc/passwd"),
             content       = IO.read(file, :all),
             :ok           = File.close(file),
➤          [_, uid, gid] <- Regex.run(~r/^xxx:.*?:(\d+):(\d+)/, content)
          do
            "Group: #{gid}, User: #{uid}"
          end
IO.puts inspect(result)      #=> nil
```

When we try to match the user `xxx`, `Regex.run` returns `nil`. This causes the match to fail, and the `nil` becomes the value of the `with`.

A Minor Gotcha

Underneath the covers, `with` is treated by Elixir as if it were a call to a function or macro. This means that you cannot write this:

```
mean = with                                # WRONG!
  count = Enum.count(values),
  sum   = Enum.sum(values)
do
  sum/count
end
```

Instead, you can put the first parameter on the same line as the `with`:

```
mean = with count = Enum.count(values),
           sum   = Enum.sum(values)
do
  sum/count
end
```

or use parentheses:

```
mean = with(
  count = Enum.count(values),
  sum   = Enum.sum(values)
do
  sum/count
end)
```

As with all other uses of `do`, you can also use the shortcut:

```
mean = with count = Enum.count(values),  
      sum   = Enum.sum(values),  
      do:   sum/count
```

End of the Basics

We've now covered the low-level ingredients of an Elixir program. In the next two chapters we'll discuss how to create anonymous functions, modules, and named functions.

In this chapter, you'll see:

- Anonymous functions
- Pattern matching and arguments
- Higher-order functions
- Closures
- The & function literal

CHAPTER 5

Anonymous Functions

Elixir is a functional language, so it's no surprise functions are a basic type.

An anonymous function is created using the `fn` keyword.

```
fn
  parameter-list -> body
  parameter-list -> body ...
end
```

Think of `fn...end` as being a bit like the quotes that surround a string literal, except here we're returning a function as a value, not a string. We can pass that function value to other functions. We can also invoke it, passing in arguments.

At its simplest, a function has a parameter list and a body, separated by `->`.

For example, the following defines a function, binding it to the variable `sum`, and then calls it:

```
iex> sum = fn (a, b) -> a + b end
#Function<12.17052888 in :erl_eval.expr/5>
iex> sum.(1, 2)
3
```

The first line of code creates a function that takes two parameters (named `a` and `b`). The implementation of the function follows the `->` arrow (in our case it simply adds the two parameters), and the whole thing is terminated with the keyword `end`. We store the function in the variable `sum`.

On the second line of code, we invoke the function using the syntax `sum.(1,2)`. The dot indicates the function call, and the arguments are passed between parentheses. (You'll have noticed we don't use a dot for named function calls—this is a difference between named and anonymous functions.)

If your function takes no arguments, you still need the parentheses to call it:

```
iex> greet = fn -> IO.puts "Hello" end
#Function<12.17052888 in :erl_eval.expr/5>
iex> greet.()
Hello
:ok
```

You can, however, omit the parentheses in a function definition:

```
iex> f1 = fn a, b -> a * b end
#Function<12.17052888 in :erl_eval.expr/5>
iex> f1.(5,6)
30
iex> f2 = fn -> 99 end
#Function<12.17052888 in :erl_eval.expr/5>
iex> f2.()
99
```

Functions and Pattern Matching

When we call `sum.(2,3)`, it's easy to assume we simply assign 2 to the parameter `a` and 3 to `b`. But that word, *assign*, should ring some bells. Elixir doesn't have assignment. Instead it tries to match values to patterns. (We came across this when we looked at [pattern matching and assignment on page 15](#).)

If we write

```
a = 2
```

then Elixir makes the pattern match by binding `a` to the value 2. And that's exactly what happens when our `sum` function gets called. If we pass 2 and 3 as arguments, and Elixir tries to match these arguments to the parameters `a` and `b` (which it does by giving `a` the value 2 and `b` the value 3), it's the same as when we write

```
{a, b} = {2, 3}
```

This means we can perform more complex pattern matching when we call a function. For example, the following function reverses the order of elements in a two-element tuple:

```
iex> swap = fn { a, b } -> { b, a } end
#Function<12.17052888 in :erl_eval.expr/5>
iex> swap.( { 6, 8 } )
{8, 6}
```

We'll use this pattern-matching capability when we look at functions with multiple implementations in the next section.

Your Turn

► *Exercise: Functions-1*

Go into IEx. Create and run the functions that do the following:

- `list_concat.([{:a, :b}, {:c, :d}]) #=> [:a, :b, :c, :d]`
- `sum(1, 2, 3) #=> 6`
- `pair_tuple_to_list.({ 1234, 5678 }) #=> [1234, 5678]`

One Function, Multiple Bodies

A single function definition lets you define different implementations, depending on the type and contents of the arguments passed. (You cannot select based on the number of arguments—each clause in the function definition must have the same number of parameters.)

At its simplest, we can use pattern matching to select which clause to run. In the example that follows, we know the tuple returned by `File.open` has `:ok` as its first element if the file was opened, so we write a function that displays either the first line of a successfully opened file or a simple error message if the file could not be opened.

```
Line 1 iex> handle_open = fn
2 ...>   {:ok, file} -> "Read data: #{IO.read(file, :line)}"
3 ...>   {_, error} -> "Error: #{file.format_error(error)}"
4 ...> end
5 #Function<12.17052888 in :erl_eval.expr/5>
6 iex> handle_open.(File.open("code/intro/hello.exs"))    # this file exists
7 "Read data: IO.puts \"Hello, World!\\n\""
8 iex> handle_open.(File.open("nonexistent"))              # this one doesn't
9 "Error: no such file or directory"
```

Start by looking inside the function definition. On lines 2 and 3 we define two separate function bodies. Each takes a single tuple as a parameter. The first of them requires that the first term in the tuple is `:ok`. The second line uses the special variable `_` (underscore) to match any other value for the first term.

Now look at line 6. We call our function, passing it the result of calling `File.open` on a file that exists. This means the function will receive the tuple `{:ok,file}`, and this matches the clause on line 2. The corresponding code calls `IO.read` to read the first line of this file.

We then call `handle_open` again, this time with the result of trying to open a file that does not exist. The tuple that is returned (`{:error,:enoent}`) is passed to our function, which looks for a matching clause. It fails on line 2 because the

first term is not `:ok`, but it succeeds on the next line. The code in that clause formats the error as a nice string.

Note a couple of other things in this code. On line 3 we call `:file.format_error`. The `:file` part of this refers to the underlying Erlang `File` module, so we can call its `format_error` function. Contrast this with the call to `File.open` on line 6. Here the `File` part refers to Elixir's built-in module. This is a good example of the underlying environment leaking through into Elixir code. It is good that you can access all the existing Erlang libraries—there are hundreds of years of effort in there just waiting for you to use. But it is also tricky because you have to differentiate between Erlang functions and Elixir functions when you call them.

And finally, this example shows off Elixir's *string interpolation*. Inside a string, the contents of `#{}...` are evaluated and the result is substituted back in.

Working with Larger Code Examples

Our `handle_open` function is getting uncomfortably long to type directly into IEx. One typo, and we'd have to type it all in again.

Instead, let's use our editor to type it into a file in the same directory we were in when we started IEx. Let's call the file `handle_open.exs`.

```
first_steps/handle_open.exs
handle_open = fn
  {:ok, file} -> "First line: #{IO.read(file, :line)}"
  {_, error} -> "Error: #{:file.format_error(error)}"
end
IO.puts handle_open.(File.open("Rakefile"))      # call with a file that exists
IO.puts handle_open.(File.open("nonexistent"))    # and then with one that doesn't
```

Now, inside IEx, type this:

```
c "handle_open.exs"
```

This compiles and runs the code in the given file.

We can do the same thing from the command line (that is, not inside IEx) using this:

```
$ elixir handle_open.exs
```

We used the file extension `.exs` for this example. This is used for code that we want to run directly from a source file (think of the `s` as meaning *script*). For files we want to compile and use later, we'll employ the `.ex` extension.

Your Turn

► *Exercise: Functions-2*

Write a function that takes three arguments. If the first two are zero, return “FizzBuzz.” If the first is zero, return “Fizz.” If the second is zero, return “Buzz.” Otherwise return the third argument. Do not use any language features that we haven’t yet covered in this book.

► *Exercise: Functions-3*

The operator rem(a, b) returns the remainder after dividing a by b. Write a function that takes a single integer (n) and calls the function in the previous exercise, passing it rem(n,3), rem(n,5), and n. Call it seven times with the arguments 10, 11, 12, and so on. You should get “Buzz, 11, Fizz, 13, 14, FizzBuzz, 16.”

(Yes, it’s a FizzBuzz solution with no conditional logic.)¹

Functions Can Return Functions

Here’s some strange code:

```
iex> fun1 = fn -> fn -> "Hello" end end
#Function<12.17052888 in :erl_eval.expr/5>
iex> fun1.()
#Function<12.17052888 in :erl_eval.expr/5>
iex> fun1.().()
"Hello"
```

The strange thing is the first line. It’s hard to read, so let’s spread it out.

```
fun1 = fn ->
  fn ->
    "Hello"
  end
end
```

The variable fun1 is bound to a function. That function takes no parameters, and its body is a second function definition. That second function also takes no parameters, and it evaluates the string "Hello".

When we call the outer function (using fun1()), it returns the inner function. When we call that (fun1().()) the inner function is evaluated and “Hello” is returned.

1. <http://c2.com/cgi/wiki?FizzBuzzTest>

We wouldn't normally write something such as `fun1().()`. But we might call the outer function and bind the result to a separate variable. We might also use parentheses to make the inner function more obvious.

```
iex> fun1 = fn -> (fn -> "Hello" end) end
#Function<12.17052888 in :erl_eval.expr/5>
iex> other = fun1.()
#Function<12.17052888 in :erl_eval.expr/5>
iex> other.()
"Hello"
```

Functions Remember Their Original Environment

Let's take this idea of nesting functions a little further.

```
iex> greeter = fn name -> (fn -> "Hello #{name}" end) end
#Function<12.17052888 in :erl_eval.expr/5>
iex> dave_greeter = greeter.("Dave")
#Function<12.17052888 in :erl_eval.expr/5>
iex> dave_greeter.()
"Hello Dave"
```

Now the outer function has a `name` parameter. Like any parameter, `name` is available for use throughout the body of the function. In this case, we use it inside the string in the inner function.

When we call the outer function, it returns the inner function definition. It has not yet substituted the `name` into the string. But when we call the inner function (`dave_greeter.()`), the substitution takes place and the greeting appears.

But something strange happens here. The inner function uses the outer function's `name` parameter. But by the time `greeter.("Dave")` returns, that outer function has finished executing and the parameter has gone out of scope. And yet when we run the inner function, it uses that parameter's value.

This works because functions in Elixir automatically carry with them the bindings of variables in the scope in which they are defined. In our example, the variable `name` is bound in the scope of the outer function. When the inner function is defined, it inherits this scope and carries the binding of `name` around with it. This is a *closure*—the scope encloses the bindings of its variables, packaging them into something that can be saved and used later.

Let's play with this some more.

Parameterized Functions

In the previous example, the outer function took an argument and the inner one did not. Let's try a different example where both take arguments.

```
iex> add_n = fn n -> (fn other -> n + other end) end
#Function<12.17052888 in :erl_eval.expr/5>
iex> add_two = add_n.(2)
#Function<12.17052888 in :erl_eval.expr/5>
iex> add_five = add_n.(5)
#Function<12.17052888 in :erl_eval.expr/5>
iex> add_two.(3)
5
iex> add_five.(7)
12
```

Here the inner function adds the value of its parameter `other` to the value of the outer function's parameter `n`. Each time we call the outer function, we give it a value for `n` and it returns a function that adds `n` to its own parameter.

Your Turn

► *Exercise: Functions-4*

Write a function prefix that takes a string. It should return a new function that takes a second string. When that second function is called, it will return a string containing the first string, a space, and the second string.

```
iex> mrs = prefix."Mrs"
#Function<erl_eval.6.82930912>
iex> mrs."Smith"
"Mrs Smith"
iex> prefix."Elixir".("Rocks")
"Elixir Rocks"
```

Passing Functions as Arguments

Functions are just values, so we can pass them to other functions.

```
iex> times_2 = fn n -> n * 2 end
#Function<12.17052888 in :erl_eval.expr/5>
iex> apply = fn (fun, value) -> fun.(value) end
#Function<12.17052888 in :erl_eval.expr/5>
iex> apply.(times_2, 6)
12
```

Here, `apply` is a function that takes a second function and a value. It returns the result of invoking that second function with the value as an argument.

We use the ability to pass functions around pretty much everywhere in Elixir code. For example, the built-in `Enum` module has a function called `map`. It takes two arguments: a collection and a function. It returns a list that is the result of applying that function to each element of the collection.

```
iex> list = [1, 3, 5, 7, 9]
[1, 3, 5, 7, 9]
iex> Enum.map list, fn elem -> elem * 2 end
[2, 6, 10, 14, 18]
iex> Enum.map list, fn elem -> elem * elem end
[1, 9, 25, 49, 81]
iex> Enum.map list, fn elem -> elem > 6 end
[false, false, false, true, true]
```

Pinned Values and Function Parameters

When we originally looked at pattern matching, we saw that the pin operator (^) allowed us to use the current value of a variable in a pattern. You can use this with function parameters, too.

```
functions/pin.exs
defmodule Greeter do
  def for(name, greeting) do
    fn
      (^name) -> "#{greeting} #{name}"
      (_)       -> "I don't know you"
    end
  end
end

mr_valim = Greeter.for("José", "Oi!")
IO.puts mr_valim."José"      # => Oi! José
IO.puts mr_valim."Dave"      # => I don't know you
```

Here, the Greeter.for function returns a function with two heads. The first head matches when its first parameter is the value of the name passed to for.

The & Notation

The strategy of creating short helper functions is so common that Elixir provides a shortcut. Let's look at it in use before we explore what's going on.

```
iex> add_one = &(&1 + 1)          # same as add_one = fn (n) -> n + 1 end
#Function<6.17052888 in :erl_eval.expr/5>
iex> add_one.(44)
45
iex> square = &(&1 * &1)
#Function<6.17052888 in :erl_eval.expr/5>
iex> square.(8)
64
iex> speak = &(IO.puts(&1))
&IO.puts/1
iex> speak.("Hello")
Hello
:ok
```

The & operator converts the expression that follows into a function. Inside that expression, the placeholders &1, &2, and so on correspond to the first, second, and subsequent parameters of the function. So &(&1 + &2) will be converted to fn p1, p2 -> p1 + p2 end.

If you think that's clever, take a look at the speak line in the previous code. Normally Elixir would have generated an anonymous function, so &(IO.puts(&1)) would become fn x -> IO.puts(x) end. But Elixir noticed that the body of the anonymous function was simply a call to a named function (the IO function puts) and that the parameters were in the correct order (that is, the first parameter to the anonymous function was the first parameter to the named function, and so on). So Elixir optimized away the anonymous function, replacing it with a direct reference to the function, IO.puts/1.

For this to work, the arguments must be in the correct order:

```
iex> rnd = &(Float.round(&1, &2))
&Float.round/2
iex> rnd = &(Float.round(&2, &1))
#Function<12.17052888 in :erl_eval.expr/5>
```

You might see references to Erlang pop up when you define functions this way. That's because Elixir runs on the Erlang VM. There's more evidence of this if you try something like &abs(&1). Here Elixir maps your use of the abs function directly into the underlying Erlang library, and returns &:erlang.abs/1.

Because [] and {} are operators in Elixir, literal lists and tuples can also be turned into functions. Here's a function that returns a tuple containing the quotient and remainder of dividing two integers:

```
iex> divrem = &{ div(&1,&2), rem(&1,&2) }
#Function<12.17052888 in :erl_eval.expr/5>
iex> divrem.(13, 5)
{2, 3}
```

Finally, the & capture operator works with string (and string-like) literals:

```
iex> s = &"bacon and #{&1}"
#Function<6.99386804/1 in :erl_eval.expr/5>
iex> s.("custard")
"bacon and custard"

iex> match_end = &~r/.*/#{&1}$/
#Function<6.99386804/1 in :erl_eval.expr/5>
iex> "cat" =~ match_end.("t")
true
iex> "cat" =~ match_end.("!")
false
```

There's a second form of the & function capture operator. You can give it the name and arity (number of parameters) of an existing function, and it will return an anonymous function that calls it. The arguments you pass to the anonymous function will in turn be passed to the named function. We've already seen this: when we entered `&(IO.puts(&1))` into iex, it displayed the result as `&IO.puts/1`. In this case, puts is a function in the IO module, and it takes one argument. The Elixir way of naming this is `IO.puts/1`. If we place an & in front of this, we wrap it in a function. Here are some other examples:

```
iex> l = &length/1
&:erlang.length/1
iex> l.([1,3,5,7])
4

iex> len = &Enum.count/1
&Enum.count/1
iex> len.([1,2,3,4])
4

iex> m = &Kernel.min/2    # This is an alias for the Erlang function
&:erlang.min/2
iex> m.(99,88)
88
```

This works with named functions we write, as well (but we haven't covered how to write them yet).

The & shortcut gives us a wonderful way to pass functions to other functions.

```
iex> Enum.map [1,2,3,4], &(&1 + 1)
[2, 3, 4, 5]
iex> Enum.map [1,2,3,4], &(&1 * &1)
[1, 4, 9, 16]
iex> Enum.map [1,2,3,4], &(&1 < 3)
[true, true, false, false]
```

Your Turn

► *Exercise: Functions-5*

Use the & notation to rewrite the following:

- `Enum.map [1,2,3,4], fn x -> x + 2 end`
- `Enum.each [1,2,3,4], fn x -> IO.inspect x end`

Functions Are the Core

At the start of the book, we said the basis of programming is transforming data. Functions are the little engines that perform that transformation. They are at the very heart of Elixir.

So far we've been looking at anonymous functions—although we can bind them to variables, the functions themselves have no names. Elixir also has named functions. In the next chapter we'll cover how to work with them.

In this chapter, you'll see:

- Modules, the basic units of code
- Defining public and private named functions
- Guard clauses
- Module directives and attributes
- Calling functions in Erlang modules

CHAPTER 6

Modules and Named Functions

Once a program grows beyond a couple of lines, you'll want to structure it. Elixir makes this easy. You break your code into *named functions* and organize these functions into *modules*. In fact, in Elixir named functions *must* be written inside modules.

Let's look at a simple example. Navigate to a working directory and create an Elixir source file called `times.exs`.

```
mm/times.exs
defmodule Times do
  def double(n) do
    n * 2
  end
end
```

Here we have a module named `Times`. It contains a single function, `double`. Because our function takes a single argument and because the number of arguments forms part of the way we identify Elixir functions, you'll see this function name written as `double/1`.

Compiling a Module

Let's look at two ways to compile this file and load it into IEx. First, if you're at the command line, you can do this:

```
$ iex times.exs
iex> Times.double(4)
8
```

Give IEx a source file's name, and it compiles and loads the file before it displays a prompt.

If you're already in IEx, you can use the `c` helper to compile your file without returning to the command line.

```
iex> c "times.exs"
[Times]
iex> Times.double(4)
8
iex> Times.double(123)
246
```

The line `c "times.exs"` compiles your source file and loads it into IEx. We then call the `double` function in the `Times` module a couple of times using `Times.double`.

What happens if we make our function fail by passing it a string rather than a number?

```
iex> Times.double("cat")
** (ArithmetricError) bad argument in arithmetic expression
  times.exs:3: Times.double/1
```

An exception (`ArithmetricError`) gets raised, and we see a stack backtrace. The first line tells us what went wrong (we tried to perform arithmetic on a string), and the next line tells us where. But look at what it writes for the name of our function: `Times.double/1`.

In Elixir a named function is identified by both its name and its number of parameters (its *arity*). Our `double` function takes one parameter, so Elixir knows it as `double/1`. If we had another version of `double` that took three parameters, it would be known as `double/3`. These two functions are totally separate as far as Elixir is concerned. But from a human perspective, you'd imagine that if two functions have the same name they are somehow related, even if they have a different number of parameters. For that reason, don't use the same name for two functions that do unrelated things.

The Function's Body Is a Block

The `do...end` block is one way of grouping expressions and passing them to other code. They are used in module and named function definitions, control structures...any place in Elixir where code needs to be handled as an entity.

However, `do...end` is not actually the underlying syntax. The actual syntax looks like this:

```
def double(n), do: n * 2
```

You can pass multiple lines to `do:` by grouping them with parentheses.

```
def greet(greeting, name), do: (
  IO.puts greeting
  IO.puts "How're you doing, #{name}?"
)
```

The `do...end` form is just a lump of syntactic sugar—during compilation it is turned into the `do:` form. (And the `do:` form itself is nothing special; it is simply a term in a keyword list.) Typically people use the `do:` syntax for single-line blocks, and `do...end` for multiline ones.

This means our times example would probably be written as follows:

```
mm/times1.exs
defmodule Times do
  def double(n), do: n * 2
end
```

We could even write it as

```
defmodule Times, do: (def double(n), do: n*2)
```

(but please don't).

Your Turn

- [Exercise: ModulesAndFunctions-1](#)
Extend the Times module with a triple function that multiplies its parameter by three.
- [Exercise: ModulesAndFunctions-2](#)
Run the result in IEx. Use both techniques to compile the file.
- [Exercise: ModulesAndFunctions-3](#)
Add a quadruple function. (Maybe it could call the double function....)

Function Calls and Pattern Matching

In the previous chapter we covered how anonymous functions use pattern matching to bind their parameter list to the passed arguments. The same is true of named functions. The difference is that we write the function multiple times, each time with its own parameter list and body. Although this looks like multiple function definitions, purists will tell you it's multiple clauses of the same definition (and they'd be right).

When you call a named function, Elixir tries to match your arguments with the parameter list of the first definition (clause). If it cannot match them, it tries the next definition of the same function (remember, this must have the same arity) and checks to see if it matches. It continues until it runs out of candidates.

Let's play with this. The factorial of n , written $n!$, is the product of all numbers from 1 to n . $0!$ is defined to be 1.

Another way of expressing this is to say

- $\text{factorial}(0) \rightarrow 1$
 - $\text{factorial}(n) \rightarrow n * \text{factorial}(n-1)$

This is a specification of the concept of *factorial*, but it is also close to an Elixir implementation:

```
mm/factorial1.exs
defmodule Factorial do
  def of(0), do: 1
  def of(n), do: n * of(n-1)
end
```

Here we have two definitions of *the same function*. If we call Factorial.of(2), Elixir matches the 2 against the first function's parameter, 0. This fails, so it tries the second definition, which succeeds when Elixir binds 2 to n. It then evaluates the body of this function, which calls Factorial.of(1). The same process applies, and the second definition is run. This, in turn, calls Factorial.of(0), which is matched by the first function definition. This function returns 1 and the recursion ends. Elixir now unwinds the stack, performing all the multiplications, and returns the answer. This factorial implementation works, but it could be significantly improved. We'll do that improvement when we look at tail recursion on page 202.

Let's play with this code:

This pattern of design and coding is very common in Elixir (and almost all functional languages). First look for the simplest possible case, one that has a definite answer. This will be the anchor. Then look for a recursive solution that will end up calling the anchor case.

Here are a couple of examples:

Sum of the first n numbers

- The sum of the first 0 numbers is 0.
- The sum of the numbers up to n is $n +$ the sum of the numbers up to $n-1$.

Length of a list

- The length of an empty list is 0.
- The length of any other list is 1 + the length of the tail of that list.

One point worth stressing: the order of these clauses can make a difference when you translate them into code. Elixir tries functions from the top down, executing the first match. So the following code will not work:

```
mm/factorial1-bad.exs
defmodule BadFactorial do
  def of(n), do: n * of(n-1)
  def of(0), do: 1
end
```

The first function definition will always match and the second will never be called. But Elixir has you covered—when you try to compile this, you’ll get a warning:

```
iex> c "factorial1-bad.exs"
.../factorial1-bad.ex:3: this clause cannot match because a previous clause at
line 2 always matches
```

One more thing: when you have multiple implementations of the same function, they should be adjacent in the source file.

Your Turn

► *Exercise: ModulesAndFunctions-4*

Implement and run a function `sum(n)` that uses recursion to calculate the sum of the integers from 1 to n . You’ll need to write this function inside a module in a separate file. Then load up IEx, compile that file, and try your function.

► *Exercise: ModulesAndFunctions-5*

Write a function `gcd(x,y)` that finds the greatest common divisor between two nonnegative integers. Algebraically, $gcd(x,y)$ is x if y is zero; it’s $gcd(y, rem(x,y))$ otherwise.

Guard Clauses

We've seen that pattern matching allows Elixir to decide which function to invoke based on the arguments passed. But what if we need to distinguish based on the argument types or on some test involving their values? For this, use *guard clauses*. These are predicates that are attached to a function definition using one or more `when` keywords. When doing pattern matching, Elixir first does the conventional parameter-based match and then evaluates any `when` predicates, executing the function only if at least one predicate is true.

```
mm/guard.exs
defmodule Guard do
  def what_is(x) when is_number(x) do
    IO.puts "#{x} is a number"
  end
  def what_is(x) when is_list(x) do
    IO.puts "#{inspect(x)} is a list"
  end
  def what_is(x) when is_atom(x) do
    IO.puts "#{x} is an atom"
  end
end

Guard.what_is(99)      # => 99 is a number
Guard.what_is(:cat)    # => cat is an atom
Guard.what_is([1,2,3])  # => [1,2,3] is a list
```

Recall our [factorial example on page 55](#):

```
mm/factorial1.exs
defmodule Factorial do
  def of(0), do: 1
  def of(n), do: n * of(n-1)
end
```

If we were to pass it a negative number, it would loop forever—no matter how many times you decrement `n`, it will never be zero. So it is a good idea to add a guard clause to stop this from happening.

```
mm/factorial2.exs
defmodule Factorial do
  def of(0), do: 1
  def of(n) when is_integer(n) and n > 0 do
    n * of(n-1)
  end
end
```

If you run this with a negative argument, none of the functions will match:

```
iex> c "factorial2.exs"
[Factorial]
iex> Factorial.of -100
** (FunctionClauseError) no function clause matching in Factorial.of/1...
```

Notice we've also added a type constraint: the parameter must be an integer.

Guard-Clause Limitations

You can write only a subset of Elixir expressions in guard clauses. The following list comes from the Getting Started guide.¹

Comparison operators

`==, !=, ===, !==, >, <, <=, >=`

Boolean and negation operators

`or, and, not, !.` Note that `||` and `&&` are not allowed.

Arithmetic operators

`+, -, *, /`

Join operators

`<>` and `++`, as long as the left side is a literal

The `in` operator

Membership in a collection or range

Type-check functions

These built-in Erlang functions return true if their argument is a given type. You can find their documentation online.²

<code>is_atom</code>	<code>is_binary</code>	<code>is_bitstring</code>	<code>is_boolean</code>	<code>is_exception</code>	<code>is_float</code>
<code>is_function</code>	<code>is_integer</code>	<code>is_list</code>	<code>is_map</code>	<code>is_number</code>	<code>is_pid</code>
<code>is_port</code>	<code>is_record</code>	<code>is_reference</code>	<code>is_tuple</code>		

Other functions

These built-in functions return values (not true or false). Their documentation is online, on the same page as the type-check functions.

<code>abs(number)</code>	<code>bit_size(bitstring)</code>	<code>byte_size(bitstring)</code>	<code>div(number,number)</code>
<code>elem(tuple, n)</code>	<code>float(term)</code>	<code>hd(list)</code>	<code>length(list)</code>
<code>node()</code>	<code>node(pid ref port)</code>	<code>rem(number,number)</code>	<code>round(number)</code>
<code>self()</code>	<code>tl(list)</code>	<code>trunc(number)</code>	<code>tuple_size(tuple)</code>

1. <http://elixir-lang.org/getting-started/case-cond-and-if.html#expressions-in-guard-clauses>

2. http://erlang.org/doc/man/erlang.html#is_atom-1

Guard Clauses vs. Conditional Logic

Have another look at our factorial function:

```
def of(0), do: 1
def of(n) when is_integer(n) and n > 0 do
  n * of(n-1)
end
```

You might think of writing the code like this:

```
def of(0), do: 1
def of(n) do
  if n < 0 do
    raise "factorial called on a negative number"
  else
    n * of(n-1)
  end
end
```

Logically these are the same, right? Both versions raise an exception when passed a negative number.

But they aren't. In the second case, the `of/1` function is defined for any input. But in the first case, it isn't defined at all for negative parameters. And that's what we really want: the first example makes it explicit that the domain of our function is nonnegative integers.

The difference between the two examples is subtle, but the first communicates what we want more accurately.

Default Parameters

When you define a named function, you can give a default value to any of its parameters by using the syntax `param \\
value`. When you call a function that is defined with default parameters, Elixir compares the number of arguments you are passing with the number of required parameters for the function. If you're passing fewer arguments than the number of required parameters, then there's no match. If the two numbers are equal, then the required parameters take the values of the passed arguments, and the other parameters take their default values. If the count of passed arguments is greater than the number of required parameters, Elixir uses the excess to override the default values of some or all parameters. Parameters are matched left to right.

```
mm/default_params.exs
defmodule Example do
  def func(p1, p2 \\ 2, p3 \\ 3, p4) do
    IO.inspect [p1, p2, p3, p4]
  end
end
```

```
Example.func("a", "b")           # => ["a", 2, 3, "b"]
Example.func("a", "b", "c")       # => ["a", "b", 3, "c"]
Example.func("a", "b", "c", "d")  # => ["a", "b", "c", "d"]
```

Default arguments can behave surprisingly when Elixir does pattern matching. For example, compile the following:

```
def func(p1, p2 \\ 2, p3 \\ 3, p4) do
  IO.inspect [p1, p2, p3, p4]
end

def func(p1, p2) do
  IO.inspect [p1, p2]
end
```

and you'll get this error:

```
** (CompileError) default_params.exs:7: def func/2 conflicts with
 defaults from def func/4
```

That's because the first function definition (with the default parameters) matches any call with two, three, or four arguments.

There's one more thing with default parameters. Here's a function with multiple heads that also has a default parameter:

```
mm/default_params1.exs
defmodule DefaultParams1 do

  def func(p1, p2 \\ 123) do
    IO.inspect [p1, p2]
  end

  def func(p1, 99) do
    IO.puts "you said 99"
  end

end
```

If you compile this, you'll get an error:

```
warning: definitions with multiple clauses and default values require a
function head. Instead of this:
```

```
def foo(:first_clause, b \\ :default) do ... end
def foo(:second_clause, b) do ... end
```

one should write this:

```
def foo(a, b \\ :default)
def foo(:first_clause, b) do ... end
def foo(:second_clause, b) do ... end
```

```
def func/2 has multiple clauses and defines defaults in a clause with a body
code/mm/default_params1.exs:8
```

```
warning: variable p1 is unused
  code/mm/default_params1.exs:8

warning: this clause cannot match because a previous clause at
  line 4 always matches code/mm/default_params1.exs:8
```

The intent is to reduce confusion that can arise with defaults. Add a function head with no body that contains the default parameters, and use regular parameters for the rest. The defaults will apply to all calls to the function.

```
mm/default_params2.exs
defmodule Params do
  def func(p1, p2 \\ 123)
  def func(p1, p2) when is_list(p1) do
    "You said #{p2} with a list"
  end
  def func(p1, p2) do
    "You passed in #{p1} and #{p2}"
  end
end

IO.puts Params.func(99)          # You passed in 99 and 123
IO.puts Params.func(99, "cat")    # You passed in 99 and cat
IO.puts Params.func([99])        # You said 123 with a list
IO.puts Params.func([99], "dog")  # You said dog with a list
```

Your Turn

► *Exercise: ModulesAndFunctions-6*

I'm thinking of a number between 1 and 1000....

The most efficient way to find the number is to guess halfway between the low and high numbers of the range. If our guess is too big, then the answer lies between the bottom of the range and one less than our guess. If our guess is too small, then the answer lies between one more than our guess and the end of the range.

Your API will be `guess(actual, range)`, where `range` is an Elixir range. Your output should look similar to this:

```
iex> Chop.guess(273, 1..1000)
Is it 500
Is it 250
Is it 375
Is it 312
Is it 281
Is it 265
Is it 273
273
```

Hints:

- You may need to implement helper functions with an additional parameter (the currently guessed number).
- The `div(a,b)` function performs integer division.
- Guard clauses are your friends.
- Patterns can match the low and high parts of a range (`a..b=4..8`).

Private Functions

The `defp` macro defines a private function—one that can be called only within the module that declares it.

You can define private functions with multiple heads, just as you can with `def`. However, you cannot have some heads private and others public. That is, the following code is not valid:

```
def fun(a) when is_list(a), do: true
defp fun(a), do: false
```

The Amazing Pipe Operator: |>

I've saved the best for last, at least when it comes to functions.

We've all seen code like this:

```
people = DB.find_customers
orders = Orders.for_customers(people)
tax    = sales_tax(orders, 2018)
filing = prepare_filing(tax)
```

Bread-and-butter programming. We did it because the alternative was to write

```
filing = prepare_filing(sales_tax(Orders.for_customers(DB.find_customers), 2018))
```

and that's the kind of code that you use to get kids to eat their vegetables. Not only is it hard to read, but you have to read it inside out if you want to see the order in which things get done.

Elixir has a better way of writing it:

```
filings = DB.find_customers
         |> Orders.for_customers
         |> sales_tax(2018)
         |> prepare_filing
```

The `|>` operator takes the result of the expression to its left and inserts it as the first parameter of the function invocation to its right. So the list of customers the first call returns becomes the argument passed to the `for_customers` function. The resulting list of orders becomes the first argument to `sales_tax`, and the given parameter, `2018`, becomes the second.

`val |> f(a,b)` is basically the same as calling `f(val,a,b)`, and

```
list
|> sales_tax(2018)
|> prepare_filing
```

is the same as `prepare_filing(sales_tax(list, 2018))`.

In the previous example, I wrote each term in the expression on a separate line, and that's perfectly valid Elixir. But you can also chain terms on the same line:

```
iex> (1..10) |> Enum.map(&(&1*&1)) |> Enum.filter(&(&1 < 40))
[1, 4, 9, 16, 25, 36]
```

Note that I had to use parentheses in that code—the `&` shortcut and the pipe operator fight otherwise.

Let me repeat that—you should always use parentheses around function parameters in pipelines.

The key aspect of the pipe operator is that it lets you write code that pretty much follows your spec's form. For the sales-tax example, you might have jotted this on some paper:

- Get the customer list.
- Generate a list of their orders.
- Calculate tax on the orders.
- Prepare the filing.

To take this from a napkin spec to running code, you just put `|>` between the items and implement each as a function.

```
DB.find_customers
|> Orders.for_customers
|> sales_tax(2018)
|> prepare_filing
```

Programming is transforming data, and the `|>` operator makes that transformation explicit.

And now this book's subtitle makes sense.

Modules

Modules provide namespaces for things you define. We've already seen them encapsulating named functions. They also act as wrappers for macros, structs, protocols, and other modules.

If we want to reference a function defined in a module from outside that module, we need to prefix the reference with the module's name. We don't need that prefix if code references something inside the same module as itself, as in the following example:

```
defmodule Mod do
  def func1 do
    IO.puts "in func1"
  end
  def func2 do
    func1
    IO.puts "in func2"
  end
end

Mod.func1
Mod.func2
```

func2 can call func1 directly because it is inside the same module. Outside the module, you have to use the fully qualified name, Mod.func1.

Just as you do in your favorite language, Elixir programmers use nested modules to impose structure for readability and reuse. After all, every programmer is a library writer.

To access a function in a nested module from the outside scope, prefix it with all the module names. To access it within the containing module, use either the fully qualified name or just the inner module name as a prefix.

```
defmodule Outer do
  defmodule Inner do
    def inner_func do
    end
  end

  def outer_func do
    Inner.inner_func
  end
end

Outer.outer_func
Outer.Inner.inner_func
```

Module nesting in Elixir is an illusion—all modules are defined at the top level. When we define a module inside another, Elixir simply prepends the outer module name to the inner module name, putting a dot between the two. This means we can directly define a nested module.

```
defmodule Mix.Tasks.Doctest do
  def run do
    end
end

Mix.Tasks.Doctest.run
```

It also means there's no particular relationship between the modules `Mix` and `Mix.Tasks.Doctest`.

Directives for Modules

Elixir has three directives that simplify working with modules. All three are executed as your program runs, and the effect of all three is *lexically scoped*—it starts at the point the directive is encountered, and stops at the end of the enclosing scope. This means a directive in a module definition takes effect from the place you wrote it until the end of the module; a directive in a function definition runs to the end of the function.

The import Directive

The `import` directive brings a module's functions and/or macros into the current scope. If you use a particular module a lot in your code, `import` can cut down the clutter in your source files by eliminating the need to repeat the module name time and again.

For example, if you import the `flatten` function from the `List` module, you'd be able to call it in your code without having to specify the module name.

```
mm/import.exs
defmodule Example do
  def func1 do
    List.flatten [1,[2,3],4]
  end
  def func2 do
    import List, only: [flatten: 1]
    flatten [5,[6,7],8]
  end
end
```

The full syntax of `import` is

```
import Module [, only: |except: ]
```

The optional second parameter lets you control which functions or macros are imported. You write only: or except:, followed by a list of name: arity pairs. It is a good idea to use import in the smallest possible enclosing scope and to use only: to import just the functions you need.

```
import List, only: [ flatten: 1, duplicate: 2 ]
```

Alternatively, you can give only: one of the atoms :functions or :macros, and import will bring in only functions or macros.

The alias Directive

The alias directive creates an alias for a module. One obvious use is to cut down on typing.

```
defmodule Example do
  def compile_and_go(source) do
    alias My.Other.Module.Parser, as: Parser
    alias My.Other.Module.Runner, as: Runner
    source
    |> Parser.parse()
    |> Runner.execute()
  end
end
```

We could have abbreviated these alias directives to

```
alias My.Other.Module.Parser
alias My.Other.Module.Runner
```

because the **as:** parameters default to the last part of the module name. We could even take this further, and do this:

```
alias My.Other.Module.{Parser, Runner}
```

The require Directive

You require a module if you want to use any macros it defines. This ensures that the macro definitions are available when your code is compiled. We'll talk about require when we discuss [macros on page 303](#).

Module Attributes

Elixir modules each have associated metadata. Each item of metadata is called an *attribute* of the module and is identified by a name. Inside a module, you can access these attributes by prefixing the name with an at sign (@). You give an attribute a value using the syntax

```
@name value
```

This works only at the top level of a module—you can't set an attribute inside a function definition. You can, however, access attributes inside functions.

```
mm/attributes.exs
defmodule Example do
  @author "Dave Thomas"
  def get_author do
    @author
  end
end
IO.puts "Example was written by #{Example.get_author}"
```

You can set the same attribute multiple times in a module. If you access that attribute in a named function in that module, the value you see will be the value in effect when the function is defined.

```
mm/attributes1.exs
defmodule Example do
  @attr "one"
  def first, do: @attr
  @attr "two"
  def second, do: @attr
end
IO.puts "#{Example.second} #{Example.first}" # => two one
```

These attributes are not variables in the conventional sense. Use them for configuration and metadata only. (Many Elixir programmers employ them where Java or Ruby programmers might use constants.)

Module Names: Elixir, Erlang, and Atoms

When we write modules in Elixir, they have names such as `String` or `PhotoAlbum`. We call functions in them using calls such as `String.length("abc")`.

What's happening here is subtle. Internally, module names are just atoms. When you write a name starting with an uppercase letter, such as `IO`, Elixir converts it internally into an atom of the same name with `Elixir.` prepended. So `IO` becomes `Elixir.IO` and `Dog` becomes `Elixir.Dog`.

```
iex> is_atom IO
true
iex> to_string IO
"Elixir.IO"
iex> :"Elixir.IO" === IO
true
```

So a call to a function in a module is really an atom followed by a dot followed by the function name. And, indeed, we can call functions like this:

```
iex> IO.puts 123
123
:ok
iex> :"Elixir.IO".puts 123
123
:ok
```

and even

```
iex> my_io = IO
IO
iex> my_io.puts 123
123
:ok
```

Calling a Function in an Erlang Library

The Erlang conventions for names are different—variables start with an uppercase letter and atoms are simple lowercase names. So, for example, the Erlang module `timer` is called just that, the atom `timer`. In Elixir we write that as `:timer`. If you want to refer to the `tc` function in `timer`, you'd write `:timer.tc`. (Note the colon at the start.)

Say we want to output a floating-point number in a three-character-wide field with one decimal place. Erlang has a function for this. A search for `erlang format` takes us to the description of the `format` function in the Erlang `io` module.³

Reading the description, we see that Erlang expects us to call `io.format`. So, in Elixir we simply change the Erlang module name to an Elixir atom:

```
iex> :io.format("The number is ~3.1f~n", [5.678])
The number is 5.7
:ok
```

Finding Libraries

If you're looking for a library to use in your app, you'll want to look first for existing Elixir modules. The built-in ones are documented on the Elixir website,⁴ and others are listed at [hex.pm](#) and on GitHub (search for `elixir`).

If that fails, search for a built-in Erlang library or search the web.⁵ If you find something written in Erlang, you'll be able to use it in your project (we'll cover how in the chapter on [projects, on page 141](#)). But be aware that the Erlang documentation for a library follows Erlang conventions. Variables start

3. <http://erlang.org/doc/man/io.html#format-2>

4. <http://elixir-lang.org/docs/>

5. <http://erlang.org/doc/> and <http://erldocs.com/19.0/> (Note that the latter is slightly out of date.)

with uppercase letters, and identifiers that start with a lowercase letter are atoms (so Erlang would say tomato and Elixir would say :tomato). A summary of the differences between Elixir and Erlang is available online.⁶

Now that we've looked at functions, let's move on to the data they manipulate. And where better to start than with lists? They're the subject of the next chapter.

Your Turn

► *Exercise: ModulesAndFunctions-7*

Find the library functions to do the following, and then use each in IEx. (If the word *Elixir* or *Erlang* appears at the end of the challenge, then you'll find the answer in that set of libraries.)

- Convert a float to a string with two decimal digits. (Erlang)
- Get the value of an operating-system environment variable. (Elixir)
- Return the extension component of a file name (so return .exs if given "dave/test.exs"). (Elixir)
- Return the process's current working directory. (Elixir)
- Convert a string containing JSON into Elixir data structures. (Just find; don't install.)
- Execute a command in your operating system's shell.

6. <http://elixir-lang.org/crash-course.html>

In this chapter, you'll see:

- The recursive structure of lists
- Traversing and building lists
- Accumulators
- Implementing map and reduce

CHAPTER 7

Lists and Recursion

When we program with lists in conventional languages, we treat them as things to be iterated—it seems natural to loop over them. So why do we have a chapter on *lists and recursion*? Because if you look at the problem in the right way, recursion is a perfect tool for processing lists.

Heads and Tails

Earlier we said a list may either be empty or consist of a head and a tail. The head contains a value and the tail is itself a list. This is a recursive definition.

We'll represent the empty list like this: [].

Let's imagine we could represent the split between the head and the tail using a pipe character: |. The single-element list we normally write as [3] can be written as the value 3 joined to the empty list:

[3 | []]

(I've highlighted the inner list.)

When we see the pipe character, we say that what's on the left is the head of a list and what's on the right is the tail.

Let's look at the list [2, 3]. The head is 2, and the tail is the single-element list containing 3. And we know what that list looks like—it is our previous example. So we could write [2,3] as

[2 | [3 | []]]

At this point, part of your brain is telling you to go read today's XKCD—this list stuff can't be useful. Ignore that small voice, just for a second. We're about to do something magical. But before we do, let's add one more term, making

our list [1, 2, 3]. This is the head 1 followed by the list [2, 3], which is what we derived a moment ago:

```
[ 1 | [ 2 | [ 3 | [] ] ] ]
```

This is valid Elixir syntax. Type it into IEx.

```
iex> [ 1 | [ 2 | [ 3 | [] ] ] ]
[1, 2, 3]
```

And here's the magic. When we discussed pattern matching, we said the pattern could be a list, and the values in that list would be assigned from the right-hand side.

```
iex> [a, b, c] = [1, 2, 3]
[1, 2, 3]
iex> a
1
iex> b
2
iex> c
3
```

We can also use the pipe character in the pattern. What's to the left of it matches the head value of the list, and what's to the right matches the tail.

```
iex> [ head | tail ] = [1, 2, 3]
[1, 2, 3]
iex> head
1
iex> tail
[2, 3]
```

Using Head and Tail to Process a List

Now we can split a list into its head and its tail, and we can construct a list from a value and a list, which become the head and tail of that new list.

So why talk about lists after we talk about modules and functions? Because lists and recursive functions go together like fish and chips. Let's look at finding the length of a list.

- The length of an empty list is 0.
- The length of a list is 1 plus the length of that list's tail.

How IEx Displays Lists

In [Chapter 11, Strings and Binaries, on page 117](#), you'll see that Elixir has two representations for strings. One is the familiar sequence of characters in consecutive memory locations. Literals written with double quotes use this form.

The second form, using single quotes, represents a string as a list of integer codepoints. So the string 'cat' is the three codepoints: 99, 97, and 116.

This is a headache for IEx. When it sees a list like [99,97,116] it doesn't know if it is supposed to be the string 'cat' or a list of three numbers. So it uses a heuristic. If all the values in a list represent printable characters, it displays the list as a string; otherwise it displays a list of integers.

```
iex> [99, 97, 116]
'cat'
iex> [99, 97, 116, 0]  # '0' is nonprintable
[99, 97, 116, 0]
```

In [Chapter 11, Strings and Binaries, on page 117](#), we'll cover how to bypass this behavior. In the meantime, don't be surprised if a string pops up when you were expecting a list.

Writing the list-length algorithm in Elixir is easy:

```
lists/mylist.exs
defmodule MyList do
  def len([]), do: 0
  def len([head|tail]), do: 1 + len(tail)
end
```

The only tricky part is the definition of the function's second variant:

```
def len([ head | tail ]) ...
```

This is a pattern match for any nonempty list. When it does match, the variable head will hold the value of the first element of the list, and tail will hold the rest of the list. (And remember that every list is terminated by an empty list, so the tail can be []).

Let's see this at work with the list [11, 12, 13, 14, 15]. At each step, we take off the head and add 1 to the length of the tail:

```
len([11,12,13,14,15])
= 1 + len([12,13,14,15])
= 1 + 1 + len([13,14,15])
= 1 + 1 + 1 + len([14,15])
= 1 + 1 + 1 + 1 + len([15])
= 1 + 1 + 1 + 1 + 1 + len([])
= 1 + 1 + 1 + 1 + 1 + 0
= 5
```

Let's try our code to see if theory works in practice:

```
iex> c "mylist.exs"
...mylist.exs:3: variable head is unused
[MyList]
iex> MyList.len([])
0
iex> MyList.len([11,12,13,14,15])
5
```

It works, but we have a compilation warning—we never used the variable `head` in the body of our function. We can fix that, and make our code more explicit, using the special variable `_` (underscore), which acts as a placeholder. We can also use an underscore in front of any variable name to turn off the warning if that variable isn't used. I sometimes like to do this to document the unused parameter.

```
lists/mylist1.exs
defmodule MyList do
  def len([]),           do: 0
  def len([ _head | tail ]), do: 1 + len(tail)
end
```

When we compile, the warning is gone. (However, if you compile the second version of `MyList`, you may get a warning about “redefining module `MyList`.” This is just Elixir being cautious.)

```
iex> c "mylist1.exs"
[MyList]
iex> MyList.len([1,2,3,4,5])
5
iex> MyList.len(["cat", "dog"])
2
```

Using Head and Tail to Build a List

Let's get more ambitious. Let's write a function that takes a list of numbers and returns a new list containing the square of each. We don't show it, but these definitions are also inside the `MyList` module.

```
lists/mylist1.exs
def square([]),           do: []
def square([ head | tail ]), do: [ head*head | square(tail) ]
```

There's a lot going on here. First, look at the parameter patterns for the two definitions of `square`. The first matches an empty list and the second matches all other lists.

Second, look at the body of the second definition:

```
def square([ head | tail ]), do: [ head*head | square(tail) ]
```

When we match a nonempty list, we return a new list whose head is the square of the original list's head and whose tail is a list of squares of the tail. This is the recursive step.

Let's try it:

```
iex> c "mylist1.exs"
[MyList]
iex> MyList.square []
[]                                     # this calls the 1st definition
[]
iex> MyList.square [4,5,6]    # and this calls the 2nd
[16, 25, 36]
```

Let's do something similar—a function that adds 1 to every element in the list:

```
lists/mylist1.exs
def add_1([]),                  do: []
def add_1([ head | tail ]), do: [ head+1 | add_1(tail) ]
```

And call it:

```
iex> c "mylist1.exs"
[MyList]
iex> MyList.add_1 [1000]
[1001]
iex> MyList.add_1 [4,6,8]
[5, 7, 9]
```

Creating a Map Function

With both `square` and `add_1`, all the work is done in the second function definition. And that definition looks about the same for each—it returns a new list whose head is the result of either squaring or incrementing the head of its argument and whose tail is the result of calling itself recursively on the tail of the argument. Let's generalize this. We'll define a function called `map` that takes a list and a function and returns a new list containing the result of applying that function to each element in the original.

```
lists/mylist1.exs
def map([], _func),           do: []
def map([ head | tail ], func), do: [ func.(head) | map(tail, func) ]
```

The `map` function is pretty much identical to the `square` and `add_1` functions. It returns an empty list if passed an empty list; otherwise it returns a list where the head is the result of calling the passed-in function and the tail is a recursive call to itself. Note that in the case of an empty list, we use `_func` as

the second parameter. The underscore prevents Elixir from warning us about an unused variable.

To call this function, pass in a list and a function (defined using `fn`):

```
iex> c "mylist1.exs"
[MyList]
iex> MyList.map [1,2,3,4], fn (n) -> n*n end
[1, 4, 9, 16]
```

A function is just a built-in type, defined between `fn` and the `end`. Here we pass a function as the second argument (`func`) to `map`. This is invoked inside `map` using `func.(head)`, which squares the value in `head`, using the result to build the new list.

We can call `map` with a different function:

```
iex> MyList.map [1,2,3,4], fn (n) -> n+1 end
[2, 3, 4, 5]
```

And another:

```
iex> MyList.map [1,2,3,4], fn (n) -> n > 2 end
[false, false, true, true]
```

And we can do the same using the `&` shortcut notation:

```
iex> MyList.map [1,2,3,4], &(&l + 1)
[2, 3, 4, 5]
iex> MyList.map [1,2,3,4], &(&l > 2)
[false, false, true, true]
```

Reducing a List to a Single Value

The `map/2` function we just wrote abstracts the idea of applying a function to each element of a list independently.

But what if we want to apply that function across the elements? How could we create an abstraction that would let us sum a list, or find the product of its elements, or find the largest element?

The `sum` function reduces a collection to a single value. Other functions need to do something similar—return the greatest/least value, the product of the elements, a string containing a concatenation of elements, and so on. How can we write a general-purpose function that reduces a collection to a value?

We know it has to take a collection. We also know we need to pass in some initial value (just like our `sum/1` function passed a 0 as an initial value to its

helper). Additionally, we need to pass in a function that takes the current value of the reduction along with the next element of the collection, and returns the next value of the reduction. So, it looks like our reduce function will be called with three arguments:

```
reduce(collection, initial_value, fun)
```

Now let's think about the recursive design:

- $\text{reduce}([], \text{value}, \text{fun}) \rightarrow \text{value}$
- $\text{reduce}([\text{head} | \text{tail}], \text{value}, \text{fun}) \rightarrow \text{reduce}(\text{tail}, \text{fun}.\text{(head, value)}), \text{fun}$

reduce applies the function to the list's head and the current value, and passes the result as the new current value when reducing the list's tail.

Here's our code for reduce. See how closely it follows the design:

```
lists/reduce.exs
defmodule MyList do
  def reduce([], value, _) do
    value
  end
  def reduce([head | tail], value, func) do
    reduce(tail, func.(head, value), func)
  end
end
```

And, again, we can use the shorthand notation to pass in the function:

```
iex> c "reduce.exs"
[MyList]
iex> MyList.reduce([1,2,3,4,5], 0, &(&1 + &2))
15
iex> MyList.reduce([1,2,3,4,5], 1, &(&1 * &2))
120
```

Your Turn

► *Exercise: ListsAndRecursion-1*

Write a mapsum function that takes a list and a function. It applies the function to each element of the list and then sums the result, so

```
iex> MyList.mapsum [1, 2, 3], &(&1 * &1)
14
```

► *Exercise: ListsAndRecursion-2*

Write a max(list) that returns the element with the maximum value in the list. (This is slightly trickier than it sounds.)

► *Exercise: ListsAndRecursion-3*

An Elixir single-quoted string is actually a list of individual character codes. Write a `caesar(list, n)` function that adds `n` to each list element, wrapping if the addition results in a character greater than `z`.

```
iex> MyList.caesar('ryvkve', 13)
?????? :)
```

More Complex List Patterns

Not every list problem can be easily solved by processing one element at a time. Fortunately, the join operator, `|`, supports multiple values to its left. Thus, you could write

```
iex> [ 1, 2, 3 | [ 4, 5, 6 ]]
[1, 2, 3, 4, 5, 6]
```

The same thing works in patterns, so you can match multiple individual elements as the head. For example, this program swaps pairs of values in a list:

```
lists/swap.exs
defmodule Swapper do
  def swap([]), do: []
  def swap([ a, b | tail ]), do: [ b, a | swap(tail) ]
  def swap([_]), do: raise "Can't swap a list with an odd number of elements"
end
```

We can play with it in IEx:

```
iex> c "swap.exs"
[Swapper]
iex> Swapper.swap [1,2,3,4,5,6]
[2, 1, 4, 3, 6, 5]
iex> Swapper.swap [1,2,3,4,5,6,7]
** (RuntimeError) Can't swap a list with an odd number of elements
```

The third definition of `swap` matches a list with a single element. This will happen if we get to the end of the recursion and have only one element left. Given that we take two values off the list on each cycle, the initial list must have had an odd number of elements.

Lists of Lists

Let's imagine we had recorded temperatures and rainfall at a number of weather stations. Each reading looks like this:

```
[ timestamp, location_id, temperature, rainfall ]
```

Our code is passed a list containing a number of these readings, and we want to report on the conditions for one particular location, number 27.

```
lists/weather.exs
defmodule WeatherHistory do
  def for_location_27([], do: [])
  def for_location_27([ [time, 27, temp, rain] | tail]), do:
    [ [time, 27, temp, rain] | for_location_27(tail) ]
  end
  def for_location_27([ _ | tail]), do: for_location_27(tail)
end
```

This is a standard *recurse until the list is empty* stanza. But look at our function definition's second clause. Where we'd normally match into a variable called head, here the pattern is

```
for_location_27([ [ time, 27, temp, rain ] | tail])
```

For this to match, the head of the list must itself be a four-element list, and the second element of this sublist must be 27. This function will execute only for entries from the desired location. But when we do this kind of filtering, we also have to remember to deal with the case when our function doesn't match. That's what the third line does. We could have written

```
for_location_27([ [ time, _, temp, rain ] | tail])
```

but in reality we don't care *what* is in the head at this point.

In the same module we define some simple test data:

```
lists/weather.exs
def test_data do
  [
    [1366225622, 26, 15, 0.125],
    [1366225622, 27, 15, 0.45],
    [1366225622, 28, 21, 0.25],
    [1366229222, 26, 19, 0.081],
    [1366229222, 27, 17, 0.468],
    [1366229222, 28, 15, 0.60],
    [1366232822, 26, 22, 0.095],
    [1366232822, 27, 21, 0.05],
    [1366232822, 28, 24, 0.03],
    [1366236422, 26, 17, 0.025]
  ]
end
```

We can use that to play with our function in IEx. To make this easier, I'm using the import function. This adds the functions in WeatherHistory to our local

name scope. After calling `import` we don't have to put the module name in front of every function call.

```
iex> c "weather.exs"
[WeatherHistory]
iex> import WeatherHistory
WeatherHistory
iex> for_location_27(test_data)
[[1366225622, 27, 15, 0.45], [1366229222, 27, 17, 0.468],
 [1366232822, 27, 21, 0.05]]
```

Our function is specific to a particular location, which is pretty limiting. We'd like to be able to pass in the location as a parameter. We can use pattern matching for this.

```
lists/weather2.exs
defmodule WeatherHistory do
  def for_location([], _target_loc), do: []
  ➤ def for_location([ [time, target_loc, temp, rain] | tail], target_loc) do
    [time, target_loc, temp, rain] | for_location(tail, target_loc)
  end
  def for_location([ _ | tail], target_loc), do: for_location(tail, target_loc)
end
```

Now the second function fires only when the location extracted from the list head equals the target location passed as a parameter.

But we can improve on this. Our filter doesn't care about the other three fields in the head—it just needs the location. But we do need the value of the head itself to create the output list. Fortunately, Elixir pattern matching is recursive and we can match patterns inside patterns.

```
lists/weather3.exs
defmodule WeatherHistory do
  def for_location([], _target_loc), do: []
  ➤ def for_location([ head = [_, target_loc, _, _] | tail], target_loc) do
    [head | for_location(tail, target_loc)]
  end
  def for_location([ _ | tail], target_loc), do: for_location(tail, target_loc)
end
```

The key change here is this line:

```
def for_location([ head = [_, target_loc, _, _] | tail], target_loc)
```

Compare that with the previous version:

```
def for_location([ [ time, target_loc, temp, rain ] | tail], target_loc)
```

In the new version, we use placeholders for the fields we don't care about. But we also match the entire four-element array into the parameter head. It's as if we said, "Match the head of the list where the second element is matched to target_loc and then match that whole head with the variable head." We've extracted an individual component of the sublist as well as the entire sublist.

In the original body of for_location, we generated our result list using the individual fields:

```
def for_location([ [ time, target_loc, temp, rain ] | tail], target_loc)
  [ [ time, target_loc, temp, rain ] | for_location(tail, target_loc) ]
end
```

In the new version, we can just use the head, making it a lot clearer:

```
def for_location([ head = [_, target_loc, _, _ ] | tail], target_loc) do
  [ head | for_location(tail, target_loc) ]
end
```

Your Turn

► *Exercise: ListsAndRecursion-4*

Write a function `MyList.span(from, to)` that returns a list of the numbers from from up to to.

The List Module in Action

The List module provides a set of functions that operate on lists.

```
#  
# Concatenate lists  
#  
iex> [1,2,3] ++ [4,5,6]  
[1, 2, 3, 4, 5, 6]  
#  
# Flatten  
#  
iex> List.flatten([[1], 2], [[[3]]])  
[1, 2, 3]  
#  
# Folding (like reduce, but can choose direction)  
#  
iex> List.foldl([1,2,3], "", fn value, acc -> "#{value}(#{acc})" end)  
"3(2(1()))"  
iex> List.foldr([1,2,3], "", fn value, acc -> "#{value}(#{acc})" end)  
"1(2(3()))"
```

```

#
# Updating in the middle (not a cheap operation)
#
iex> list = [ 1, 2, 3 ]
[ 1, 2, 3 ]
iex> List.replace_at(list, 2, "buckle my shoe")
[1, 2, "buckle my shoe"]
#
# Accessing tuples within lists
#
iex> kw = [{:name, "Dave"}, {:likes, "Programming"}, {:where, "Dallas", "TX"}]
[{:name, "Dave"}, {:likes, "Programming"}, {:where, "Dallas", "TX"}]
iex> List.keyfind(kw, "Dallas", 1)
{:where, "Dallas", "TX"}
iex> List.keyfind(kw, "TX", 2)
{:where, "Dallas", "TX"}
iex> List.keyfind(kw, "TX", 1)
nil
iex> List.keyfind(kw, "TX", 1, "No city called TX")
"No city called TX"
iex> kw = List.keydelete(kw, "TX", 2)
[name: "Dave", likes: "Programming"]
iex> kw = List.keyreplace(kw, :name, 0, {:first_name, "Dave"})
[first_name: "Dave", likes: "Programming"]

```

Get Friendly with Lists

Lists are the natural data structure to use when you have a stream of values to handle. You'll use them to parse data, handle collections of values, and record the results of a series of function calls. It's worth spending a while getting comfortable with them.

Next we'll look at the various dictionary types, including maps. These let us organize data into collections of key/value pairs.

In this chapter, you'll see:

- The two and a half dictionary data types
- Pattern matching and updating maps
- Structs
- Nested data structures

CHAPTER 8

Maps, Keyword Lists, Sets, and Structs

A dictionary is a data type that associates keys with values.

We've already looked briefly at the dictionary types: maps and keyword lists. In this short chapter we'll cover how to use them with pattern matching and how to update them. Then we'll dive into structs, a special kind of map with a fixed structure. Finally we'll explore nested data structures and see how to alter fields in a map inside another map inside another map....

First, though, let's answer a common question—how do we choose an appropriate dictionary type for a particular need?

How to Choose Between Maps, Structs, and Keyword Lists

Ask yourself these questions (in this order):

- Do I want to pattern-match against the contents (for example, matching a dictionary that has a key of :name somewhere in it)?

If so, use a map.

- Will I want more than one entry with the same key?

If so, you'll have to use the Keyword module.

- Do I need to guarantee the elements are ordered?

If so, again, use the Keyword module.

- Do I have a fixed set of fields (that is, is the structure of the data always the same)?

If so, use a struct.

- Otherwise, if you've reached this point,

Use a map.

Keyword Lists

Keyword lists are typically used in the context of options passed to functions.

```
maps/keywords.exs
defmodule Canvas do
  @defaults [ fg: "black", bg: "white", font: "Merriweather" ]

  def draw_text(text, options \\ []) do
    options = Keyword.merge(@defaults, options)
    IO.puts "Drawing text #{inspect(text)}"
    IO.puts "Foreground: #{options[:fg]}"
    IO.puts "Background: #{Keyword.get(options, :bg)}"
    IO.puts "Font: #{Keyword.get(options, :font)}"
    IO.puts "Pattern: #{Keyword.get(options, :pattern, "solid")}"
    IO.puts "Style: #{inspect Keyword.get_values(options, :style)}"
  end
end

Canvas.draw_text("hello", fg: "red", style: "italic", style: "bold")

# =>
# Drawing text "hello"
# Foreground: red
# Background: white
# Font: Merriweather
# Pattern: solid
# Style: ["italic", "bold"]
```

For simple access, you can use the access operator, `kwlist[key]`. In addition, all the functions of the Keyword and Enum modules are available.^{1,2}

Maps

Maps are the go-to key/value data structure in Elixir. They have good performance at all sizes.

Let's play with the Map API:³

```
iex> map = %{ name: "Dave", likes: "Programming", where: "Dallas" }
%{likes: "Programming", name: "Dave", where: "Dallas"}
iex> Map.keys map
[:likes, :name, :where]
iex> Map.values map
["Programming", "Dave", "Dallas"]
iex> map[:name]
```

- <http://elixir-lang.org/docs/master/elixir/Keyword.html>
- <http://elixir-lang.org/docs/master/elixir/Enum.html>
- <http://elixir-lang.org/docs/master/elixir/Map.html>

```

"Dave"
iex> map.name
"Dave"
iex> map1 = Map.drop map, [:where, :likes]
%{name: "Dave"}
iex> map2 = Map.put map, :also_likes, "Ruby"
%{also_likes: "Ruby", likes: "Programming", name: "Dave", where: "Dallas"}
iex> Map.keys map2
[:also_likes, :likes, :name, :where]
iex> Map.has_key? map1, :where
false
iex> { value, updated_map } = Map.pop map2, :also_likes
{"Ruby", %{likes: "Programming", name: "Dave", where: "Dallas"}}
iex> Map.equal? map, updated_map
true

```

Pattern Matching and Updating Maps

The question we most often ask of our maps is, “Do you have the following keys (and maybe values)?” For example, given this map:

```
person = %{ name: "Dave", height: 1.88 }
```

- Is there an entry with the key :name?

```

iex> %{ name: a_name } = person
%{height: 1.88, name: "Dave"}
iex> a_name
"Dave"

```

- Are there entries for the keys :name and :height?

```

iex> %{ name: _, height: _ } = person
%{height: 1.88, name: "Dave"}

```

- Does the entry with key :name have the value "Dave"?

```

iex> %{ name: "Dave" } = person
%{height: 1.88, name: "Dave"}

```

Our map does not have the key :weight, so the following pattern match fails:

```

iex> %{ name: _, weight: _ } = person
** (MatchError) no match of right hand side value: %{height: 1.88, name: "Dave"}

```

It’s worth noting how the first pattern match destructured the map, extracting the value associated with the key :name. We can use this in many ways. Here’s one example. The for construct lets us iterate over a collection, filtering as we go. We cover it when we talk about [enumerating on page 111](#). The following example uses for to iterate over a list of people. Destructuring is used to extract the height value, which is used to filter the results.

maps/query.exs

```
people = [
  %{ name: "Grumpy",      height: 1.24 },
  %{ name: "Dave",        height: 1.88 },
  %{ name: "Dopey",       height: 1.32 },
  %{ name: "Shaquille",   height: 2.16 },
  %{ name: "Sneezy",      height: 1.28 }
]

IO.inspect(for person = %{height: height} <- people, height > 1.5, do: person)
```

This produces

```
[%{height: 1.88, name: "Dave"}, %{height: 2.16, name: "Shaquille"}]
```

In this code, we feed a list of maps to our comprehension. The generator clause binds each map (as a whole) to person and binds the height from that map to height. The filter selects only those maps where the height exceeds 1.5, and the do block returns the people that match. The comprehension as a whole returns a list of these people, which IO.inspect prints.

Clearly pattern matching is just pattern matching, so this maps capability works equally well in cond expressions, function head matching, and any other circumstances in which patterns are used.

maps/book_room.exs

```
defmodule HotelRoom do

  def book(%{name: name, height: height})
  when height > 1.9 do
    IO.puts "Need extra-long bed for #{name}"
  end

  def book(%{name: name, height: height})
  when height < 1.3 do
    IO.puts "Need low shower controls for #{name}"
  end

  def book(person) do
    IO.puts "Need regular bed for #{person.name}"
  end

end

people |> Enum.each(&HotelRoom.book/1)

#=> Need low shower controls for Grumpy
#  Need regular bed for Dave
#  Need regular bed for Dopey
#  Need extra-long bed for Shaquille
#  Need low shower controls for Sneezy
```

Pattern Matching Can't Bind Keys

You can't bind a value to a key during pattern matching. You can write this:

```
iex> %{ 2 => state } = %{ 1 => :ok, 2 => :error }
%{1 => :ok, 2 => :error}
iex> state
:error
```

but not this:

```
iex> %{ item => :ok } = %{ 1 => :ok, 2 => :error }
** (CompileError) iex:5: illegal use of variable item in map key...
```

Pattern Matching Can Match Variable Keys

When we looked at basic pattern matching, we saw that the [pin operator on page 19](#) uses the value already in a variable on the left-hand side of a match. We can do the same with the keys of a map:

```
iex> data = %{ name: "Dave", state: "TX", likes: "Elixir" }
%{likes: "Elixir", name: "Dave", state: "TX"}
iex> for key <- [ :name, :likes ] do
...>   %{ ^key => value } = data
...>   value
...> end
["Dave", "Elixir"]
```

Updating a Map

Maps let us add new key/value entries and update existing entries without traversing the whole structure. But as with all values in Elixir, a map is immutable, and so the result of the update is a new map.

The simplest way to update a map is with this syntax:

```
new_map = %{ old_map | key => value, ... }
```

This creates a new map that is a copy of the old, but the values associated with the keys on the right of the pipe character are updated:

```
iex> m = %{ a: 1, b: 2, c: 3 }
%{a: 1, b: 2, c: 3}
iex> m1 = %{ m | b: "two", c: "three" }
%{a: 1, b: "two", c: "three"}
iex> m2 = %{ m1 | a: "one" }
%{a: "one", b: "two", c: "three"}
```

However, this syntax will not add a new key to a map. To do this, you have to use the `Map.put_new/3` function.

Structs

When Elixir sees `%{ ... }` it knows it is looking at a map. But it doesn't know much more than that. In particular, it doesn't know what you intend to do with the map, whether only certain keys are allowed, or whether some keys should have default values.

That's fine for anonymous maps. But what if we want to create a typed map—a map that has a fixed set of fields and default values for those fields, and that you can pattern-match by type as well as content.

Enter the *struct*.

A struct is just a module that wraps a limited form of map. It's limited because the keys must be atoms and because these maps don't have Dict capabilities. The name of the module becomes the name of the map type.

Inside the module, you use the `defstruct` macro to define the struct's members.

```
maps/defstruct.exs
defmodule Subscriber do
  defstruct name: "", paid: false, over_18: true
end
```

Let's play with this in IEx:

```
$ iex defstruct.exs
iex> s1 = %Subscriber{}
%Subscriber{name: "", over_18: true, paid: false}
iex> s2 = %Subscriber{ name: "Dave" }
%Subscriber{name: "Dave", over_18: true, paid: false}
iex> s3 = %Subscriber{ name: "Mary", paid: true }
%Subscriber{name: "Mary", over_18: true, paid: true}
```

The syntax for creating a struct is the same as the syntax for creating a map—you simply add the module name between the `%` and the `{`.

You access the fields in a struct using dot notation or pattern matching:

```
iex> s3.name
"Mary"
iex> %Subscriber{name: a_name} = s3
%Subscriber{name: "Mary", over_18: true, paid: true}
iex> a_name
"Mary"
```

And updates follow suit:

```
iex> s4 = %Subscriber{ s3 | name: "Marie" }
%Subscriber{name: "Marie", over_18: true, paid: true}
```

Why are structs wrapped in a module? The idea is that you are likely to want to add struct-specific behavior.

```
maps/defstruct1.exs
defmodule Attendee do
  defstruct name: "", paid: false, over_18: true

  def may_attend_after_party(attendee = %Attendee{}) do
    attendee.paid && attendee.over_18
  end

  def print_vip_badge(%Attendee{name: name}) when name != "" do
    IO.puts "Very cheap badge for #{name}"
  end

  def print_vip_badge(%Attendee{}) do
    raise "missing name for badge"
  end
end

$ iex defstruct1.exs
iex> a1 = %Attendee{name: "Dave", over_18: true}
%Attendee{name: "Dave", over_18: true, paid: false}
iex> Attendee.may_attend_after_party(a1)
false
iex> a2 = %Attendee{a1 | paid: true}
%Attendee{name: "Dave", over_18: true, paid: true}
iex> Attendee.may_attend_after_party(a2)
true
iex> Attendee.print_vip_badge(a2)
Very cheap badge for Dave
:ok
iex> a3 = %Attendee{}
%Attendee{name: "", over_18: true, paid: false}
iex> Attendee.print_vip_badge(a3)
** (RuntimeError) missing name for badge...
```

Notice how we could call the functions in the `Attendee` module to manipulate the associated struct.

Structs also play a large role when implementing polymorphism, which we'll see when we look at [protocols on page 329](#).

Nested Dictionary Structures

The various dictionary types let us associate keys with values. But those values can themselves be dictionaries. For example, we may have a bug-reporting system. We could represent this using the following:

```
maps/nested.exs
defmodule Customer do
  defstruct name: "", company: ""
end

defmodule BugReport do
  defstruct owner: %Customer{}, details: "", severity: 1
end
```

Let's create a simple report:

```
iex> report = %BugReport{owner: %Customer{name: "Dave", company: "Pragmatic"}, ...
...>                               details: "broken"}
%BugReport{details: "broken", severity: 1,
          owner: %Customer{company: "Pragmatic", name: "Dave"}}
```

The owner attribute of the report is itself a Customer struct.

We can access nested fields using regular dot notation:

```
iex> report.owner.company
"Pragmatic"
```

But now our customer complains the company name is incorrect—it should be PragProg. Let's fix it:

```
iex> report = %BugReport{ report | owner:
...>                           %Customer{ report.owner | company: "PragProg" }}
%BugReport{details: "broken",
          owner: %Customer{company: "PragProg", name: "Dave"}, 
          severity: 1}
```

Ugly stuff! We had to update the overall bug report's owner attribute with an updated customer structure. This is verbose, hard to read, and error prone.

Fortunately, Elixir has a set of nested dictionary-access functions. One of these, `put_in`, lets us set a value in a nested structure:

```
iex> put_in(report.owner.company, "PragProg")
%BugReport{details: "broken",
          owner: %Customer{company: "PragProg", name: "Dave"}, 
          severity: 1}
```

This isn't magic—it's simply a macro that generates the long-winded code we'd have to have written otherwise.

The `update_in` function lets us apply a function to a value in a structure.

```
iex> update_in(report.owner.name, &("Mr. " <> &l))
%BugReport{details: "broken",
  owner: %Customer{company: "PragProg", name: "Mr. Dave"},  

  severity: 1}
```

The other two nested access functions are `get_in` and `get_and_update_in`. The documentation in IEx contains everything you need for these. However, both of these functions support a cool trick: nested access.

Nested Accessors and Nonstructs

If you are using the nested accessor functions with maps or keyword lists, you can supply the keys as atoms:

```
iex> report = %{ owner: %{ name: "Dave", company: "Pragmatic" }, severity: 1}
%{owner: %{company: "Pragmatic", name: "Dave"}, severity: 1}
iex> put_in(report[:owner][:company], "PragProg")
%{owner: %{company: "PragProg", name: "Dave"}, severity: 1}
iex> update_in(report[:owner][:name], &("Mr. " <> &l))
%{owner: %{company: "Pragmatic", name: "Mr. Dave"}, severity: 1}
```

Dynamic (Runtime) Nested Accessors

The nested accessors we've seen so far are macros—they operate at compile time. As a result, they have some limitations:

- The number of keys you pass a particular call is static.
- You can't pass the set of keys as parameters between functions.

These are a natural consequence of the way the macros bake their parameters into code at compile time.

To overcome this, `get_in`, `put_in`, `update_in`, and `get_and_update_in` can all take a list of keys as a separate parameter. Adding this parameter changes them from macros to function calls, so they become dynamic.

	Macro	Function
get_in	<code>no</code>	(dict, keys)
put_in	(path, value)	(dict, keys, value)
update_in	(path, fn)	(dict, keys, fn)
get_and_update_in	(path, fn)	(dict, keys, fn)

Here's a simple example:

```
maps/dynamic_nested.exs
nested = %{
  buttercup: %{
    actor: %{
      first: "Robin",
      last: "Wright"
    },
    role: "princess"
  },
  westley: %{
    actor: %{
      first: "Cary",
      last: "Elwes"      # typo!
    },
    role: "farm boy"
  }
}

IO.inspect get_in(nested, [:buttercup])
# => %{actor: %{first: "Robin", last: "Wright"}, role: "princess"}

IO.inspect get_in(nested, [:buttercup, :actor])
# => %{first: "Robin", last: "Wright"}

IO.inspect get_in(nested, [:buttercup, :actor, :first])
# => "Robin"

IO.inspect put_in(nested, [:westley, :actor, :last], "Elwes")
# => %{buttercup: %{actor: %{first: "Robin", last: "Wright"}, role: "princess"}, westley: %{actor: %{first: "Cary", last: "Elwes"}, role: "farm boy"}}

```

There's a cool trick that the dynamic versions of both `get_in` and `get_and_update_in` support—if you pass a function as a key, that function is invoked to return the corresponding values.

```
maps/get_in_func.exs
authors = [
  %{ name: "José", language: "Elixir" },
  %{ name: "Matz", language: "Ruby" },
  %{ name: "Larry", language: "Perl" }
]

languages_with_an_r = fn (:get, collection, next_fn) ->
  for row <- collection do
    if String.contains?(row.language, "r") do
      next_fn.(row)
    end
  end
end

IO.inspect get_in(authors, [languages_with_an_r, :name])
#=> [ "José", nil, "Larry" ]

```

The Access Module

The Access module provides a number of predefined functions to use as parameters to get and get_and_update_in. These functions act as simple filters while traversing the structures.

The all and at functions only work on lists. all returns all elements in the list, while at returns the n^{th} element (counting from zero).

```
maps/access1.exs
cast = [
  %{
    character: "Buttercup",
    actor: %{
      first: "Robin",
      last: "Wright"
    },
    role: "princess"
  },
  %{
    character: "Westley",
    actor: %{
      first: "Cary",
      last: "Elwes"
    },
    role: "farm boy"
  }
]
IO.inspect get_in(cast, [Access.all(), :character])
#=> ["Buttercup", "Westley"]

IO.inspect get_in(cast, [Access.at(1), :role])
#=> "farm boy"

IO.inspect get_and_update_in(cast, [Access.all(), :actor, :last],
                           fn (val) -> {val, String.upcase(val)} end)
#=> [{"Wright", "Elwes"}, #<-- %{actor: %{first: "Robin", last: "WRIGHT"}, character: "Buttercup", role: "princess"}, #<-- %{actor: %{first: "Cary", last: "ELWES"}, character: "Westley", role: "farm boy"}]}
#       ^-- actor: %{{first: "Robin", last: "WRIGHT"}, character: "Buttercup", role: "princess"}]
```

The elem function works on tuples:

```
maps/access2.exs
cast = [
  %{
    character: "Buttercup",
    actor: {"Robin", "Wright"},
    role: "princess"
  },
```

```
%{
  character: "Westley",
  actor:     {"Carey", "Elwes"},
  role:      "farm boy"
}
]

IO.inspect get_in(cast, [Access.all(), :actor, Access.elem(1)])
#=> ["Wright", "Elwes"]

IO.inspect get_and_update_in(cast, [Access.all(), :actor, Access.elem(1)],
                           fn (val) -> {val, String.reverse(val)} end)
#=> [{"Wright", "Elwes"}, #<%{actor: {"Robin", "thgirW"}, character: "Buttercup", role: "princess"}, #<%{actor: {"Carey", "sewlE"}, character: "Westley", role: "farm boy"}]}
#<%{actor: {"Robin", "Wright"}, character: "Buttercup", role: "princess"}, #<%{actor: {"Carey", "Elwes"}, character: "Westley", role: "farm boy"}]}
```

The key and key! functions work on dictionary types (maps and structs):

```
maps/access3.exs
cast = %{
  buttercup: %{
    actor:     {"Robin", "Wright"},
    role:      "princess"
  },
  westley: %{
    actor:     {"Carey", "Elwes"},
    role:      "farm boy"
  }
}

IO.inspect get_in(cast, [Access.key(:westley), :actor, Access.elem(1)])
#=> "Elwes"

IO.inspect get_and_update_in(cast, [Access.key(:buttercup), :role],
                           fn (val) -> {val, "Queen"} end)
#=> {"princess", #<%{buttercup: %{actor: {"Robin", "Wright"}, role: "Queen"}, westley: %{actor: {"Carey", "Elwes"}, role: "farm boy"}}}
```

Finally, Access.pop lets you remove the entry with a given key from a map or keyword list. It returns a tuple containing the value associated with the key and the updated container. nil is returned for the value if the key isn't in the container.

The name has nothing to do with the pop stack operation.

```
iex> Access.pop(%{name: "Elixir", creator: "Valim"}, :name)
{"Elixir", %{creator: "Valim"}}
iex> Access.pop([name: "Elixir", creator: "Valim"], :name)
{"Elixir", [creator: "Valim"]}
iex> Access.pop(%{name: "Elixir", creator: "Valim"}, :year)
{nil, %{creator: "Valim", name: "Elixir"}}
```

Sets

Sets are implemented using the module `MapSet`.

```
iex> set1 = 1..5 |> Enum.into(MapSet.new)
#MapSet<[1, 2, 3, 4, 5]>
iex> set2 = 3..8 |> Enum.into(MapSet.new)
#MapSet<[3, 4, 5, 6, 7, 8]>
iex> MapSet.member? set1, 3
true
iex> MapSet.union set1, set2
#MapSet<[1, 2, 3, 4, 5, 6, 7, 8]>
iex> MapSet.difference set1, set2
#MapSet<[1, 2]>
iex> MapSet.difference set2, set1
#MapSet<[6, 7, 8]>
iex> MapSet.intersection set2, set1
#MapSet<[3, 4, 5]>
```

With Great Power Comes Great Temptation

The dictionary types are clearly a powerful tool—you’ll use them all the time. But you might also be tempted to abuse them. Structs in particular might lead you into the darkness because you can associate functions with them in their module definitions. At some point, the old object-orientation neurons still active in the nether regions of your brain might burst into life and you might think, “Hey, this is a bit like a class definition.” And you’d be right. You *can* write something akin to object-oriented code using structs (or maps) and modules.

This is a bad idea—not because objects are intrinsically bad, but because you’ll be mixing paradigms and diluting the benefits a functional approach gives you.

Stay pure, young coder. Stay pure.

As a way of refocusing you away from the dark side, the next chapter is a mini diversion into the benefits of separating functions and the data they work on. And we disguise it in a discussion of types.

An Aside—What Are Types?

The preceding two chapters described the basics of lists and maps. But you may have noticed that, although I talked about them as types, I didn't really say what I meant.

The first thing to understand is that the primitive data types are not necessarily the same as the types they can represent. For example, a primitive Elixir list is just an ordered group of values. We can use the [...] literal to create a list, and the | operator to deconstruct and build lists.

Then there's another layer. Elixir has the List module, which provides a set of functions that operate on lists. Often these functions simply use recursion and the | operator to add this extra functionality.

In my mind, there's a difference between the primitive list and the functionality of the List module. The primitive list is an implementation, whereas the List module adds a layer of abstraction. Both implement types, but the type is different. Primitive lists, for example, don't have a flatten function.

Maps are also a primitive type. And, like lists, they have an Elixir module that implements a richer, derived map type.

The code that provides the Keyword type is an Elixir module. But the type is represented as a list of tuples:

```
options = [ {:width, 72}, {:style, "light"}, {:style, "print"} ]
```

Clearly this is still a list, and all the list functions will work on it. But Elixir adds functionality to give you dictionary-like behavior.

```
iex> options = [ {:width, 72}, {:style, "light"}, {:style, "print"} ]
[width: 72, style: "light", style: "print"]
iex> List.last options
{:style, "print"}
iex> Keyword.get_values options, :style
["light", "print"]
```

In a way, this is a form of the duck typing that is talked about in dynamic object-oriented languages.¹ The Keyword module doesn't have an underlying primitive data type. It simply assumes that any value it works on is a list that has been structured a certain way.

This means the APIs for collections in Elixir are fairly broad. Working with a keyword list, you have access to the APIs in the primitive list type, and the List and Keyword modules. You also get Enum and Collectable, which we talk about next.

1. http://en.wikipedia.org/wiki/Duck_ttyping

In this chapter, you'll see:

- The Enum module
- The Stream module
- The Collectable protocol
- Comprehensions

CHAPTER 10

Processing Collections—Enum and Stream

Elixir comes with a number of types that act as collections. We've already seen lists and maps. Ranges, files, and even functions can also act as collections. And as we'll discuss when we look at [protocols on page 329](#), you can also define your own.

Collections differ in their implementation. But they all share something: you can iterate through them. Some of them share an additional trait: you can add things to them.

Technically, things that can be iterated are said to implement the Enumerable protocol.

Elixir provides two modules that have a bunch of iteration functions. The Enum module is the workhorse for collections. You'll use it all the time. I strongly recommend getting to know it.

The Stream module lets you enumerate a collection lazily. This means that the next value is calculated only when it is needed. You'll use this less often, but when you do it's a lifesaver.

I don't want to fill this book with a list of all the APIs. You'll find the definitive (and up-to-date) list online.¹ Instead, I'll illustrate some common uses and let you browse the documentation for yourself. (But please do remember to do so. Much of Elixir's power comes from these libraries.)

Enum—Processing Collections

The Enum module is probably the most used of all the Elixir libraries. Employ it to iterate, filter, combine, split, and otherwise manipulate collections. Here are some common tasks:

1. <http://elixir-lang.org/docs/>

- Convert any collection into a list:

```
iex> list = Enum.to_list 1..5
[1, 2, 3, 4, 5]
```

- Concatenate collections:

```
iex> Enum.concat([1,2,3], [4,5,6])
[1, 2, 3, 4, 5, 6]
iex> Enum.concat [1,2,3], 'abc'
[1, 2, 3, 97, 98, 99]
```

- Create collections whose elements are some function of the original:

```
iex> Enum.map(list, &(&l * 10))
[10, 20, 30, 40, 50]
iex> Enum.map(list, &String.duplicate("*", &l))
["**", "***", "****", "*****", "*****"]
```

- Select elements by position or criteria:

```
iex> Enum.at(10..20, 3)
13
iex> Enum.at(10..20, 20)
nil
iex> Enum.at(10..20, 20, :no_one_here)
:no_one_here
iex> Enum.filter(list, &(&l > 2))
[3, 4, 5]
iex> require Integer      # to get access to is_even
Integer
iex> Enum.filter(list, &Integer.is_even/1)
[2, 4]
iex> Enum.reject(list, &Integer.is_even/1)
[1, 3, 5]
```

- Sort and compare elements:

```
iex> Enum.sort ["there", "was", "a", "crooked", "man"]
["a", "crooked", "man", "there", "was"]
iex> Enum.sort ["there", "was", "a", "crooked", "man"],
...>     &(String.length(&l) <= String.length(&2))
["a", "was", "man", "there", "crooked"]
iex(4)> Enum.max ["there", "was", "a", "crooked", "man"]
"was"
iex(5)> Enum.max_by ["there", "was", "a", "crooked", "man"], &String.length/1
"crooked"
```

- Split a collection:

```
iex> Enum.take(list, 3)
[1, 2, 3]
iex> Enum.take_every list, 2
```

```
[1, 3, 5]
iex> Enum.take_while(list, &(&1 < 4))
[1, 2, 3]
iex> Enum.split(list, 3)
{[1, 2, 3], [4, 5]}
iex> Enum.split_while(list, &(&1 < 4))
{[1, 2, 3], [4, 5]}
```

- Join a collection:

```
iex> Enum.join(list)
"12345"
iex> Enum.join(list, ", ")
"1, 2, 3, 4, 5"
```

- Predicate operations:

```
iex> Enum.all?(list, &(&1 < 4))
false
iex> Enum.any?(list, &(&1 < 4))
true
iex> Enum.member?(list, 4)
true
iex> Enum.empty?(list)
false
```

- Merge collections:

```
iex> Enum.zip(list, [:a, :b, :c])
[{:1, :a}, {:2, :b}, {:3, :c}]
iex> Enum.with_index(["once", "upon", "a", "time"])
[{"once", 0}, {"upon", 1}, {"a", 2}, {"time", 3}]
```

- Fold elements into a single value:

```
iex> Enum.reduce(1..100, &(&1+&2))
5050
iex> Enum.reduce(["now", "is", "the", "time"], fn word, longest ->
...>     if String.length(word) > String.length(longest) do
...>         word
...>     else
...>         longest
...>     end
...>   end)
"time"
iex> Enum.reduce(["now", "is", "the", "time"], 0, fn word, longest ->
...>     if String.length(word) > longest,
...>     do: String.length(word),
...>     else: longest
...>   end)
4
```

- Deal a hand of cards:

```
iex> import Enum
iex> deck = for rank <- '23456789TJQKA', suit <- 'CDHS', do: [suit,rank]
['C2', 'D2', 'H2', 'S2', 'C3', 'D3', ... ]
iex> deck |> shuffle |> take(13)
['DQ', 'S6', 'HJ', 'H4', 'C7', 'D6', 'SJ', 'S9', 'D7', 'HA', 'S4', 'C2', 'CT']
iex> hands = deck |> shuffle |> chunk(13)
[['D8', 'CQ', 'H2', 'H3', 'HK', 'H9', 'DK', 'S9', 'CT', 'ST', 'SK', 'D2', 'HA'],
 ['C5', 'S3', 'CK', 'HQ', 'D3', 'D4', 'CA', 'C8', 'S6', 'DQ', 'H5', 'S2', 'C4'],
 ['C7', 'C6', 'C2', 'D6', 'D7', 'SA', 'SQ', 'H8', 'DT', 'C3', 'H7', 'DA', 'HT'],
 ['S5', 'S4', 'C9', 'S8', 'D5', 'H4', 'S7', 'SJ', 'HJ', 'D9', 'DJ', 'CJ', 'H6']]
```

A Note on Sorting

In our example of sort, we used

```
iex> Enum.sort ["there", "was", "a", "crooked", "man"],
...>     &(String.length(&1) <= String.length(&2))
```

It's important to use `<=` and not just `<` if you want the sort to be *stable*.

Your Turn

► *Exercise: ListsAndRecursion-5*

Implement the following Enum functions using no library functions or list comprehensions: `all?`, `each`, `filter`, `split`, and `take`. You may need to use an `if` statement to implement filter. The syntax for this is

```
if condition do
  expression(s)
else
  expression(s)
end
```

► *Exercise: ListsAndRecursion-6*

(Hard) Write a `flatten(list)` function that takes a list that may contain any number of sublists, which themselves may contain sublists, to any depth. It returns the elements of these lists as a flat list.

```
iex> MyList.flatten([ 1, [ 2, 3, [4] ], 5, [[[6]]]])
```

```
[1,2,3,4,5,6]
```

Hint: You may have to use `Enum.reverse` to get your result in the correct order.

Streams—Lazy Enumerables

In Elixir, the `Enum` module is greedy. This means that when you pass it a collection, it potentially consumes all the contents of that collection. It also means the result will typically be another collection. Look at the following pipeline:

```
enum/pipeline.exs
[ 1, 2, 3, 4, 5 ]
  #=> [ 1, 2, 3, 4, 5 ]
|> Enum.map(&(&1*&&1))
  #=> [ 1, 4, 9, 16, 25 ]
|> Enum.with_index
  #=> [ {1, 0}, {4, 1}, {9, 2}, {16, 3}, {25, 4} ]
|> Enum.map(fn {value, index} -> value - index end)
  #=> [1, 3, 7, 13, 21]
|> IO.inspect          #=> [1, 3, 7, 13, 21]
```

The first `map` function takes the original list and creates a new list of its squares. `with_index` takes this list and returns a list of tuples. The next `map` then subtracts the index from the value, generating a list that gets passed to `IO.inspect`.

So, this pipeline generates four lists on its way to outputting the final result.

Let's look at something different. Here's some code that reads lines from a file and returns the longest:

```
enum/longest_line.exs
IO.puts File.read!("/usr/share/dict/words")
|> String.split
|> Enum.max_by(&String.length/1)
```

In this case, we read the whole dictionary into memory (on my machine that's 2.4MB), then split it into a list of words (236,000 of them) before processing it to find the longest (which happens to be *formaldehydesulphoxylate*).

In both of these examples, our code is suboptimal because each call to `Enum` is self-contained. Each call takes a collection and returns a collection.

What we really want is to process the elements in the collection as we need them. We don't need to store intermediate results as full collections; we just need to pass the current element from function to function. And that's what streams do.

A Stream Is a Composable Enumerator

Here's a simple example of creating a Stream:

```
iex> s = Stream.map [1, 3, 5, 7], &(&1 + 1)
#Stream<[enum: [1, 3, 5, 7], funs: [#Function<46.3851/1 in Stream.map/2>] ]>
```

If we'd called `Enum.map`, we'd have seen the result `[2,4,6,8]` come back immediately. Instead we get back a stream value that contains a specification of what we intended.

How do we get the stream to start giving us results? Treat it as a collection and pass it to a function in the `Enum` module:

```
iex> s = Stream.map [1, 3, 5, 7], &(&1 + 1)
#Stream<[enum: [1, 3, 5, 7], funs: [#Function<46.3851/1 in Stream.map/2>] ]>
iex> Enum.to_list s
[2, 4, 6, 8]
```

Because streams are enumerable, you can also pass a stream to a stream function. Because of this, we say that streams are *composable*.

```
iex> squares = Stream.map [1, 2, 3, 4], &(&1*&1)
#Stream<[enum: [1, 2, 3, 4],
          funs: [#Function<32.133702391 in Stream.map/2>] ]>
iex> plus_ones = Stream.map squares, &(&1+1)
#Stream<[enum: [1, 2, 3, 4],
          funs: [#Function<32.133702391 in Stream.map/2>,
                 #Function<32.133702391 in Stream.map/2>] ]>
iex> odds = Stream.filter plus_ones, fn x -> rem(x,2) == 1 end
#Stream<[enum: [1, 2, 3, 4],
          funs: [#Function<26.133702391 in Stream.filter/2>,
                 #Function<32.133702391 in Stream.map/2>,
                 #Function<32.133702391 in Stream.map/2>] ]>
iex> Enum.to_list odds
[5, 17]
```

Of course, in real life we'd have written this as

```
enum/stream1.exs
[1,2,3,4]
|> Stream.map(&(&1*&1))
|> Stream.map(&(&1+1))
|> Stream.filter(fn x -> rem(x,2) == 1 end)
|> Enum.to_list
```

Note that we're never creating intermediate lists—we're just passing successive elements of each of the collections to the next in the chain. The `Stream` values shown in the previous IEx session give a hint of how this works—chained streams are represented as a list of functions, each of which is applied in turn to each element of the stream as it is processed.

Streams aren't only for lists. More and more Elixir modules now support streams. For example, here's our longest-word code written using streams:

enum/stream2.exs

```
Io.puts File.open!("/usr/share/dict/words")
    |> Io.stream(:line)
    |> Enum.max_by(&String.length/1)
```

The magic here is the call to `Io.stream`, which converts an IO device (in this case the open file) into a stream that serves one line at a time. In fact, this is such a useful concept that there's a shortcut:

enum/stream3.exs

```
Io.puts File.stream!("/usr/share/dict/words") |> Enum.max_by(&String.length/1)
```

The good news is that there is no intermediate storage. The bad news is that it runs about two times slower than the previous version. However, consider the case where we were reading data from a remote server or from an external sensor (maybe temperature readings). Successive lines might arrive slowly, and they might go on for ever. With the `Enum` implementation we'd have to wait for all the lines to arrive before we started processing. With streams we can process them as they arrive.

Infinite Streams

Because streams are lazy, there's no need for the whole collection to be available up front. For example, if I write

```
iex> Enum.map(1..10_000_000, &(&l+1)) |> Enum.take(5)
[2, 3, 4, 5, 6]
```

it takes about 8 seconds before I see the result. Elixir is creating a 10-million-element list, then taking the first five elements from it. If instead I write

```
iex> Stream.map(1..10_000_000, &(&l+1)) |> Enum.take(5)
[2, 3, 4, 5, 6]
```

the result comes back instantaneously. The `take` call just needs five values, which it gets from the stream. Once it has them, there's no more processing.

In these examples the stream is bounded, but it can equally well go on forever. To do that, we'll need to create streams based on functions.

Creating Your Own Streams

Streams are implemented solely in Elixir libraries—there is no specific runtime support. However, this doesn't mean you want to drop down to the very lowest level and create your own streamable types. The actual implementation is complex (in the same way that string theory and dating rituals are complex). Instead, you probably want to use some helpful wrapper functions to do the heavy lifting. There are a number of these, including `cycle`, `repeatedly`, `iterate`,

unfold, and resource. (If you needed proof that the internal implementation is tricky, consider the fact that these last two names give you almost no hint of their power.)

Let's start with the three simplest: cycle, repeatedly, and iterate.

Stream.cycle

Stream.cycle takes an enumerable and returns an infinite stream containing that enumerable's elements. When it gets to the end, it repeats from the beginning, indefinitely. Here's an example that generates the rows in an HTML table with alternating *green* and *white* classes:

```
iex> Stream.cycle(~w{ green white }) |>
...> Stream.zip(1..5) |>
...> Enum.map(fn {class, value} ->
...>     "<tr class='#{class}'><td>#{value}</td></tr>|n" end) |>
...> IO.puts
<tr class="green"><td>1</td></tr>
<tr class="white"><td>2</td></tr>
<tr class="green"><td>3</td></tr>
<tr class="white"><td>4</td></tr>
<tr class="green"><td>5</td></tr>
:ok
```

Stream.repeatedly

Stream.repeatedly takes a function and invokes it each time a new value is wanted.

```
iex> Stream.repeatedly(fn -> true end) |> Enum.take(3)
[true, true, true]
iex> Stream.repeatedly(&:random.uniform/0) |> Enum.take(3)
[0.7230402056221108, 0.94581636451987, 0.5014907142064751]
```

Stream.iterate

Stream.iterate(start_value, next_fun) generates an infinite stream. The first value is start_value. The next value is generated by applying next_fun to this value. This continues for as long as the stream is being used, with each value being the result of applying next_fun to the previous value.

Here are some examples:

```
iex> Stream.iterate(0, &(&1+1)) |> Enum.take(5)
[0, 1, 2, 3, 4]
iex> Stream.iterate(2, &(&1*&1)) |> Enum.take(5)
[2, 4, 16, 256, 65536]
iex> Stream.iterate([], &[&1]) |> Enum.take(5)
[], [[], [[[[]]]], [[[[[[[]]]]]]]]
```

Stream.unfold

Now we can get a little more adventurous. Stream.unfold is related to iterate, but you can be more explicit both about the values output to the stream and about the values passed to the next iteration. You supply an initial value and a function. The function uses the argument to create two values, returned as a tuple. The first is the value to be returned by this iteration of the stream, and the second is the value to be passed to the function on the next iteration of the stream. If the function returns nil, the stream terminates.

This sounds abstract, but unfold is quite useful—it is a general way of creating a potentially infinite stream of values where each value is some function of the previous state.

The key is the generating function. Its general form is

```
fn state -> { stream_value, new_state } end
```

For example, here's a stream of Fibonacci numbers:

```
iex> Stream.unfold({0,1}, fn {f1,f2} -> {f1, {f2, f1+f2}} end) |> Enum.take(15)
[0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377]
```

Here the *state* is a tuple containing the current and the next number in the sequence. We seed it with the initial state of {0, 1}. The value each iteration of the stream returns is the first of the state values. The new state moves one down the sequence, so an initial state of {f1,f2} becomes a new state of {f2,f1+f2}.

Stream.resource

At this point you might be wondering how streams can interact with external resources. We've already seen how you can turn a file's contents into a stream of lines, but how could you implement this yourself? You'd need to open the file when the stream first starts, return successive lines, and then close the file at the end. Or maybe you want to turn a database result-set cursor into a stream of values. You'd have to execute the query when the stream starts, return each row as stream values, and close the query at the end. And that's where Stream.resource comes in.

Stream.resource builds upon Stream.unfold. It makes two changes.

The first argument to unfold is the initial value to be passed to the iteration function. But if that value is a resource, we don't want to open it until the stream starts delivering values, and that might not happen until long after we create the stream. To get around this, resource takes not a value, but a function that returns the value. That's the first change.

Second, when the stream is done with the resource, we may need to close it. That's what the third argument to Stream.resource does—it takes the final accumulator value and does whatever is needed to deallocate the resource.

Here's an example from the library documentation:

```
Stream.resource(fn -> File.open!("sample") end,
  fn file ->
    case IO.read(file, :line) do
      data when is_binary(data) -> {[data], file}
      _ -> {[:halt], file}
    end
  end,
  fn file -> File.close(file) end)
```

The first function opens the file when the stream becomes active, and passes it to the second function. This reads the file, line by line, returning either a line and the file as a tuple, or a `:halt` tuple at the end of the file. The third function closes the file.

Let's finish with a different kind of resource: time. We'll implement a timer that counts down the number of seconds until the start of the next minute. It uses a stream resource to do this. The allocation function returns the number of seconds left until the next minute starts. It does this each time the stream is evaluated, so we'll get a countdown that varies depending on when it is called.

The iteration function looks at the time left. If zero, it returns `{:halt, 0}`; otherwise it sleeps for a second and returns the current countdown as a string, along with the decremented counter.

In this case there's no resource deallocation, so the third function does nothing.

Here's the code:

```
enum/countdown.exs
defmodule Countdown do
  def sleep(seconds) do
    receive do
      after seconds*1000 -> nil
    end
  end

  def say(text) do
    spawn fn -> :os.cmd('say #{text}') end
  end
```

```

def timer do
  Stream.resource(
    fn ->           # the number of seconds to the start of the next minute
      {_h,_m,s} = :erlang.time
      60 - s - 1
    end,
    fn             # wait for the next second, then return its countdown
      0 ->
      {:halt, 0}
    count ->
      sleep(1)
      { [inspect(count)], count - 1 }
    end,
    fn _ -> nil end  # nothing to deallocate
  )
end
end

```

(The eagle-eyed among you will have noticed a function called `say` in the `Countdown` module. This executes the shell command `say`, which, on OS X, speaks its argument. You could substitute `espeak` on Linux and `ptts` on Windows.)

Let's play with the code.

```

$ iex countdown.exs
iex> counter = Countdown.timer
#Function<17.133702391/2 in Stream.resource/3>

iex> printer = counter |> Stream.each(&IO.puts/1)
#Stream[enum: #Function<17.133702391/2 in Stream.resource/3>,
  funs: [#Function<0.133702391/1 in Stream.each/2>] >

iex> speaker = printer |> Stream.each(&Countdown.say/1)
#Stream[enum: #Function<17.133702391/2 in Stream.resource/3>,
  funs: [#Function<0.133702391/1 in Stream.each/2>,
  #Function<0.133702391/1 in Stream.each/2>] >

```

So far, we've built a stream that creates time events, prints the countdown value, and speaks it. But there's been no output, as we haven't yet asked the stream for any values. Let's do that now:

```

iex> speaker |> Enum.take(5)
37    ** numbers are output once
36    ** per second. Even cooler, the
35    ** computer says
34    ** "thirty seven", "thirty six"...
33
["37", "36", "35", "34", "33"]

```

Cool—we must have started it around 22 seconds into a minute, so the countdown starts at 37. Let's use the same stream again, a few seconds later:

```
iex> speaker |> Enum.take(3)
29
28
27
["29", "28", "27"]
```

Wait some more seconds, and this time let it run to the top of the minute:

```
iex> speaker |> Enum.to_list
6
5
4
3
2
1
["6", "5", "4", "3", "2", "1"]
```

This is clearly not great code, as it fails to correct the sleep time for any delays introduced by our code. But it illustrates a very cool point. Lazy streams let you deal with resources that are asynchronous to your code, and the fact that they are initialized every time they are used means they're effectively side effect-free. Every time we pipe our stream to an `Enum` function, we get a fresh set of values, computed at that time.

Streams in Practice

In the same way that functional programming requires you to look at problems in a new way, streams ask you to look at iteration and collections afresh. Not every situation where you're iterating requires a stream. But consider using a stream when you want to defer processing until you need the data, and when you need to deal with large numbers of things without necessarily generating them all at once.

The Collectable Protocol

The `Enumerable` protocol lets you iterate over the elements in a type—given a collection, you can get the elements. `Collectable` is in some sense the opposite—it allows you to build a collection by inserting elements into it.

Not all collections are collectable. Ranges, for example, cannot have new entries added to them.

The collectable API is pretty low-level, so you'll typically access it via `Enum.into` and when using comprehensions (which we cover in the next section). For example, we can inject the elements of a range into an empty list using

```
iex> Enum.into 1..5, []
[1, 2, 3, 4, 5]
```

If the list is not empty, the new elements are tacked onto the end:

```
iex> Enum.into 1..5, [ 100, 101 ]
[100, 101, 1, 2, 3, 4, 5]
```

Output streams are collectable, so the following code lazily copies standard input to standard output:

```
iex> Enum.into IO.stream(:stdio, :line), IO.stream(:stdio, :line)
```

Comprehensions

When you're writing functional code, you often map and filter collections of things. To make your life easier (and your code easier to read), Elixir provides a general-purpose shortcut for this: the *comprehension*.

The idea of a comprehension is fairly simple: given one or more collections, extract all combinations of values from each, optionally filter the values, and then generate a new collection using the values that remain.

The general syntax for comprehensions is deceptively simple:

`result = for generator or filter... [, into: value], do: expression`

Let's see a couple of basic examples before we get into the details.

```
iex> for x <- [ 1, 2, 3, 4, 5 ], do: x * x
[1, 4, 9, 16, 25]
iex> for x <- [ 1, 2, 3, 4, 5 ], x < 4, do: x * x
[1, 4, 9]
```

A generator specifies how you want to extract values from a collection.

`pattern <- enumerable_thing`

Any variables matched in the pattern are available in the rest of the comprehension (including the block). For example, `x <- [1,2,3]` says that we want to first run the rest of the comprehension with `x` set to 1. Then we run it with `x` set to 2, and so on. If we have two generators, their operations are nested, so

```
x <- [1,2], y <- [5,6]
```

will run the rest of the comprehension with `x=1, y=5; x=1, y=6; x=2, y=5; and x=2, y=6`. We can use those values of `x` and `y` in the `do` block:

```
iex> for x <- [1,2], y <- [5,6], do: x * y
[5, 6, 10, 12]
iex> for x <- [1,2], y <- [5,6], do: {x, y}
[{:1, 5}, {:1, 6}, {:2, 5}, {:2, 6}]
```

You can use variables from generators in later generators:

```
iex> min_maxes = [{1,4}, {2,3}, {10, 15}]
[{:1, 4}, {:2, 3}, {:10, 15}]
iex> for {min,max} <- min_maxes, n <- min..max, do: n
[1, 2, 3, 4, 2, 3, 10, 11, 12, 13, 14, 15]
```

A filter is a predicate. It acts as a gatekeeper for the rest of the comprehension—if the condition is false, then the comprehension moves on to the next iteration without generating an output value.

For example, the code that follows uses a comprehension to list pairs of numbers from 1 to 8 whose product is a multiple of 10. It uses two generators (to cycle through the pairs of numbers) and two filters. The first filter allows only pairs in which the first number is at least the value of the second. The second filter checks to see if the product is a multiple of 10.

```
iex> first8 = [ 1,2,3,4,5,6,7,8 ]
[1, 2, 3, 4, 5, 6, 7, 8]
iex> for x <- first8, y <- first8, x >= y, rem(x*y, 10)==0, do: { x, y }
[{:5, 2}, {:5, 4}, {:6, 5}, {:8, 5}]
```

This comprehension iterates 64 times, with $x=1, y=1$; $x=1, y=2$; and so on. However, the first filter cuts the iteration short when x is less than y . This means the second filter runs only 36 times.

Because the first term in a generator is a pattern, we can use it to deconstruct structured data. Here's a comprehension that swaps the keys and values in a keyword list.

```
iex> reports = [ dallas: :hot, minneapolis: :cold, dc: :muggy, la: :smoggy ]
[dallas: :hot, minneapolis: :cold, dc: :muggy, la: :smoggy]
iex> for { city, weather } <- reports, do: { weather, city }
[hot: :dallas, cold: :minneapolis, muggy: :dc, smoggy: :la]
```

Comprehensions Work on Bits, Too

A bitstring (and, by extension, a binary or a string) is simply a collection of ones and zeroes. So it's probably no surprise that comprehensions work on bits, too. What might be surprising is the syntax:

```
iex> for << ch <- "hello" >>, do: ch
'hello'
iex> for << ch <- "hello" >>, do: <<ch>>
["h", "e", "l", "l", "o"]
```

Here the generator is enclosed in << and >>, indicating a binary. In the first case, the do block returns the integer code for each character, so the resulting list is [104, 101, 108, 108, 111], which IEx displays as 'hello'.

In the second case, we convert the code back into a string, and the result is a list of those one-character strings.

Again, the thing to the left of the <- is a pattern, and so we can use binary pattern matching. Let's convert a string into the octal representation of its characters:

```
iex> for << b1::size(2), b2::size(3), b3::size(3) >> <- "hello" >>,
...> do: "0#{b1}#{b2}#{b3}"
["0150", "0145", "0154", "0154", "0157"]
```

Scoping and Comprehensions

All variable assignments inside a comprehension are local to that comprehension—you will not affect the value of a variable in the outer scope.

```
iex> name = "Dave"
Dave
iex> for name <- [ "cat", "dog" ], do: String.upcase(name)
["CAT", "DOG"]
iex> name
Dave
iex>
```

The Value Returned by a Comprehension

In our examples thus far, the comprehension has returned a list. The list contains the values returned by the do expression for each iteration of the comprehension.

This behavior can be changed with the into: parameter. This takes a collection that is to receive the results of the comprehension. For example, we can populate a map using

```
iex> for x <- ~w{ cat dog }, into: %{}, do: { x, String.upcase(x) }
%{"cat" => "CAT", "dog" => "DOG"}
```

It might be more clear to use Map.new in this case:

```
iex> for x <- ~w{ cat dog }, into: Map.new, do: { x, String.upcase(x) }
%{"cat" => "CAT", "dog" => "DOG"}
```

The collection doesn't have to be empty:

```
iex> for x <- ~w{ cat dog }, into: %{"ant" => "ANT"}, do: { x, String.upcase(x) }
%{"ant" => "ANT", "cat" => "CAT", "dog" => "DOG"}
```

In [Chapter 24, Protocols—Polymorphic Functions, on page 329](#), we'll look at protocols, which let us specify common behaviors across different types. The `into:` option takes values that implement the `Collectable` protocol. These include lists, binaries, functions, maps, files, hash dicts, hash sets, and IO streams, so we can write things such as

```
iex> for x <- ~w{ cat dog }, into: IO.stream(:stdio,:line), do: "<>#{x}>\n"
<<cat>>
<<dog>>
%IO.Stream{device: :standard_io, line_or_bytes: :line, raw: false}
```

Your Turn

► *Exercise: ListsAndRecursion-7*

In the last exercise of [Chapter 7, Lists and Recursion, on page 71](#), you wrote a `span` function. Use it and list comprehensions to return a list of the prime numbers from 2 to n .

► *Exercise: ListsAndRecursion-8*

The Pragmatic Bookshelf has offices in Texas (TX) and North Carolina (NC), so we have to charge sales tax on orders shipped to these states. The rates can be expressed as a keyword list (I wish it were that simple....):

```
tax_rates = [ NC: 0.075, TX: 0.08 ]
```

Here's a list of orders:

```
orders = [
  [ id: 123, ship_to: :NC, net_amount: 100.00 ],
  [ id: 124, ship_to: :OK, net_amount: 35.50 ],
  [ id: 125, ship_to: :TX, net_amount: 24.00 ],
  [ id: 126, ship_to: :TX, net_amount: 44.80 ],
  [ id: 127, ship_to: :NC, net_amount: 25.00 ],
  [ id: 128, ship_to: :MA, net_amount: 10.00 ],
  [ id: 129, ship_to: :CA, net_amount: 102.00 ],
  [ id: 130, ship_to: :NC, net_amount: 50.00 ] ]
```

Write a function that takes both lists and returns a copy of the orders, but with an extra field, `total_amount`, which is the net plus sales tax. If a shipment is not to NC or TX, there's no tax applied.

Moving Past Divinity

L. Peter Deutsch once penned, “To iterate is human, to recurse divine.” And that’s certainly the way I felt when I first started coding Elixir. The joy of pattern-matching lists in sets of recursive functions drove my designs. After a while, I realized that perhaps I was taking this too far.

In reality, most of our day-to-day work is better handled using the various enumerators built into Elixir. They make your code smaller, easier to understand, and probably more efficient.

Part of the process of learning to be effective in Elixir is working out for yourself when to use recursion and when to use enumerators. I recommend enumerating when you can.

Next we'll look at string handling in Elixir (and Erlang).

In this chapter, you'll see:

- Strings and string literals
- Character lists (single-quoted literals)
- Pattern matching and processing strings

CHAPTER 11

Strings and Binaries

We've been using strings without really discussing them. Let's rectify that.

String Literals

Elixir has two kinds of string: single-quoted and double-quoted. They differ significantly in their internal representation. But they also have many things in common.

- Strings can hold characters in UTF-8 encoding.
- They may contain escape sequences:

\a	BEL (0x07)	\b	BS (0x08)	\d	DEL (0x7f)
\e	ESC (0x1b)	\f	FF (0x0c)	\n	NL (0x0a)
\r	CR (0x0d)	\s	SP (0x20)	\t	TAB (0x09)
\v	VT (0x0b)	\uhhh	1–6 hex digits	\xhh	2 hex digits

- They allow interpolation on Elixir expressions using the syntax `#{}:`

```
iex> name = "dave"
"dave"
iex> "Hello, #{String.capitalize name}!"
"Hello, Dave!"
```

- Characters that would otherwise have special meaning can be escaped with a backslash.
- They support *heredocs*.

Heredocs

Any string can span several lines. To illustrate this, we'll use both `IO.puts` and `IO.write`. We use `write` for the multiline string because `puts` always appends a newline, and we want to see the contents without this.

```
IO.puts "start"
IO.write "
  my
  string
"
IO.puts "end"

produces

start
  my
  string
end
```

Notice how the multiline string retains the leading and trailing newlines and the leading spaces on the intermediate lines.

The *heredoc* notation fixes this. Triple the string delimiter (" or "") and indent the trailing delimiter to the same margin as your string contents, and you get this:

```
IO.puts "start"
IO.write """
  my
  string
"""
IO.puts "end"

which produces

start
my
string
end
```

Heredocs are used extensively to add documentation to functions and modules.

Sigils

Like Ruby, Elixir has an alternative syntax for some literals. We've already seen it with regular expressions, where we wrote `~r{...}`. In Elixir, these `~-`-style literals are called *sigils* (symbols with magical powers).

A sigil starts with a tilde, followed by an upper- or lowercase letter, some delimited content, and perhaps some options. The delimiters can be `<...>`, `{...}`, `[...]`, `(...)`, `|...|`, `/.../`, `"..."`, and `'...'`.

The letter determines the sigil's type:

- `~C` A character list with no escaping or interpolation
- `~c` A character list, escaped and interpolated just like a single-quoted string
- `~D` A Date in the format `yyyy-mm-dd`
- `~N` A naive (raw) `DateTime` in the format `yyyy-mm-dd hh:mm:ss[.ddd]`
- `~R` A regular expression with no escaping or interpolation
- `~r` A regular expression, escaped and interpolated
- `~S` A string with no escaping or interpolation
- `~s` A string, escaped and interpolated just like a double-quoted string
- `~T` A Time in the format `hh:mm:ss[.dddd]`
- `~W` A list of whitespace-delimited words, with no escaping or interpolation
- `~w` A list of whitespace-delimited words, with escaping and interpolation

Here are some examples of sigils, using a variety of delimiters:

```
iex> ~C[1\n2#{1+2}]
'1\\n2#{1+2}'
iex> ~c"1\n2#{1+2}"
'1\n23'
iex> ~S[1\n2#{1+2}]
"1\\n2#{1+2}"
iex> ~s/1\n2#{1+2}/
"1\n23"
iex> ~W[the c#{'a'}t sat on the mat]
["the", "c#{'a'}t", "sat", "on", "the", "mat"]
iex> ~w[the c#{'a'}t sat on the mat]
["the", "cat", "sat", "on", "the", "mat"]
iex> ~D<1999-12-31>
~D[1999-12-31]
iex> ~T[12:34:56]
~T[12:34:56]
iex> ~N{1999-12-31 23:59:59}
~N[1999-12-31 23:59:59]
```

The `~W` and `~w` sigils take an optional type specifier, `a`, `c`, or `s`, which determines whether it returns a list of atoms, character lists, or strings. (We've already seen the `~r` options.)

```
iex> ~w[the c#{'a'}t sat on the mat]a
[:the, :cat, :sat, :on, :the, :mat]
iex> ~w[the c#{'a'}t sat on the mat]c
['the', 'cat', 'sat', 'on', 'the', 'mat']
iex> ~w[the c#{'a'}t sat on the mat]s
["the", "cat", "sat", "on", "the", "mat"]
```

The delimiter can be any nonword character. If it is (, [, {, or <, then the terminating delimiter is the corresponding closing character. Otherwise the terminating delimiter is the next nonescaped occurrence of the opening delimiter.

Elixir does not check the nesting of delimiters, so the sigil `~s{a{b}}` is the three-character string `a{b}`.

If the opening delimiter is three single or three double quotes, the sigil is treated as a heredoc.

```
iex> ~w"""
...> the
...> cat
...> sat
...> """
...> """
["the", "cat", "sat"]
```

If you want to specify modifiers with heredoc sigils (most commonly you'd do this with `~r`), add them after the trailing delimiter.

```
iex> ~r"""
...> hello
...> """
...> """
...> i
~r/hello\n/i
```

One of the interesting things about sigils is that you can define your own. We talk about this [in Part III, on page 347](#).

The Name “strings”

Before we get further into this, I need to explain something. In most other languages, you'd call both 'cat' and "cat" *strings*. And that's what I've been doing so far. But Elixir has a different convention.

In Elixir, the convention is that we call only double-quoted strings “strings.” The single-quoted form is a character list.

This is important. The single- and double-quoted forms are very different, and libraries that work on strings work only on the double-quoted form.

Let's explore the differences in more detail.

Single-Quoted Strings—Lists of Character Codes

Single-quoted strings are represented as a list of integer values, each value corresponding to a codepoint in the string. For this reason, we refer to them as *character lists* (or *char lists*).

```
iex> str = 'wombat'
'wombat'
iex> is_list str
true
iex> length str
6
iex> Enum.reverse str
'tabmow'
```

This is confusing: IEx *says* it is a list, but it shows the value as a string. That's because IEx prints a list of integers as a string if it believes each number in the list is a printable character. You can try this for yourself:

```
iex> [ 67, 65, 84 ]
'CAT'
```

You can look at the internal representation in a number of ways:

```
iex> str = 'wombat'
'wombat'
iex> :io.format "~w~n", [ str ]
[119,111,109,98,97,116]
:ok
iex> List.to_tuple str
{119, 111, 109, 98, 97, 116}
iex> str ++ [0]
[119, 111, 109, 98, 97, 116, 0]
```

The `~w` in the format string forces `str` to be written as an Erlang term—the underlying list of integers. The `~n` is a newline.

The last example creates a new character list with a null byte at the end. IEx no longer thinks all the bytes are printable, and so returns the underlying character codes.

If a character list contains characters Erlang considers nonprintable, you'll see the list representation.

```
iex> 'ðx/ðy'
[8706, 120, 47, 8706, 121]
```

Because a character list is a list, we can use the usual pattern matching and List functions.

```
iex> 'pole' ++ 'vault'
'polevault'
iex> 'pole' -- 'vault'
'poe'
iex> List.zip [ 'abc', '123' ]
[{:97, 49}, {:98, 50}, {:99, 51}]
iex> [ head | tail ] = 'cat'
'cat'
iex> head
99
iex> tail
'at'
iex> [ head | tail ]
'cat'
```

Why is the head of 'cat' 99 and not *c*? Remember that a char list is just a list of integer character codes, so each individual entry is a number. It happens that 99 is the code for a lowercase *c*.

In fact, the notation `?c` returns the integer code for the character *c*. This is often useful when employing patterns to extract information from character lists. Here's a simple module that parses the character-list representation of an optionally signed decimal number.

```
strings/parse.exs
defmodule Parse do
  def number([ ?- | tail ]), do: _number_digits(tail, 0) * -1
  def number([ ?+ | tail ]), do: _number_digits(tail, 0)
  def number(str),           do: _number_digits(str, 0)

  defp _number_digits([], value), do: value
  defp _number_digits([ digit | tail ], value)
    when digit in '0123456789' do
      _number_digits(tail, value*10 + digit - ?0)
    end
  defp _number_digits([ non_digit | _ ], _) do
    raise "Invalid digit '#{{non_digit}}'"
  end
end
```

Let's try it in IEx.

```
iex> c("parse.exs")
[Parse]
iex> Parse.number('123')
123
iex> Parse.number('-123')
-123
iex> Parse.number('+123')
123
```

```
iex> Parse.number('+9')
9
iex> Parse.number('+a')
** (RuntimeError) Invalid digit 'a'
```

Your Turn

► *Exercise: StringsAndBinaries-1*

Write a function that returns true if a single-quoted string contains only printable ASCII characters (space through tilde).

► *Exercise: StringsAndBinaries-2*

Write an anagram?(word1, word2) that returns true if its parameters are anagrams.

► *Exercise: StringsAndBinaries-3*

Try the following in IEx:

```
iex> [ 'cat' | 'dog' ]
['cat',100,111,103]
```

Why does IEx print 'cat' as a string, but 'dog' as individual numbers?

► *Exercise: StringsAndBinaries-4*

(Hard) Write a function that takes a single-quoted string of the form *number [+/*/] number* and returns the result of the calculation. The individual numbers do not have leading plus or minus signs.

```
calculate('123 + 27') # => 150
```

Binaries

The binary type represents a sequence of bits.

A binary literal looks like `<< term,... >>`.

The simplest term is just a number from 0 to 255. The numbers are stored as successive bytes in the binary.

```
iex> b = << 1, 2, 3 >>
<<1, 2, 3>>
iex> byte_size b
3
iex> bit_size b
24
```

You can specify modifiers to set any term's size (in bits). This is useful when working with binary formats such as media files and network packets.

```
iex> b = << 1::size(2), 1::size(3) >>      # 01 001
<<9::size(5)>>                           # = 9 (base 10)
iex> byte_size b
1
iex> bit_size b
5
```

You can store integers, floats, and other binaries in binaries.

```
iex> int = << 1 >>
<<1>>
iex> float = << 2.5 :: float >>
<<64, 4, 0, 0, 0, 0, 0, 0>>
iex> mix = << int :: binary, float :: binary >>
<<1, 64, 4, 0, 0, 0, 0, 0>>
```

Let's finish an initial look at binaries with an example of bit extraction. An IEEE 754 float has a sign bit, 11 bits of exponent, and 52 bits of mantissa. The exponent is biased by 1023, and the mantissa is a fraction with the top bit assumed to be 1. So we can extract the fields and then use `:math.pow`, which performs exponentiation, to reassemble the number:

```
iex> << sign::size(1), exp::size(11), mantissa::size(52) >> = << 3.14159::float >>
iex> (1 + mantissa / :math.pow(2, 52)) * :math.pow(2, exp-1023) * (1 - 2*sign)
3.14159
```

Double-Quoted Strings Are Binaries

Whereas single-quoted strings are stored as char lists, the contents of a double-quoted string (`dqs`) are stored as a consecutive sequence of bytes in UTF-8 encoding. Clearly this is more efficient in terms of memory and certain forms of access, but it does have two implications.

First, because UTF-8 characters can take more than a single byte to represent, the size of the binary is not necessarily the length of the string.

```
iex> dqs = "ðx/ðy"
"ðx/ðy"
iex> String.length dqs
5
iex> byte_size dqs
9
iex> String.at(dqs, 0)
"ð"
iex> String.codepoints(dqs)
[ "ð", "x", "/", "ð", "y" ]
iex> String.split(dqs, "/")
[ "ðx", "ðy" ]
```

Second, because you're no longer using lists, you need to learn and work with the binary syntax alongside the list syntax in your code.

Strings and Elixir Libraries

When Elixir library documentation uses the word *string* (and most of the time it uses the word *binary*), it means double-quoted strings.

The String module defines functions that work with double-quoted strings.

`at(str, offset)`

Returns the grapheme at the given offset (starting at 0). Negative offsets count from the end of the string.

```
iex> String.at("dog", 0)
"ð"
iex> String.at("dog", -1)
"g"
```

`capitalize(str)`

Converts str to lowercase, and then capitalizes the first character.

```
iex> String.capitalize "école"
"École"
iex> String.capitalize "íííííí"
"Íííííí"
```

`codepoints(str)`

Returns the codepoints in str.

```
iex> String.codepoints("José's øg")
["J", "o", "s", "é", "'", "s", " ", "ø", "ø", "g"]
```

`downcase(str)`

Converts str to lowercase.

```
iex> String.downcase "ØRSted"
"ørsted"
```

`duplicate(str, n)`

Returns a string containing n copies of str.

```
iex> String.duplicate "Ho! ", 3
" Ho! Ho! Ho! "
```

`ends_with?(str, suffix | [suffixes])`

Returns true if str ends with any of the given suffixes.

```
iex> String.ends_with? "string", ["elix", "stri", "ring"]
true
```

first(str)

Returns the first grapheme from str.

```
iex> String.first "õog"
"õ"
```

graphemes(str)

Returns the graphemes in the string. This is different from the codepoints function, which lists combining characters separately. The following example uses a combining diaeresis along with the letter e to represent é. (It might not display properly on your ereader.)

```
iex> String.codepoints "noe\u0308l"
["n", "o", "e", "", "l"]
iex> String.graphemes "noe\u0308l"
["n", "o", "e", "é", "l"]
```

jaro_distance

Returns a float between 0 and 1 indicating the likely similarity of two strings.

```
iex> String.jaro_distance("jonathan", "jonathon")
0.9166666666666666
iex> String.jaro_distance("josé", "john")
0.6666666666666666
```

last(str)

Returns the last grapheme from str.

```
iex> String.last "õog"
"g"
```

length(str)

Returns the number of graphemes in str.

```
iex> String.length "õx/õy"
5
```

myers_difference

Returns the list of transformations needed to convert one string to another.

```
iex> String.myers_difference("banana", "panama")
[del: "b", ins: "p", eq: "ana", del: "n", ins: "m", eq: "a"]
```

`next_codepoint(str)`

Splits str into its leading codepoint and the rest, or nil if str is empty. This may be used as the basis of an iterator.

```
defmodule MyString do
  def each(str, func), do: _each(String.next_codepoint(str), func)
  defp _each({codepoint, rest}, func) do
    func.(codepoint)
    _each(String.next_codepoint(rest), func)
  end
  defp _each(nil, _), do: []
end

MyString.each "dog", fn c -> IO.puts c end
```

produces

```
ð
o
g
```

`next_grapheme(str)`

Same as `next_codepoint`, but returns graphemes (`:no_grapheme` on completion).

`pad_leading(str, new_length, padding || "")`

Returns a new string, at least `new_length` characters long, containing `str` right-justified and padded with `padding`.

```
iex> String.pad_leading("cat", 5, ">")
">>cat"
```

`pad_trailing(str, new_length, padding || "")`

Returns a new string, at least `new_length` characters long, containing `str` left-justified and padded with `padding`.

```
iex> String.pad_trailing("cat", 5)
"cat "
```

`printable?(str)`

Returns true if `str` contains only printable characters.

```
iex> String.printable? "José"
true
iex> String.printable? "|x00 a null"
false
```

`replace(str, pattern, replacement, options || [global: true, insert_replaced: nil])`

Replaces pattern with replacement in str under control of options.

If the :global option is true, all occurrences of the pattern are replaced; otherwise only the first is replaced.

If :insert_replaced is a number, the pattern is inserted into the replacement at that offset. If the option is a list, it is inserted multiple times.

```
iex> String.replace "the cat on the mat", "at", "AT"
"the cAT on the mAT"
iex> String.replace "the cat on the mat", "at", "AT", global: false
"the cAT on the mat"
iex> String.replace "the cat on the mat", "at", "AT", insert_replaced: 0
"the catAT on the matAT"
iex> String.replace "the cat on the mat", "at", "AT", insert_replaced: [0,2]
"the catATat on the matATat"
```

`reverse(str)`

Reverses the graphemes in a string.

```
iex> String.reverse "pupils"
"slipup"
iex> String.reverse "Σf÷∂"
"∂÷fΣ"
```

`slice(str, offset, len)`

Returns a len character substring starting at offset (measured from the end of str if negative).

```
iex> String.slice "the cat on the mat", 4, 3
"cat"
iex> String.slice "the cat on the mat", -3, 3
"mat"
```

`split(str, pattern || nil, options || [global: true])`

Splits str into substrings delimited by pattern. If :global is false, only one split is performed. pattern can be a string, a regular expression, or nil. In the latter case, the string is split on whitespace.

```
iex> String.split " the cat on the mat "
["the", "cat", "on", "the", "mat"]
iex> String.split "the cat on the mat", "t"
["", "he ca", " on ", "he ma", ""]
iex> String.split "the cat on the mat", ~r{[ae]}
["th", " c", "t on th", " m", "t"]
iex> String.split "the cat on the mat", ~r{[ae]}, parts: 2
["th", " cat on the mat"]
```

`starts_with?(str, prefix | [prefixes])`

Returns true if str starts with any of the given prefixes.

```
iex> String.starts_with? "string", ["elix", "stri", "ring"]
true
```

`trim(str)`

Trims leading and trailing whitespace from str.

```
iex> String.trim "\t Hello \r\n"
"Hello"
```

`trim(str, character)`

Trims leading and trailing instances of character from str.

```
iex> String.trim "!!!SALE!!!", "!"
"SALE"
```

`trim_leading(str)`

Trims leading whitespace from str.

```
iex> String.trim_leading "\t\f      Hello\t\n"
"Hello\t\n"
```

`trim_leading(str, character)`

Trims leading copies of character (an integer codepoint) from str.

```
iex> String.trim_leading "!!!SALE!!!", "!"
"SALE!!!"
```

`trim_trailing(str)`

Trims trailing whitespace from str.

```
iex> String.trim_trailing(" line \r\n")
" line"
```

`trim_trailing(str, character)`

Trims trailing occurrences of character from str.

```
iex> String.trim_trailing "!!!SALE!!!", "!"
"!!!SALE"
```

`upcase(str)`

```
iex> String.upcase "José Ørstüd"
"JOSÉ ØRSTÜD"
```

`valid?(str)`

Returns true if str is a string containing valid codepoints.

```
iex> String.valid? "ø"
true
iex> String.valid? "øøg"
true
iex> String.valid? << 0x80, 0x81 >>
false
```

Your Turn

► *Exercise: StringsAndBinaries-5*

Write a function that takes a list of double-quoted strings and prints each on a separate line, centered in a column that has the width of the longest string. Make sure it works with UTF characters.

```
iex> center(["cat", "zebra", "elephant"])
  cat
  zebra
  elephant
```

Binaries and Pattern Matching

The first rule of binaries is “if in doubt, specify the type of each field.” Available types are binary, bits, bitstring, bytes, float, integer, utf8, utf16, and utf32. You can also add qualifiers:

- `size(n)`: The size of the field, in bits.
- `signed` or `unsigned`: For integer fields, should it be interpreted as signed?
- `endianness`: big, little, or native.

Use hyphens to separate multiple attributes for a field:

```
<< length::unsigned-integer-size(12), flags::bitstring-size(4) >> = data
```

However, unless you’re doing a lot of work with binary file or protocol formats, the most common use of all this scary stuff is to process UTF-8 strings.

String Processing with Binaries

When we process lists, we use patterns that split the head from the rest of the list. With binaries that hold strings, we can do the same kind of trick. We have to specify the type of the head (UTF-8), and make sure the tail remains a binary.

```
strings(utf-iterate.ex)
defmodule Utf8 do
  def each(str, func) when is_binary(str), do: _each(str, func)

  defp _each(<< head :: utf8, tail :: binary >>, func) do
    func.(head)
    _each(tail, func)
  end

  defp _each(<<>>, _func), do: []
end

Utf8.each "dog", fn char -> IO.puts char end
```

produces

```
8706
111
103
```

The parallels with list processing are clear, but the differences are significant. Rather than use [head | tail], we use << head::utf8, tail::binary >>. And rather than terminate when we reach the empty list, [], we look for an empty binary, <<>>.

Your Turn

► *Exercise: StringsAndBinaries-6*

Write a function to capitalize the sentences in a string. Each sentence is terminated by a period and a space. Right now, the case of the characters in the string is random.

```
iex> capitalize_sentences("oh. a DOG. woof. ")
"Oh. A dog. Woof. "
```

► *Exercise: StringsAndBinaries-7*

Chapter 7 had an exercise about [calculating sales tax on page 114](#). We now have the sales information in a file of comma-separated id, ship_to, and amount values. The file looks like this:

```
id,ship_to,net_amount
123,:NC,100.00
124,:OK,35.50
125,:TX,24.00
126,:TX,44.80
127,:NC,25.00
128,:MA,10.00
129,:CA,102.00
120,:NC,50.00
```

Write a function that reads and parses this file and then passes the result to the `sales_tax` function. Remember that the data should be formatted into a keyword list, and that the fields need to be the correct types (so the `id` field is an integer, and so on).

You'll need the library functions `File.open`, `IO.read(file, :line)`, and `IO.stream(file)`.

Familiar Yet Strange

String handling in Elixir is the result of a long evolutionary process in the underlying Erlang environment. If we were starting from scratch, things would probably look a little different. But once you get over the slightly strange way that strings are matched using binaries, you'll find that it works out well. In particular, pattern matching makes it very easy to look to strings that start with a particular sequence, which in turn makes simple parsing tasks a pleasure to write.

You may have noticed that we're a long way into the book and haven't yet talked about control-flow constructs such as `if` and `case`. This is deliberate: we use them less often in Elixir than in more conventional languages. However, we still need them, so they are the subject of the next chapter.

In this chapter, you'll see:

- if and unless
- cond (a multiway if)
- case (a pattern-matching switch)
- Exceptions

CHAPTER 12

Control Flow

Elixir code tries to be declarative, not imperative.

In Elixir we write lots of small functions, and a combination of guard clauses and pattern matching of parameters replaces most of the control flow seen in other languages.

However, Elixir does have a small set of control-flow constructs. The reason I've waited so long to introduce them is that I want you to try not to use them much. You definitely will, and should, drop the occasional cond or case into your code. But before you do, consider more functional alternatives. The benefit will become obvious as you write more code—functions written without explicit control flow tend to be shorter and more focused. They're easier to read, test, and reuse. If you end up with a 10- or 20-line function in an Elixir program, it is pretty much guaranteed that it will contain one of the constructs in this chapter and that you can simplify it.

So, forewarned, let's go.

if and unless

In Elixir, if and its evil twin, unless, take two parameters: a condition and a keyword list, which can contain the keys do: and else:. If the condition is truthy, the if expression evaluates the code associated with the do: key; otherwise it evaluates the else: code. The else: branch may be absent.

```
iex> if 1 == 1, do: "true part", else: "false part"  
"true part"
```

```
iex> if 1 == 2, do: "true part", else: "false part"  
"false part"
```

Just as it does with function definitions, Elixir provides some syntactic sugar. You can write the first of the previous examples as follows:

```
iex> if 1 == 1 do
...>   "true part"
...> else
...>   "false part"
...> end
true part
```

unless is similar:

```
iex> unless 1 == 1, do: "error", else: "OK"
"OK"
iex> unless 1 == 2, do: "OK", else: "error"
"OK"
iex> unless 1 == 2 do
...>   "OK"
...> else
...>   "error"
...> end
"OK"
```

The value of if and unless is the value of the expression that was evaluated.

cond

The cond macro lets you list out a series of conditions, each with associated code. It executes the code corresponding to the first truthy conditions.

In the game of FizzBuzz, children count up from 1. If the number is a multiple of three, they say “Fizz.” For multiples of five, they say “Buzz.” For multiples of both, they say “FizzBuzz.” Otherwise, they say the number.

In Elixir, we could code this as follows:

```
control/fizzbuzz.ex
Line 1 defmodule FizzBuzz do
-
-  def upto(n) when n > 0, do: _upto(1, n, [])
-
5   defp _upto(_current, 0, result), do: Enum.reverse result
-
-  defp _upto(current, left, result) do
-    next_answer =
-      cond do
10     rem(current, 3) == 0 and rem(current, 5) == 0 ->
-       "FizzBuzz"
-     rem(current, 3) == 0 ->
-       "Fizz"
-     rem(current, 5) == 0 ->
```

```

15      "Buzz"
-      true ->
-      current
-      end
-      _upto(current+1, left-1, [ next_answer | result ])
20  end
- end

```

The cond starts on line 8. We assign the value of the cond expression to next_answer. Inside the cond, we have four alternatives—the current number is a multiple of 3 and 5, just 3, just 5, or neither. Elixir examines each in turn and returns the value of the expression following the `->` for the first true one. The `_upto` function then recurses to find the next value. Note the use of `true ->` to handle the case where none of the previous conditions match. This is the equivalent of the `else` or `default` stanza of a more traditional case statement.

There's a minor problem, though. The result list we build always has the most recent value as its head. When we finish, we'll end up with a list that has the answers in reverse order. That's why in the anchor case (when `left` is zero), we reverse the result before returning it. This is a very common pattern. And don't worry about performance—list reversal is highly optimized.

Let's try the code in IEx:

```

iex> c("fizzbuzz.ex")
[FizzBuzz]
iex> FizzBuzz.upto(20)
[1, 2, "Fizz", 4, "Buzz", "Fizz", 7, 8, "Fizz", "Buzz", 11, "Fizz",
.. 13, 14, "FizzBuzz", 16, 17, "Fizz", 19, "Buzz"]

```

In this case, we could do something different and remove the call to `reverse`. If we process the numbers in reverse order (so we start at `n` and end at 1), the resulting list will be in the correct order.

```

control/fizzbuzz1.ex
defmodule FizzBuzz do

  def upto(n) when n > 0, do: _downto(n, [])
  defp _downto(0, result), do: result
  defp _downto(current, result) do
    next_answer =
      cond do
        rem(current, 3) == 0 and rem(current, 5) == 0 ->
          "FizzBuzz"
        rem(current, 3) == 0 ->
          "Fizz"
        rem(current, 5) == 0 ->
          "Buzz"
        true ->

```

```

        current
    end
    _downto(current-1, [ next_answer | result ])
end
end

```

This code is quite a bit cleaner than the previous version. However, it is also slightly less idiomatic—readers will expect to traverse the numbers in a natural order and reverse the result.

There's a third option. FizzBuzz transforms a number into a string. We like to code things as transformations, so let's use `Enum.map` to transform the range of numbers from 1 to `n` to the corresponding FizzBuzz words.

```
control/fizzbuzz2.ex
defmodule FizzBuzz do
  def upto(n) when n > 0 do
    1..n |> Enum.map(&fizzbuzz/1)
  end

  defp fizzbuzz(n) do
    cond do
      rem(n, 3) == 0 and rem(n, 5) == 0 ->
        "FizzBuzz"
      rem(n, 3) == 0 ->
        "Fizz"
      rem(n, 5) == 0 ->
        "Buzz"
      true ->
        n
    end
  end
end
```

This section is intended to show you how `cond` works, but you'll often find that it's better not to use it, and instead to take advantage of pattern matching in function calls. The choice is yours.

```
control/fizzbuzz3.ex
defmodule FizzBuzz do
  def upto(n) when n > 0, do: 1..n |> Enum.map(&fizzbuzz/1)

  defp fizzbuzz(n), do: _fizzword(n, rem(n, 3), rem(n, 5))

  defp _fizzword(_n, 0, 0), do: "FizzBuzz"
  defp _fizzword(_n, 0, _), do: "Fizz"
  defp _fizzword(_n, _, 0), do: "Buzz"
  defp _fizzword( n, _, _), do: n
end
```

case

case lets you test a value against a set of patterns, executes the code associated with the first pattern that matches, and returns the value of that code. The patterns may include guard clauses.

For example, the File.open function returns a two-element tuple. If the open is successful, it returns `{:ok, file}`, where file is an identifier for the open file. If the open fails, it returns `{:error, reason}`. We can use case to take the appropriate action when we open a file. (Here, the code opens its own source file.)

```
control/case.ex
case File.open("case.ex") do
{:ok, file} ->
  IO.puts "First line: #{IO.read(file, :line)}"
{:error, reason} ->
  IO.puts "Failed to open file: #{reason}"
end
```

produces

```
First line: case File.open("case.ex") do
```

If we change the file name to something that doesn't exist and then rerun the code, we instead get Failed to open file: enoent.

We can use the full power of nested pattern matches:

```
control/case1.exs
defmodule Users do
  dave = %{name: "Dave", state: "TX", likes: "programming"}
  case dave do
    %{state: some_state} = person ->
      IO.puts "#{person.name} lives in #{some_state}"
    _ ->
      IO.puts "No matches"
  end
end
```

We've seen how to employ guard clauses to refine the pattern used when matching functions. We can do the same with case.

```
control/case2.exs
dave = %{name: "Dave", age: 27}
case dave do
  person = %{age: age} when is_number(age) and age >= 21 ->
    IO.puts "You are cleared to enter the Foo Bar, #{person.name}"
  _ ->
    IO.puts "Sorry, no admission"
end
```

Raising Exceptions

First, the official warning: exceptions in Elixir are *not* control-flow structures. Instead, Elixir exceptions are intended for things that should never happen in normal operation. That means the database going down or a name server failing to respond could be considered exceptional. Failing to open a configuration file whose name is fixed could be seen as exceptional. However, failing to open a file whose name a user entered is not. (You could anticipate that a user might mistype it every now and then.)

Raise an exception with the `raise` function. At its simplest, you pass it a string and it generates an exception of type `RuntimeError`.

```
iex> raise "Giving up"
** (RuntimeError) Giving up
```

You can also pass the type of the exception, along with other optional attributes. All exceptions implement at least the `message` attribute.

```
iex> raise RuntimeError
** (RuntimeError) runtime error
iex> raise RuntimeError, message: "override message"
** (RuntimeError) override message
```

You use exceptions far less in Elixir than in other languages—the design philosophy is that errors should propagate back up to an external, supervising process. We'll cover this when we talk about [OTP supervisors on page 247](#).

Elixir has all the usual exception-catching mechanisms. To emphasize how little you should use them, I've described them in [an appendix on page 355](#).

Designing with Exceptions

If `File.open` succeeds, it returns `{:ok, file}`, where `file` is the service that gives you access to the file. If it fails, it returns `{:error, reason}`. So, for code that knows a file open might not succeed and wants to handle the fact, you might write

```
case File.open(user_file_name) do
{:ok, file} ->
  process(file)
{:error, message} ->
  IO.puts :stderr, "Couldn't open #{user_file_name}: #{message}"
end
```

If instead you expect the file to open successfully every time, you could raise an exception on failure.

```
case File.open("config_file") do
{:ok, file} ->
  process(file)
{:error, message} ->
  raise "Failed to open config file: #{message}"
end
```

Or you could let Elixir raise an exception for you and write

```
{ :ok, file } = File.open("config_file")
process(file)
```

If the pattern match on the first line fails, Elixir will raise a `MatchError` exception. It won't be as informative as our version that handled the error explicitly, but if the error should never happen, this form is probably good enough (at least until it triggers the first time and the operations folks say they'd like more information).

An even better way to handle this is to use `File.open!`. The trailing exclamation point in the method name is an Elixir convention—if you see it, you know the function will raise an exception on error, and that exception will be meaningful. So we could simply write

```
file = File.open!("config_file")
```

and get on with our lives.

Doing More with Less

Elixir has just a few forms of control flow: `if`, `unless`, `cond`, `case`, and (perhaps) `raise`. But surprisingly, this doesn't matter in practice. Elixir programs are rich and expressive without a lot of branching code. And they're easier to work with as a result.

That concludes our basic tour of Elixir. Now let's start putting it all together and implement a full project.

Your Turn

► *Exercise: ControlFlow-1*

Rewrite the FizzBuzz example using `case`.

► *Exercise: ControlFlow-2*

We now have three different implementations of FizzBuzz. One uses `cond`, one uses `case`, and one uses separate functions with guard clauses.

Take a minute to look at all three. Which do you feel best expresses the problem. Which will be easiest to maintain?

The case style and the implementation using guard clauses are different from control structures in most other languages. If you feel that one of these was the best implementation, can you think of ways to remind yourself to investigate these options as you write Elixir code in the future?

► *Exercise: ControlFlow-3*

Many built-in functions have two forms. The `xxx` form returns the tuple `{:ok, data}` and the `xxx!` form returns data on success but raises an exception otherwise. However, some functions don't have the `xxx!` form.

Write an `ok!` function that takes an arbitrary parameter. If the parameter is the tuple `{:ok, data}`, return the data. Otherwise, raise an exception containing information from the parameter.

You could use your function like this:

```
file = ok! File.open("somefile")
```

In this chapter, you'll see:

- Project structure
- The mix build tool
- ExUnit testing framework
- Documentation

CHAPTER 13

Organizing a Project

Let's stop hacking and get serious.

You'll want to organize your source code, write tests, and handle any dependencies. And you'll want to follow Elixir conventions, because that way you'll get support from the tools.

In this chapter we'll look at `mix`, the Elixir build tool. We'll investigate the directory structure it uses and see how to manage external dependencies. And we'll end up using `ExUnit` to write tests for our code (and to validate the examples in our code's documentation). To motivate this, we'll write a tool that downloads and lists the n oldest issues from a GitHub project. Along the way, we'll need to find some libraries and make some design decisions typical of an Elixir project. We'll call our project `issues`.

The Project: Fetch Issues from GitHub

GitHub provides a nice web API for fetching issues.¹ Simply issue a GET request to

`https://api.github.com/repos/user/project/issues`

and you'll get back a JSON list of issues. We'll reformat this, sort it, and filter out the oldest n , presenting the result as a table:

#	created_at	title
889	2013-03-16T22:03:13Z	MIX_PATH environment variable (of sorts)
892	2013-03-20T19:22:07Z	Enhanced mix test --cover
893	2013-03-21T06:23:00Z	mix test time reports
898	2013-03-23T19:19:08Z	Add mix compile --warnings-as-errors

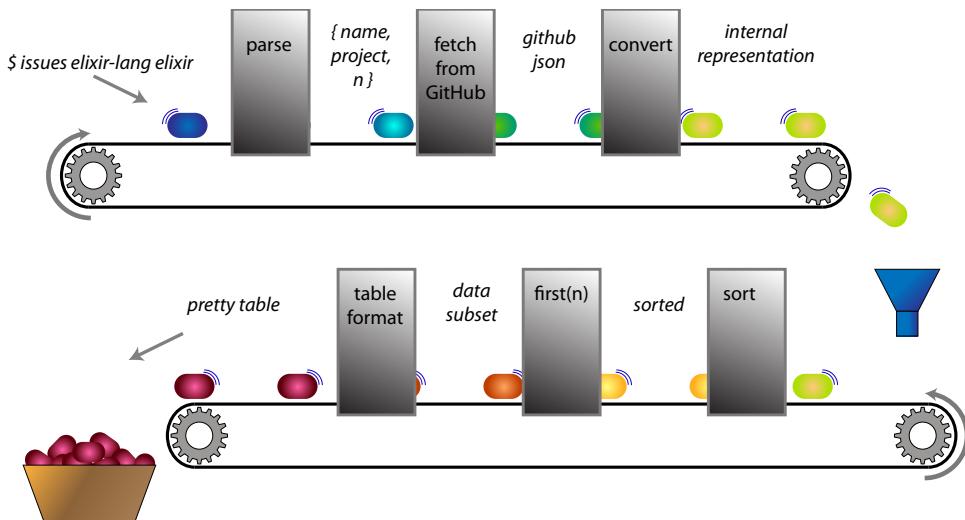
1. <http://developer.github.com/v3/>

How Our Code Will Do It

Our program will run from the command line. We'll need to pass in a GitHub user name, a project name, and an optional count. This means we'll need some basic command-line parsing.

We'll need to access GitHub as an HTTP client, so we'll have to find a library that gives us the client side of HTTP. The response that comes back will be in JSON, so we'll need a library that handles JSON, too. We'll need to be able to sort the resulting structure. And finally, we'll need to lay out selected fields in a table.

We can think of this data transformation in terms of a production line. Raw data enters at one end and is transformed by each of the stations in turn.



Here we see data, starting at the command line and ending at *pretty table*. At each stage, it undergoes a transformation (parse, fetch, and so on). These transformations are the functions we write. We'll cover each one in turn.

Step 1: Use Mix to Create Our New Project

Mix is a command-line utility that manages Elixir projects. Use it to create new projects, manage a project's dependencies, run tests, and run your code. If you have Elixir installed, you also have mix. Try running it now:

```
$ mix help
mix           # Run the default task (current: mix run)
mix archive   # List all archives
mix archive.build
:           :
mix new       # Create a new Elixir project
mix run       # Run the given file or expression
mix test      # Run a project's tests
iex -S mix    # Start IEx and run the default task
```

This is a list of the standard tasks that come with mix. (Your list may be a little different, depending on your version of Elixir.) For more information on a particular task, use `mix help taskname`.

```
$ mix help deps
```

List all dependencies and their status.

Dependencies must be specified in the `mix.exs` file in one of the following formats:

You can write your own mix tasks, for a project and to share between projects.²

Create the Project Tree

Each Elixir project lives in its own directory tree. If you use mix to manage this tree, then you'll follow the mix conventions (which are also the conventions of the Elixir community). We'll use these conventions in the rest of this chapter.

We'll call our project issues, so it will go in a directory named `issues`. We'll create this directory using mix.

At the command line, navigate to a place where you want this new project to live, and type

```
$ mix new issues
* creating README.md
:           :
* creating test
* creating test/test_helper.exs
* creating test/issues_test.exs
```

Your mix project was created successfully.

You can use mix to compile it, test it, and more:

```
cd issues
mix test
```

Run `'mix help'` for more commands.

2. <http://elixir-lang.org/getting-started/mix-otp/introduction-to-mix.html>

In tree form, the newly created files and directories look like this:

```
issues
├── .formatter.exs
├── .gitignore
├── README.md
└── config
    └── config.exs
lib
└── issues.ex
mix.exs
test
    ├── issues_test.exs
    └── test_helper.exs
```

Change into the `issues/` directory. This is a good time to set up version control. I use Git, so I do

```
$ git init
$ git add .
$ git commit -m "Initial commit of new project"
```

(I don't want to clutter the book with version-control stuff, so that's the last time I'll mention it. Make sure you follow your own version-control practices as we go along.)

Our new project contains three directories and some files.

`.formatter.exs`

Configuration used by the source code formatter

`.gitignore`

Lists the files and directories generated as by-products of the build and not to be saved in the repository.

`README.md`

A place to put a description of your project (in Markdown format). If you store your project on GitHub, this file's contents will appear on the project's home page.

`config/`

Eventually we'll put some application-specific configuration here.

`lib/`

This is where our project's source lives. Mix has already added a top-level module (`issues.ex` in our case).

`mix.exs`

This source file contains our project’s configuration options. We will be adding stuff to this as our project progresses.

`test/`

A place to store our tests. Mix has already created a helper file and a stub for unit tests of the issues module.

Now our job is to add our code. But before we do, let’s think a little about the implementation.

Transformation: Parse the Command Line

Let’s start with the command line. We really don’t want to couple the handling of command-line options into the main body of our program, so let’s write a separate module to interface between what the user types and what our program does. By convention this module is called *Project.CLI* (so our code would be in `Issues.CLI`). Also by convention, the main entry point to this module will be a function called `run` that takes an array of command-line arguments.

Where should we put this module?

Elixir has a convention. Inside the `lib/` directory, create a subdirectory with the same name as the project (so we’d create the directory `lib/issues/`). This directory will contain the main source for our application, one module per file. And each module will be namespaced inside the `Issues` module—the module naming follows the directory naming.

In this case, the module we want to write is `Issues.CLI`—it is the `CLI` module nested inside the `Issues` module. Let’s reflect that in the directory structure and put `cli.ex` in the `lib/issues` directory:

```
lib
└── issues
    └── cli.ex
    └── issues.ex
```

Elixir comes bundled with an option-parsing library,³ so we will use that. We’ll tell it that `-h` and `--help` are possible switches, and anything else is an argument. It returns a tuple, where the first element is a keyword list of the options and the second is a list of the remaining arguments. Our initial `CLI` module looks like the following:

3. <http://elixir-lang.org/docs/stable/elixir/OptionParser.html>

```
project/0/issues/lib/issues/cli.ex
defmodule Issues.CLI do
  @default_count 4
  @moduledoc """
  Handle the command line parsing and the dispatch to
  the various functions that end up generating a
  table of the last _n_ issues in a github project
  """
  def run(argv) do
    parse_args(argv)
  end
  @doc """
  `argv` can be -h or --help, which returns :help.

  Otherwise it is a github user name, project name, and (optionally)
  the number of entries to format.

  Return a tuple of `{ user, project, count }`, or `{:help}` if help was given.
  """
  def parse_args(argv) do
    parse = OptionParser.parse(argv, switches: [ help: :boolean ],
                               aliases: [ h: :help ])
    case parse do
      { [ help: true ], _, _ }           -> :help
      { _, [ user, project, count ], _ } -> { user, project, count }
      { _, [ user, project ], _ }        -> { user, project, @default_count }
      _ -> :help
    end
  end
end
```

Write Some Basic Tests

At this point, I get a little nervous if I don't have some tests. Fortunately, Elixir comes with a wonderful (and simple) testing framework called ExUnit.

Have a look at the file test/issues_test.exs.

```
project/0/issues/test/issues_test.exs
defmodule IssuesTest do
  use ExUnit.Case
  doctest Issues

  test "greets the world" do
    assert Issues.hello() == :world
  end
end
```

It acts as a template for all the test files you write. I just copy and paste the boilerplate into separate test files as I need them. So let's write tests for our CLI module, putting those tests into the file test/cli_test.exs. (Test file names must end with _test.) We'll test that the option parser successfully detects the -h and --help options, and that it returns the arguments otherwise. We'll also check that it supplies a default value for the count if only two arguments are given.

```
project/1/issues/test/cli_test.exs
defmodule CliTest do
  use ExUnit.Case
  doctest Issues

  import Issues.CLI, only: [ parse_args: 1 ]

  test ":help returned by option parsing with -h and --help options" do
    assert parse_args(["-h", "anything"]) == :help
    assert parse_args(["--help", "anything"]) == :help
  end

  test "three values returned if three given" do
    assert parse_args(["user", "project", "99"]) == { "user", "project", 99 }
  end

  test "count is defaulted if two values given" do
    assert parse_args(["user", "project"]) == { "user", "project", 4 }
  end
end
```

These tests all use the basic assert macro that ExUnit provides. This macro is clever—if an assertion fails, it can extract the values from the expression you pass it, giving you a nice error message.

To run our tests, we'll use the mix test task.

```
issues$ mix test
Compiled lib/issues.ex
Compiled lib/issues/cli.ex
Generated issues app
..
Failures:

1) test three values returned if three given (CliTest)
   test/cli_test.exs:11
     Assertion with == failed
     code: parse_args(["user", "project", "99"]) == {"user", "project", 99}
     lhs:  {"user", "project", "99"}
     rhs:  {"user", "project", 99}
     stacktrace:
       test/cli_test.exs:13
.
Finished in 0.01 seconds
4 tests, 1 failures
```

One of the four tests failed. When we pass a count as the third parameter, our code blows up. See how the assertion shows you its type (== in this case), the line of code that failed, and the two values that we compared. You can see the difference between the left-hand side (lhs), which is the value returned by `parse_args`, and the expected value (the rhs)—if your terminal and your eyes support it, you'll see that the "99" in the line labelled `lhs:` is colored red, and the 99 in the next line is green. We were expecting to get a number as the count, but we got a string.

That's easily fixed. The built-in function `String.to_integer` converts a binary (a string) into an integer.

```
project/1/issues/lib/issues/cli.ex
def parse_args(argv) do
  parse = OptionParser.parse(argv, switches: [ help: :boolean ],
                            aliases: [ h:      :help    ])
  case parse do
    { [ help: true ], _, _ } -> :help
    { _, [ user, project, count ], _ } -> { user, project,
                                                String.to_integer(count) }
    { _, [ user, project ], _ } -> { user, project, @default_count }
    _ -> :help
  end
end
```

Your Turn

► *Exercise: Organizing A Project-1*

Do what I did. Honest. Create the project and write and test the option parser. It's one thing to read about it, but you'll be doing this a lot, so you may as well start now.

Refactor: Big Function Alert

Our parse_args function is waving two red flags. First, it contains conditional logic. Second, it is too long. Let's split it up.

```
project/1a/issues/lib/issues/cli.ex
def parse_args(argv) do
  OptionParser.parse(argv, switches: [ help: :boolean],
                     aliases: [ h:     :help   ])
  |> elem(1)
  |> args_to_internal_representation()
end

def args_to_internal_representation([user, project, count]) do
  { user, project, String.to_integer(count) }
end

def args_to_internal_representation([user, project]) do
  { user, project, @default_count }
end

def args_to_internal_representation(_) do # bad arg or --help
  :help
end
```

And run the tests:

```
issues$ mix test
.....
Finished in 0.05 seconds
2 doctests, 4 tests, 0 failures
```

Transformation: Fetch from GitHub

Now let's continue down our data-transformation chain. Having parsed our arguments, we need to transform them by fetching data from GitHub. So we'll extend our run function to call a process function, passing it the value returned from the parse_args function. We could have written this:

```
process(parse_args(argv))
```

But to understand this code, you have to read it right to left. I prefer to make the chain more explicit using the Elixir pipe operator:

```
project/1a/issues/lib/issues/cli.ex
def run(argv) do
  argv
  |> parse_args
  |> process
end
```

We need two variants of the process function. One handles the case where the user asked for help and parse_args returned :help. The other handles the case where a user, a project, and a count are returned.

```
project/1a/issues/lib/issues/cli.ex
def process(:help) do
  IO.puts """
  usage: issues <user> <project> [ count | #{@default_count} ]
  """
  System.halt(0)
end

def process({user, project, _count}) do
  Issues.GithubIssues.fetch(user, project)
end
```

We can use mix to run our function. Let's first see if help gets displayed.

```
$ mix run -e 'Issues.CLI.run(["-h"])'
usage: issues <user> <project> [ count | 4 ]
```

You pass mix run an Elixir expression, which gets evaluated in the context of your application. Mix will recompile your application, as it is out of date, before executing the expression.

If we pass it user and project names, however, it'll blow up because we haven't written that code yet.

```
% mix run -e 'Issues.CLI.run(["elixir-lang", "elixir"])'
** (UndefinedFunctionError) undefined function: Issues.GithubIssues.fetch/2
      GithubIssues.fetch("elixir-lang", "elixir")
```

Let's write that code now. Our program will act as an HTTP client, accessing GitHub through its web API. So, it looks like we'll need an external library.

Step 2: Use Libraries

Elixir comes with a bunch of libraries preinstalled. Some are written in Elixir, and others in Erlang.

The first port of call is <http://elixir-lang.org/docs.html>, the Elixir documentation. Often you'll find a built-in library that does what you want.

Next, see if any standard Erlang libraries do what you need. This isn't a simple task. Visit <http://erlang.org/doc/> and look in the left sidebar for *Application Groups*. There you'll find libraries sorted by top-level category.

If you find what you're looking for in either of these places, you're all set, as all these libraries are already available to your application. But if the built-in libraries don't contain what you need, you'll have to add an external dependency.

Finding an External Library

Package managers: Ruby has RubyGems, Python has pip, Node.js has npm.

And Elixir has *hex*.

Visit <https://hex.pm> and search its list of packages that integrate nicely with a mix-based project.

If all else fails, Google and GitHub are your friends. Search for terms such as *elixir http client* or *erlang distributed logger*, and you're likely to turn up the libraries you need.

In our case, we need an HTTP client. We find that Elixir has nothing built in, but hex.pm has a number of HTTP client libraries.

To me, HTTPoison looks like a good option. So how do we include it in our project?

Adding a Library to Your Project

Mix takes the view that all external libraries should be copied into the project's directory structure. The good news is that it handles all this for us—we just need to list the dependencies, and it does the rest. Remember the `mix.exs` file at the top level of our project? Here is that original version.

```
project/0/issues/mix.exs
defmodule Issues.MixProject do
  use Mix.Project

  def project do
    [
      app: :issues,
      version: "0.1.0",
      elixir: "~> 1.6-dev",
      start_peristent: Mix.env() == :prod,
      deps: deps()
    ]
  end

  # Run "mix help compile.app" to learn about applications.
  def application do
    [
      extra_applications: [:logger]
    ]
  end

  # Run "mix help deps" to learn about dependencies.
  defp deps do
    [
      # {:dep_from_hexpm, "~> 0.3.0"},
      # {:dep_from_git, git: "https://github.com/elixir-lang/my_dep.git", tag: "0.1.0"},
    ]
  end
end
```

We add new dependencies to the deps function. As the HTTPoison package is in hex.pm, that's very simple. We just give the name and the version we want.

```
project/1a/issues/mix.exs
defp deps do
  [
    { :httpoison, "~> 1.0.0" }
  ]
end
```

In this case, we give the version as " $\sim> 1.0.0$ ". This matches any version of HTTPoison with a major version of 1 and a minor version of 0 or greater. In IEx, type `h Version` for more details.

Once your mix.exs file is updated, you're ready to have mix manage your dependencies.

Use mix deps to list the dependencies and their status:

```
$ mix deps
* httpoison (package)
  the dependency is not available, run `mix deps.get`
```

Download the dependencies with mix deps.get:

```
Resolving Hex dependencies...
Dependency resolution completed:
certifi 2.0.0
hackney 1.10.1
httpoison 0.13.0
idna 5.1.0
metrics 1.0.1
mimerl 1.0.2
ssl_verify_fun 1.1.1
unicode_util_compat 0.3.1
* Getting httpoison (Hex package)
  Checking package (https://repo.hex.pm/tarballs/httpoison-0.13.0.tar)
  Using locally cached package
  . . .
```

Run mix deps again:

```
* mimerl (Hex package) (rebar3)
  locked at 1.0.2 (mimerl) 993f9b0e
  the dependency build is outdated, please run "mix deps.compile"
* metrics (Hex package) (rebar3)
  locked at 1.0.1 (metrics) 25f094de
  the dependency build is outdated, please run "mix deps.compile"
* unicode_util_compat (Hex package) (rebar3)
  locked at 0.3.1 (unicode_util_compat) a1f612a7
  the dependency build is outdated, please run "mix deps.compile"
  . . .
* httpoison (Hex package) (mix)
  locked at 0.9.0 (httpoison) 68187a2d
  the dependency build is outdated, please run "mix deps.compile"
```

This shows that the HTTPoison library is installed but that it hasn't yet been compiled. Mix also remembers the exact version of each library it installs in the file mix.lock. This means that at any point in the future you can get the same version of the library you use now.

Don't worry that the library isn't compiled—mix will automatically compile it the first time we need it.

If you look at your project tree, you'll find a new directory called deps containing your dependencies. Note that these dependencies are themselves just projects, so you can browse their source and read their documentation.

Your Turn

► [Exercise: Organizing A Project-2](#)

Add the dependency to your project and install it.

Back to the Transformation

So, back to our problem. We have to write the function `GithubIssues.fetch`, which transforms a user name and project into a data structure containing that project's issues. The `HTTPoison` page on GitHub gives us a clue,⁴ and we write a new module, `Issues.GithubIssues`:

```
project/1a/issues/lib/issues/github_issues.ex
defmodule Issues.GithubIssues do
  @user_agent [ {"User-agent", "Elixir dave@pragprog.com"} ]
  def fetch(user, project) do
    issues_url(user, project)
    |> HTTPoison.get(@user_agent)
    |> handle_response
  end
  def issues_url(user, project) do
    "https://api.github.com/repos/#{user}/#{project}/issues"
  end
  def handle_response({ :ok, %{status_code: 200, body: body}}) do
    { :ok, body }
  end
  def handle_response({ _,   %{status_code: _,   body: body}}) do
    { :error, body }
  end
end
```

We simply call `get` on the GitHub URL. (We also have to pass in a user-agent header to keep the GitHub API happy.) What comes back is a structure. If we have a successful response, we return a tuple whose first element is `:ok`, along with the body. Otherwise we return an `:error` tuple, also with the body.

There's one more thing. The examples on the `HTTPoison` GitHub page call `HTTPoison.start`. That's because `HTTPoison` actually runs as a separate application, outside your main process. A lot of developers will copy this code, calling `start` inline like this.

In older versions of Elixir, you could add `HTTPoison` to the list of applications to start in `mix.exs`:

```
def application do
  [ applications: [ :logger, :httpoison ] ]
end
```

This is no longer necessary. The fact that you have listed `HTTPoison` as a dependency means that `mix` will automatically start it as an application.

4. <https://github.com/edgurgel/httpoison>

What Does Application Mean?

OTP is the framework that manages suites of running applications. But just what is an application?

I found the answer counterintuitive at first. Erlang programs—and, by extension, Elixir programs—are often structured as suites of cooperating subapplications. Frequently, the code that would be a library in another language is a subapplication in Elixir. It might help to think of these as components or services.

We can play with this in IEx. Use the `-S mix` option to run mix before dropping into interaction mode. Because this is the first time we've tried to run our code since installing the dependencies, you'll see them get compiled:

```
$ iex -S mix
Erlang/OTP 20 [erts-9.1] [source] [64-bit] [smp:4:4] [ds:4:4:10] [async-thr
eads:10] [hipe] [kernel-poll:false]

==> Compiling mimerl
==> Compiling metrics
:
Generated issues app

iex(1)>
```

Let's try it out. (The output is massaged to fit the page.)

```
iex> Issues.GithubIssues.fetch("elixir-lang", "elixir")
{:ok,
 [
  {"url": "https://api.github.com/repos/elixir-lang/elixir/issues/7121",
   "repository_url": "https://api.github.com/repos/elixir-lang/elixir",
   "labels_url": "https://api.github.com/repos/elixir-lang/elixir/labels{/name}",
   "events_url": "https://api.github.com/repos/elixir-lang/elixir/issues/7121/events",
   "html_url": "https://github.com/elixir-lang/elixir/issues/7121",
   "id": 282654795,
   "number": 7121,
   "title": "IEx.Helpers.h duplicate output for default arguments",
   "user": {
     "login": "wojtekmach",
     "id": 76071,
     "avatar_url": "https://avatars0.githubusercontent.com/u/76071?v=4",
     "gravatar_id": "",
     "url": "https://api.github.com/users/wojtekmach",
     "html_url": "https://github.com/wojtekmach",
     "followers_url": "https://api.github.com/users/wojtekmach/followers",
     ...
   }
 ]}
```

This tuple is the body of the GitHub response. The first element is set to :ok. The second element is a string containing the data encoded in JSON format.

Transformation: Convert Response

We'll need a JSON library to convert the response into a data structure. Searching hex.pm, I found the poison library (no relation to HTTPoison), so let's add its dependency to our mix.exs file.⁵

```
project/2/issues/mix.exs
defp deps do
  [
    { :httpoison, "~> 1.0.0" },
    { :poison,    "~> 3.1"   },
  ]
end
```

Run mix deps.get, and you'll end up with poison installed.

To convert the body from a string, we call the Poison.Parser.parse! function when we return the message from the GitHub API:

```
project/3/issues/lib/issues/github_issues.ex
def handle_response({ _, %{status_code: status_code, body: body}}) do
  {
    status_code |> check_for_error(),
    body       |> Poison.Parser.parse!()
  }
end

defp check_for_error(200), do: :ok
defp check_for_error(_),   do: :error
```

We also have to deal with a possible error response from the fetch, so back in the CLI module we write a function that decodes the body and returns it on a success response; the function extracts the error from the body and displays it otherwise.

```
project/3/issues/lib/issues/cli.ex
def process({user, project, _count}) do
  Issues.GithubIssues.fetch(user, project)
  |> decode_response()
end

def decode_response({:ok, body}), do: body

def decode_response({:error, error}) do
  IO.puts "Error fetching from Github: #{error["message"]}"
  System.halt(2)
end
```

5. <https://github.com/devinus/poison>

The JSON that GitHub returns for a successful response is a list of maps, where each map in the list contains a GitHub issue.

Dependencies That Aren't in Hex

The dependencies you need are likely to be in hex, so mix will probably find them automatically. However, sometimes you'll need to go further afield. The good news is that mix can also load dependencies from other sources. The most common is GitHub.

HTTPoison uses a library called Hackney. In earlier versions of the book, Hackney wasn't in hex.pm, so I had to add the following dependency to my mix.exs:

```
def deps do
  [ { . . . },
    { :hackney, github: "benoitz/hackney" }
  ]
end
```

Application Configuration

Before we move on, there's one little tweak I'd like to make. The issues_url function hard-codes the GitHub URL. Let's make this configurable.

Remember that when we created the project using mix new, it added a config/ directory containing config.exs. That file stores application-level configuration.

It should start with the line

```
use Mix.Config
```

We then write configuration information for each of the applications in our project. Here we're configuring the Issues application, so we write this code:

```
project/3a/issues/config/config.exs
use Mix.Config
config :issues, github_url: "https://api.github.com"
```

Each config line adds one or more key/value pairs to the given application's environment. If you have multiple lines for the same application, they accumulate, with duplicate keys in later lines overriding values from earlier ones.

In our code, we use Application.get_env to return a value from the environment.

```
project/3a/issues/lib/issues/github_issues.ex
# use a module attribute to fetch the value at compile time
@github_url Application.get_env(:issues, :github_url)

def issues_url(user, project) do
  "#{@github_url}/repos/#{user}/#{project}/issues"
end
```

Because the application environment is commonly used in Erlang code, you'll find yourself using the configuration facility to configure code you import, as well as code you write.

Sometimes you may want to vary the configuration, perhaps depending on your application's environment. One way is to use the `import_config` function, which reads configuration from a file. If your `config.exs` contains

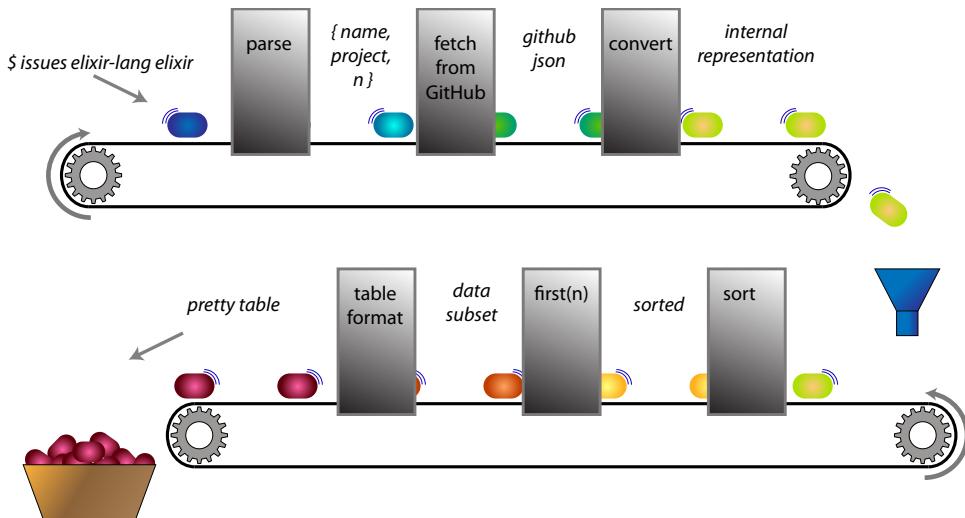
```
use Mix.Config
import_config "#{Mix.env}.exs"
```

then Elixir will read `dev.exs`, `test.exs`, or `prod.exs`, depending on your environment.

You can override the default config file name (`config/config.exs`) using the `--config` option to `elixir`.

Transformation: Sort Data

Have a look at the original design in the following figure.



We're making good progress—we've coded all the functions of the top conveyor belt. Our next transformation is to sort the data on its `created_at` field, with the newest entries first. And this can just use a standard Elixir library function, `sort/2`. We could create a new module for this, but it would be pretty lonely. For now we'll put the function in the CLI module and keep an eye out for opportunities to move it out if we add related functions later.

So now our CLI module contains this:

```
project/3b/issues/lib/issues/cli.ex
def process({user, project, _count}) do
  Issues.GithubIssues.fetch(user, project)
  |> decode_response()
  |> sort_into_descending_order()
end

def sort_into_descending_order(list_of_issues) do
  list_of_issues
  |> Enum.sort(fn i1, i2 ->
    i1["created_at"] >= i2["created_at"]
  end)
end
```

That `sort_into_descending_order` function worries me a little—I get the comparison the wrong way around about 50% of the time, so let's write a little CLI test.

```
project/3b/issues/test/cli_test.exs
test "sort descending orders the correct way" do
  result = sort_into_descending_order(fake_created_at_list(["c", "a", "b"]))
  issues = for issue <- result, do: Map.get(issue, "created_at")
  assert issues == ~w{ c b a }
end

defp fake_created_at_list(values) do
  for value <- values,
    do: %{"created_at" => value, "other_data" => "xxx"}
end
```

Update the import line at the top of the test:

```
import Issues.CLI, only: [ parse_args: 1,
                           sort_into_descending_order: 1 ]
```

and run it:

```
$ mix test
....
Finished in 0.00 seconds
5 tests, 0 failures
```

Lookin' fine; mighty fine.

Transformation: Take First n Items

Our next transformation is to extract the first $count$ entries from the list. Resisting the temptation to write the function ourselves (How *would* you write such a function?), we discover the built-in `Enum.take`:

```

def process({user, project, count}) do
  Issues.GithubIssues.fetch(user, project)
  |> decode_response()
  |> sort_into_descending_order()
  |> last(count)
end

def last(list, count) do
  list
  |> Enum.take(count)
  |> Enum.reverse
end

```

Your Turn

► *Exercise: OrganizingAProject-3*

Bring your version of this project in line with the code here.

► *Exercise: OrganizingAProject-4*

(Tricky) Before reading the next section, see if you can write the code to format the data into columns, like the sample output at the start of the chapter. This is probably the longest piece of Elixir code you'll have written. Try to do it without using if or cond.

Transformation: Format the Table

All that's left from our design is to create the formatted table. The following would be a nice interface:

```

def process({user, project, count}) do
  Issues.GithubIssues.fetch(user, project)
  |> decode_response()
  |> sort_intoAscendingOrder()
  |> last(count)
  |> printTableForColumns(["number", "created_at", "title"])
end

```

We pass the formatter the list of columns to include in the table, and it writes the table to standard output. The formatter doesn't add any new project- or design-related techniques, so we'll just show the listing.

```
project/4/issues/lib/issues/table_formatter.ex
defmodule Issues.TableFormatter do

  import Enum, only: [ each: 2, map: 2, map_join: 3, max: 1 ]

  def printTableForColumns(rows, headers) do
    with data_by_columns = splitIntoColumns(rows, headers),
          column_widths  = widthsOf(data_by_columns),
          format         = formatFor(column_widths)
```

```

do
  puts_one_line_in_columns(headers, format)
  IO.puts(separator(column_widths))
  puts_in_columns(data_by_columns, format)
end
end

def split_into_columns(rows, headers) do
  for header <- headers do
    for row <- rows, do: printable(row[header])
  end
end

def printable(str) when is_binary(str), do: str
def printable(str), do: to_string(str)

def widths_of(columns) do
  for column <- columns, do: column |> map(&String.length/1) |> max
end

def format_for(column_widths) do
  map_join(column_widths, " | ", fn width -> "~#{width}s" end) <> "~n"
end

def separator(column_widths) do
  map_join(column_widths, "-+-", fn width -> List.duplicate("-", width) end)
end

def puts_in_columns(data_by_columns, format) do
  data_by_columns
  |> List.zip
  |> map(&Tuple.to_list/1)
  |> each(&puts_one_line_in_columns/1, format)
end

def puts_one_line_in_columns(fields, format) do
  :io.format(format, fields)
end
end

```

And here are the tests for it:

```

project/4/issues/test/table_formatter_test.exs
defmodule TableFormatterTest do
  use ExUnit.Case          # bring in the test functionality
  import ExUnit.CaptureIO # And allow us to capture stuff sent to stdout

  alias Issues.TableFormatter, as: TF

  @simple_test_data [
    [ c1: "r1 c1", c2: "r1 c2", c3: "r1 c3", c4: "r1++c4" ],
    [ c1: "r2 c1", c2: "r2 c2", c3: "r2 c3", c4: "r2 c4" ],
    [ c1: "r3 c1", c2: "r3 c2", c3: "r3 c3", c4: "r3 c4" ],
    [ c1: "r4 c1", c2: "r4++c2", c3: "r4 c3", c4: "r4 c4" ]
  ]

```

```

@headers [ :c1, :c2, :c4 ]

def split_with_three_columns do
  TF.split_into_columns(@simple_test_data, @headers)
end

test "split_into_columns" do
  columns = split_with_three_columns()
  assert length(columns) == length(@headers)
  assert List.first(columns) == ["r1 c1", "r2 c1", "r3 c1", "r4 c1"]
  assert List.last(columns) == ["r1++c4", "r2 c4", "r3 c4", "r4 c4"]
end

test "column_widths" do
  widths = TF.widths_of(split_with_three_columns())
  assert widths == [ 5, 6, 7 ]
end

test "correct format string returned" do
  assert TF.format_for([9, 10, 11]) == "~-9s | ~-10s | ~-11s~n"
end

test "Output is correct" do
  result = capture_io fn ->
    TF.print_table_for_columns(@simple_test_data, @headers)
  end
  assert result == """
    c1      | c2      | c4
    -----+-----+-----
    r1 c1 | r1 c2 | r1++c4
    r2 c1 | r2 c2 | r2 c4
    r3 c1 | r3 c2 | r3 c4
    r4 c1 | r4++c2 | r4 c4
    """
end
end

```

(Although you can't see it here, the output we compare against in the last test contains trailing whitespace.)

Rather than clutter the process function in the CLI module with a long module name, I chose to use import to make the print function available without a module qualifier. This goes near the top of cli.ex.

```

defmodule Issues.CLI do
  import Issues.TableFormatter, only: [ print_table_for_columns: 2 ]

```

This code also uses a great Elixir testing feature. By importing ExUnit.CaptureIO, we get access to the capture_io function. This runs the code passed to it but captures anything written to standard output, returning it as a string.

Step 3: Make a Command-Line Executable

Although we can run our code by calling the `run` function via `mix`, it isn't friendly for other users. So let's create something we can run from the command line.

`Mix` can package our code, along with its dependencies, into a single file that can be run on any Unix-based platform. This uses Erlang's `escript` utility, which can run precompiled programs stored as a Zip archive. In our case, the program will be run as `issues`.

When `escript` runs a program, it looks in your `mix.exs` file for the option `escript`. This should return a keyword list of `escript` configuration settings. The most important of these is `main_module`, which must be set to the name of a module containing a `main` function. It passes the command-line arguments to this `main` function as a list of character lists (not binaries). As this seems to be a command-line concern, we'll put the `main` function in `Issues.CLI`. Here's the update to `mix.exs`:

```
project/4/issues/mix.exs
defmodule Issues.MixProject do
  use Mix.Project

  def project do
    [
      app: :issues,
      escript: escript_config(),
      version: "0.1.0",
      elixir: "~> 1.6-dev",
      start_peristent: Mix.env() == :prod,
      deps: deps()
    ]
  end

  def application do
    [
      extra_applications: [:logger]
    ]
  end

  defp deps do
    [
      { :httpoison, "~> 1.0.0" },
      { :poison,    "~> 3.1"   },
    ]
  end
  defp escript_config do
    [
      main_module: Issues.CLI
    ]
  end
end
```

Now let's add a main function to our CLI. In fact, all we need to do is rename the existing run function:

```
project/4/issues/lib/issues/cli.ex
def main(argv) do
  argv
  |> parse_args
  |> process
end
```

Then we package our program using mix:

```
$ mix escript.build
Generated escript issues
```

Now we can run the app locally. We can also send it to a friend—it will run on any computer that has Erlang installed.

```
$ ./issues pragdave earmark 4
num | created_at           | title
----+-----+-----
159 | 2017-09-21T10:01:24Z | Block level HTML ... messes up formatting
161 | 2017-10-11T09:12:59Z | Be clear in README ... GFM are supported.
162 | 2017-10-11T16:59:50Z | Working on #161, looking at rendering
171 | 2017-12-03T11:08:40Z | Fix typespecs
```

Step 4: Add Some Logging

Imagine a large Elixir application—dozens of processes potentially running across a number of nodes. You'd really want a standard way to keep track of significant events as it runs. Enter the Elixir logger.

The default mix.exs starts the logger for your application.

```
project/5/issues/mix.exs
def application do
  [
    extra_applications: [:logger]
  ]
end
```

The logger supports four levels of message—in increasing order of severity they are debug, info, warn, and error. You select the level of logging in two ways.

First, you can determine at compile time the minimum level of logging to include. Logging below this level is not even compiled into your code. The compile-time level is set in the config/config.exs file:

```
project/5/issues/config/config.exs
use Mix.Config

config :issues,
  github_url: "https://api.github.com"

➤ config :logger,
➤   compile_time_purge_level: :info
```

Next, you can choose to change the minimum log level at runtime by calling `Logger.configure`. (Clearly, this cannot enable log levels that you excluded at compile time.)

After all this configuration, it's time to add some logging.

The basic logging functions are `Logger.debug`, `.info`, `.warn`, and `.error`. Each function takes either a string or a zero-arity function:

```
Logger.debug "Order total #{total(order)}"
Logger.debug fn -> "Order total #{total(order)}" end
```

Why have the function version? Perhaps the calculation of the order total is expensive. In the first version, we'll always call it to interpolate the value into our string, even if the runtime log level is set to ignore debug-level messages. In the function variant, though, the `total` function will be invoked only if the log message is needed.

Anyway, here's a version of our `fetch` function with some logging:

```
project/5/issues/lib/issues/github_issues.ex
defmodule Issues.GithubIssues do

➤  require Logger

  @user_agent [ {"User-agent", "Elixir dave@pragprog.com"} ]
  # use a module attribute to fetch the value at compile time
  @github_url Application.get_env(:issues, :github_url)

  def fetch(user, project) do
➤    Logger.info("Fetching #{user}'s project #{project}")
    issues_url(user, project)
    |> HTTPoison.get(@user_agent)
    |> handle_response
  end

  def issues_url(user, project) do
    "#{@github_url}/repos/#{user}/#{project}/issues"
  end
```

```

def handle_response({ _, %{status_code: status_code, body: body}}) do
  ▶ Logger.info("Got response: status code=#{status_code}")
  ▶ Logger.debug(fn -> inspect(body) end)
  {
    status_code |> check_for_error(),
    body        |> Poison.Parser.parse!()
  }
end

defp check_for_error(200), do: :ok
defp check_for_error(_),   do: :error
end

```

Note the use of require Logger at the top of the module. If you forget this (and I do every time), you'll get an error when you make the first call to Logger.

We can play with the new code in IEx:

```

iex> Issues.CLI.process {"pragdave", "earmark", 1}
19:53:44.207 [info] Fetching pragdave's project earmark
19:53:44.804 [info] Got response: status code=200
num | created_at           | title
---+-----+-----+
171 | 2017-12-03T11:08:40Z | Fix typespecs
:ok

```

Notice that the debug-level message is not displayed.

Step 5: Create Project Documentation

Java has Javadoc, Ruby has RDoc, and Elixir has ExDoc—a documentation tool that describes your project, showing the modules, the things defined in them, and any documentation you've written for them.

Using it is easy. First, add the ExDoc dependency to your mix.exs file. You'll also need to add an output formatter—I use earmark, a pure-Elixir Markdown-to-HTML convertor.

```

defp deps do
  [
    { :httpoison, "~> 1.0.0" },
    { :poison,    "~> 3.1.0" },
    { :ex_doc,    "~> 0.18.1" },
    { :earmark,   "~> 1.2.4" },
  ]
end

```

While you're in the mix.exs, you can add a project name and (if your project is in GitHub) a URL. The latter allows ExDoc to provide live links to your source code. These parameters go in the project function:

```

def project do
  [ app:      :issues,
    version:  "0.0.1",
    name:     "Issues",
    source_url: "https://github.com/pragdave/issues",
    deps:      deps ]
end

```

Then run `mix deps.get`.

To generate the documentation, just run

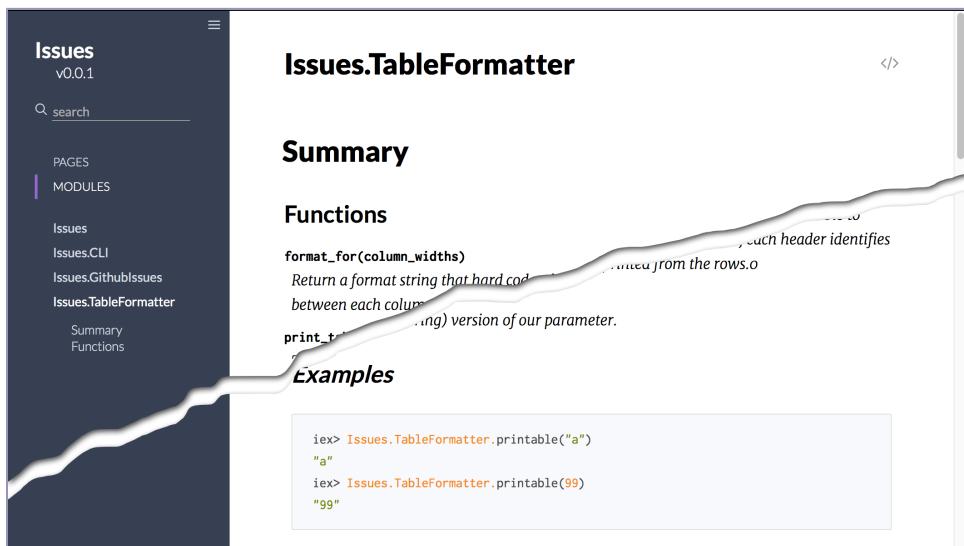
```
$ mix docs
```

Docs generated with success.

Open up `docs/index.html` in your browser to read them.

The first time you run this task, it will install ExDoc. That involves compiling some C code, so you'll need a development environment on your machine.

Open `docs/index.html` in your browser, then use the sidebar on the left to search or drill down through your modules. Here's what I see for the start of the documentation for `TableFormatter`:

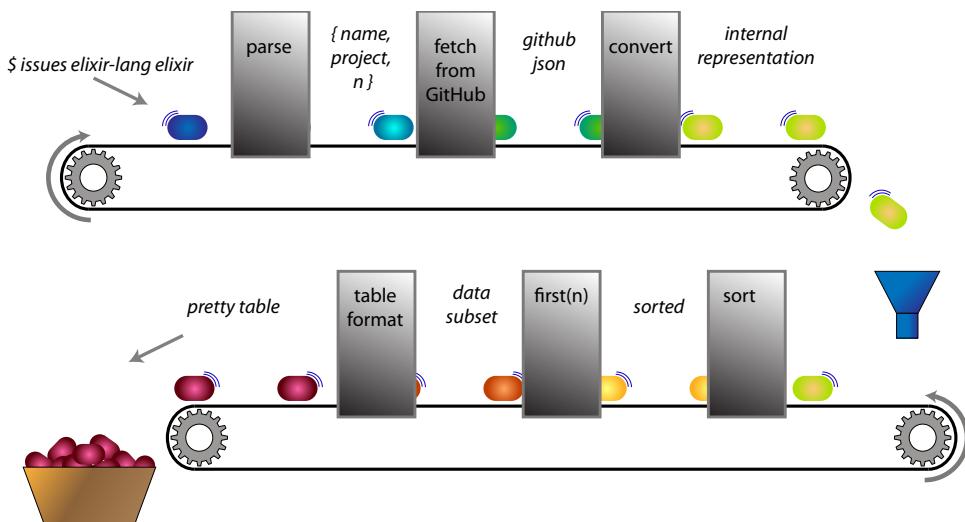


And that's it. The full project is in the source download at [project/5/issues](#).

Coding by Transforming Data

I wanted to show you how Elixir projects are written—the tools we use and the processes we follow. I wanted to illustrate how lots of small functions can transform data, how specifying that transformation acts as an outline for the program, and how easy testing can be in Elixir.

But mostly I wanted to show how enjoyable Elixir development is, and how thinking about the world in terms of data and its transformation is a productive way to code. Look at our original design:



Then have a look at the `CLI.process` function:

```
def process({user, project, count}) do
  Issues.GithubIssues.fetch(user, project)
  |> decode_response()
  |> sort_intoAscendingOrder()
  |> last(count)
  |> printTableForColumns(["number", "created_at", "title"])
end
```

This is a cool way to code. Next we'll dig into some of the tooling that makes using Elixir a joy.

Your Turn

► *Exercise: Organizing A Project-6*

In the United States, the National Oceanic and Atmospheric Administration provides hourly XML feeds of conditions at 1,800 locations.⁶ For example, the feed for a small airport close to where I'm writing this is at http://w1.weather.gov/xml/current_obs/KDFO.xml.

Write an application that fetches this data, parses it, and displays it in a nice format.

(Hint: You might not have to download a library to handle XML parsing.)

6. http://w1.weather.gov/xml/current_obs

In this chapter, you'll see:

- Debugging
- Testing
- Code exploration
- Server monitoring
- Source-code formatting

CHAPTER 14

Tooling

You'd expect that a relatively new language would come with a fairly minimal set of tools—after all, the development team will be having fun playing with the language.

Not so with Elixir. Tooling was important from the start, and the core team has spent a lot of time providing a world-class environment in which to develop code.

In this short chapter, we'll look at some aspects of this.

This chapter is not the full list. We've already seen the ExDoc tool, which creates beautiful documentation for your code. Later, when we look at [OTP applications, on page 282](#), we'll experiment with the Elixir release manager, a tool for managing releases while your application continues to run.

For now, let's look at testing, code-exploration, and server-monitoring tools.

Debugging with IEx

You already know that IEx is the go-to utility to play with Elixir code. It also has a secret and dark second life as a debugger. It isn't fancy, but it lets you get into a running program and examine the environment.

You enter the debugger when running Elixir code hits a *breakpoint*. There are two ways of creating a breakpoint. One works by adding calls into the code you want to debug. The other is initiated from inside IEx. We'll look at both using the following (buggy) code:

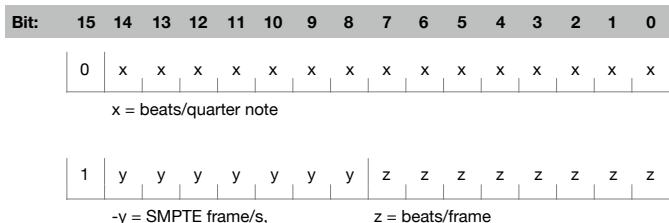
```
tooling/buggy/lib/buggy.ex
defmodule Buggy do
  def parse_header(
    <<
      format::integer-16,
      tracks::integer-16,
```

```
>     division::integer-16
    >>
) do
  IO.puts "format: #{format}"
  IO.puts "tracks: #{tracks}"
  IO.puts "division: #{decode(division)}"
end

def decode(<< 1::1, beats::15 >>) do
  "J = #{beats}"
end

def decode(<< 0::1, fps::7, beats::8 >>) do
  "#{-fps} fps, #{beats}/frame"
end
end
```

This code is supposed to decode the data part of a MIDI header frame. This contains three 16-bit fields: the format, the number of tracks, and the time division. This last field comes in one of two formats:



The `parse_header/1` function splits the overall header into the three fields, and the `decode/1` function works out which type of time division we have.

Let's run it, using a sample header I extracted from a MIDI file.

```
$ iex -S mix
iex> header = << 0, 1, 0, 8, 0, 120 >>
<<0, 1, 0, 8, 0, 120>>
iex> Buggy.parse_header header
format: 1
tracks: 8
** (FunctionClauseError) no function clause matching in Buggy.decode/1
iex>
```

Oh no! That was totally unexpected. It looks like we're not passing the correct value to decode. Let's use the debugger to find out what's going on.

Injecting Breakpoints Using IEx.pry

We can add a breakpoint to our source code using the `pry` function. For example, to stop our code just before we call `decode` we could write this:

```
def parse_header(
  <<
    format::integer-16,
    tracks::integer-16,
    division::integer-16
  >>
) do
  > require IEx; IEx.pry
  IO.puts "format: #{format}"
  IO.puts "tracks: #{tracks}"
  IO.puts "division: #{decode(division)}"
end
```

(We need the require because pry is a macro.)

Let's try the code now:

```
$ iex -S mix
iex> Buggy.parse_header << 0, 1, 0, 8, 0, 120 >>
Break reached: Buggy.parse_header/1 (lib/buggy.ex:11)

  9:
  10:   require IEx; IEx.pry
  11:   IO.puts "format: #{format}"

pry> binding
[division: 120, format: 1, tracks: 8]
iex> continue()
format: 1
tracks: 8
** (FunctionClauseError) no function clause matching in Buggy.decode/1
```

We reached the breakpoint, and IEx entered *pry mode*. It showed us the function we were in as well as the source lines surrounding the breakpoint.

At this point, IEx is running in the context of this function, so a call to binding shows the local variables. The value in the division function is 120, but that isn't matching either of the parameters to decode.

Aha! decode is expecting a binary, not an integer. Let's fix our code:

```
def parse_header(
  <<
    format::integer-16,
    tracks::integer-16,
  > division::bits-16
  >>
) do
  ...
```

The pry call is still in there, so let's recompile and try again:

```
iex> r Buggy
{:reloaded, Buggy, [Buggy]}
iex> Buggy.parse_header << 0, 1, 0, 8, 0, 120 >>
Break reached: Buggy.parse_header/1 (lib/buggy.ex:12)

10:    ) do
11:
12:      require IEx; IEx.pry
13:      IO.puts "format: #{format}"
14:      IO.puts "tracks: #{tracks}"

pry> binding
[division: <<0, 120>>, format: 1, tracks: 8]
pry> continue
format: 1
tracks: 8
division: 0 fps, 120/frame
:ok
```

Now the division is a binary, and when we continue the code runs and outputs the header fields. Except...it's parsing the time division as if it were the SMPTE version, and not the beats/quarter note version.

Setting Breakpoints with Break

The second way to create a breakpoint doesn't involve any code changes. Instead, you can use the `break!` command inside IEx to add a breakpoint on any public function. Let's remove the call to `pry` and run the code again. Inside IEx we'll add a breakpoint on the `decode` function:

```
iex> require IEx
IEx
iex> break! Buggy.decode/1
1
iex> breaks

ID  Module.function/arity  Pending stops
-----
1   Buggy.decode/1          1

iex> Buggy.parse_header << 0, 1, 0, 8, 0, 120 >>
format: 1
tracks: 8
Break reached: Buggy.decode/1 (lib/buggy.ex:21)

19:  end
20:
21:  def decode(<< 0::1, fps::7, beats::8 >>) do
22:    "#{-fps} fps, #{beats}/frame"
23:  end

pry> binding
[division: <<0, 120>>, format: 1, tracks: 8]
```

We hit the breakpoint, and we are indeed matching the wrong version of the decode function when we pass it 000000001111000. Ah, that's because I'm discriminating based on the value of the top bit, and I got it the wrong way around: the SMPTE version should be

```
def decode(<< 1::1, fps::7, beats::8 >>) do
```

and the beats version should be

```
def decode(<< 0::1, beats::15 >>) do
```

There's lots more functionality in the debugger. You can start by getting help for `IEx.break/4`.

Does This Seem a Little Artificial?

I have a confession to make. The only time I use the Elixir breakpoint facility is when I work on this section of the book. If I have to add code to the source to break in the middle of a function, then I can just raise an exception there instead to get the information I need. And the fact that I can only break at public functions from inside `IEx` means that I can't get the kind of granularity I need to diagnose issues, because 90% of my functions are private.

However, I'm an old curmudgeon—my favorite editor is a card punch. Don't let my lack of enthusiasm stop you from trying the debugger.

Testing

We already used the ExUnit framework to write tests for our Issues tracker app. But that chapter only scratched the surface of Elixir testing. Let's dig deeper.

Testing the Comments

When I document my functions, I like to include examples of the function being used—comments saying things such as, “Feed it these arguments, and you'll get this result.” In the Elixir world, a common way to do this is to show the function being used in an `IEx` session.

Let's look at an example from our Issues app. The `TableFormatter` formatter module defines a number of self-contained functions that we can document.

```
project/5/issues/lib/issues/table_formatter.ex
defmodule Issues.TableFormatter do
  import Enum, only: [ each: 2, map: 2, map_join: 3, max: 1 ]
  @doc """

```

Takes a list of row data, where each row is a Map, **and** a list of headers. Prints a table to STDOUT of the data from each row identified by each header. That is, each header identifies a column, **and** those columns are extracted **and** printed from the rows.

We calculate the width of each column to fit the longest element **in** that column.

```
"""
def print_table_for_columns(rows, headers) do
  with data_by_columns = split_into_columns(rows, headers),
       column_widths   = widths_of(data_by_columns),
       format         = format_for(column_widths)
  do
    puts_one_line_in_columns(headers, format)
    IO.puts(separator(column_widths))
    puts_in_columns(data_by_columns, format)
  end
end

@doc """
Given a list of rows, where each row contains a keyed list of columns, return a list containing lists of the data in each column. The `headers` parameter contains the list of columns to extract

## Example
iex> list = [Enum.into([{"a", "1"}, {"b", "2"}, {"c", "3"}], %{}),
...>           Enum.into([{"a", "4"}, {"b", "5"}, {"c", "6"}], %{})]
iex> Issues.TableFormatter.split_into_columns(list, [ "a", "b", "c" ])
[ ["1", "4"], ["2", "5"], ["3", "6"] ]

"""
def split_into_columns(rows, headers) do
  for header <- headers do
    for row <- rows, do: printable(row[header])
  end
end

@doc """
Return a binary (string) version of our parameter.

## Examples
iex> Issues.TableFormatter.printable("a")
"a"
iex> Issues.TableFormatter.printable(99)
"99"

"""
def printable(str) when is_binary(str), do: str
def printable(str), do: to_string(str)

@doc """
Given a list containing sublists, where each sublist contains the data for a column, return a list containing the maximum width of each column.

```

```

## Example
iex> data = [ [ "cat", "wombat", "elk"], ["mongoose", "ant", "gnu"] ]
iex> Issues.TableFormatter.widths_of(data)
[ 6, 8 ]
"""
def widths_of(columns) do
  for column <- columns, do: column |> map(&String.length/1) |> max
end

@doc """
Return a format string that hard-codes the widths of a set of columns.
We put `|` between each column.

## Example
iex> widths = [5,6,99]
iex> Issues.TableFormatter.format_for(widths)
"~-5s | ~-6s | ~-99s~n"
"""
def format_for(column_widths) do
  map_join(column_widths, " | ", fn width -> "~-#{width}s" end) <> "~n"
end

@doc """
Generate the line that goes below the column headings. It is a string of
hyphens, with + signs where the vertical bar between the columns goes.

## Example
iex> widths = [5,6,9]
iex> Issues.TableFormatter.separator(widths)
"-----+-----+-----"
"""
def separator(column_widths) do
  map_join(column_widths, "-+-", fn width -> List.duplicate("-", width) end)
end

@doc """
Given a list containing rows of data, a list containing the header selectors,
and a format string, write the extracted data under control of the format string.
"""
def puts_in_columns(data_by_columns, format) do
  data_by_columns
  |> List.zip
  |> map(&Tuple.to_list/1)
  |> each(&puts_one_line_in_columns/1, format))
end

def puts_one_line_in_columns(fields, format) do
  :io.format(format, fields)
end
end

```

Note how some of the documentation contains sample IEx sessions. I like doing this. It helps people who come along later understand how to use my

code. But, as importantly, it lets *me* understand what my code will feel like to use. I typically write these sample sessions before I start on the code, changing stuff around until the API feels right.

But the problem with comments is that they just don't get maintained. The code changes, the comment gets stale, and it becomes useless. Fortunately, ExUnit has doctest, a tool that extracts the iex sessions from your code's @doc strings, runs it, and checks that the output agrees with the comment.

To invoke it, simply add one or more

```
doctest <<ModuleName>>
```

lines to your test files. You can add them to existing test files for a module (such as table_formatter_test.exs) or create a new test file just for them. That's what we'll do here. Let's create a new test file, test/doc_test.exs, containing this:

```
project/5/issues/test/doc_test.exs
defmodule DocTest do
  use ExUnit.Case
  doctest Issues.TableFormatter
end
```

We can now run it:

```
$ mix test test/doc_test.exs
.....
Finished in 0.00 seconds
5 doctests, 0 failures
```

And, of course, these tests are integrated into the overall test suite:

```
$ mix test
.....
Finished in 0.1 seconds
5 doctests, 9 tests, 0 failures
```

Let's force an error to see what happens:

```
@doc """
Return a binary (string) version of our parameter.

## Examples

  iex> Issues.TableFormatter.printable("a")
  "a"
  iex> Issues.TableFormatter.printable(99)
  "99.0"
"""

def printable(str) when is_binary(str), do: str
def printable(str), do: to_string(str)
```

And run the tests again:

```
$ mix test test/doc_test.exs
.....
1) test doc at Issues.TableFormatter.printable/1 (3) (DocTest)
  Doctest failed
    code: " Issues.TableFormatter.printable(99) should equal \"99.0\""
    lhs: "\"99\""
    stacktrace:
      lib/issues/table_formatter.ex:52: Issues.TableFormatter (module)
6 tests, 1 failures
```

Structuring Tests

You'll often find yourself wanting to group your tests at a finer level than per module. For example, you might have multiple tests for a particular function, or multiple functions that work on the same test data. ExUnit has you covered.

Let's test this simple module:

```
tooling/pbt/lib/stats.ex
defmodule Stats do
  def sum(vals),    do: vals |> Enum.reduce(0, &+/2)
  def count(vals), do: vals |> length
  def average(vals), do: sum(vals) / count(vals)
end
```

Our tests might look something like this:

```
tooling/pbt/test/describe.exs
defmodule TestStats do
  use ExUnit.Case

  test "calculates sum" do
    list = [1, 3, 5, 7, 9]
    assert Stats.sum(list) == 25
  end

  test "calculates count" do
    list = [1, 3, 5, 7, 9]
    assert Stats.count(list) == 5
  end

  test "calculates average" do
    list = [1, 3, 5, 7, 9]
    assert Stats.average(list) == 5
  end
end
```

There are a couple of issues here. First, these tests only pass in a list of integers. Presumably we'd want to test with floats, too. So let's use the describe feature of ExUnit to document that these are the integer versions of the tests:

```
tooling/pbt/test/describe.exs
defmodule TestStats0 do
  use ExUnit.Case

  describe "Stats on lists of ints" do
    test "calculates sum" do
      list = [1, 3, 5, 7, 9]
      assert Stats.sum(list) == 25
    end

    test "calculates count" do
      list = [1, 3, 5, 7, 9]
      assert Stats.count(list) == 5
    end

    test "calculates average" do
      list = [1, 3, 5, 7, 9]
      assert Stats.average(list) == 5
    end
  end
end
```

If any of these fail, the message would include the description and test name:

```
test Stats on lists of ints calculates sum (TestStats0)
  test/describe.exs:12
  Assertion with == failed
  ...
```

A second issue with our tests is that we're duplicating the test data in each function. In this particular case this is arguably not a major problem. There are times, however, where this data is complicated to create. So let's use the setup feature to move this code into a single place. While we're at it, we'll also put the expected answers into the setup. This means that if we decide to change the test data in the future, we'll find it all in one place.

```
tooling/pbt/test/describe.exs
defmodule TestStats1 do
  use ExUnit.Case

  describe "Stats on lists of ints" do
    setup do
      [ list: [1, 3, 5, 7, 9, 11],
        sum: 36,
        count: 6
      ]
    end

    test "calculates sum", fixture do
      assert Stats.sum(fixture.list) == fixture.sum
    end
  end
end
```

```

test "calculates count", fixture do
  assert Stats.count(fixture.list) == fixture.count
end

test "calculates average", fixture do
  assert Stats.average(fixture.list) == fixture.sum / fixture.count
end
end
end

```

The setup function is invoked automatically before each test is run. (There's also a `setup_all` function that is invoked just once for the test run.) The setup function returns a keyword list of named test data. In testing circles, this data, which is used to drive tests, is called a *fixture*.

This data is passed to our tests as a second parameter, following the test name. In my tests, I've called this parameter `fixture`. I then access the individual fields using the `fixture.list` syntax.

In the code here I passed a block to `setup`. You can also pass the name of a function (as an atom).

Inside the setup code you can define callbacks using `on_exit`. These will be invoked at the end of the test. They can be used to undo changes made by the test.

There's a lot of depth in ExUnit. I'd recommend spending a little time in the ExUnit docs.¹

Property-Based Testing

When you use assertions, you work out ahead of time the result you expect your function to return. This is good, but it also has some limitations. In particular, any assumptions you made while writing the original code are likely to find their way into the tests, too.

A different approach is to consider the overall *properties* of the function you're testing. For example, if your function converts a string to uppercase, then you can predict that, whatever string you feed it,

- the length of the output will be the same as the length of the input, and
- any lowercase characters in the input will have been replaced by their uppercase counterpart.

These are intrinsic properties of the function. And we can test them statistically by simply injecting a (large) number of different strings and verifying the results honor the properties. If all the tests pass, we haven't proved the

1. http://elixir-lang.org/docs/stable/ex_unit/ExUnit.html

function is correct (although we have a lot of confidence it should be). But, more importantly, if any of the tests fail, we've found a boundary condition our function doesn't handle. And property-based testing is surprisingly good at finding these errors. Let's look again at our previous example:

```
tooling/pbt/lib/stats.ex
defmodule Stats do
  def sum(vals),    do: vals |> Enum.reduce(0, &+/2)
  def count(vals), do: vals |> length
  def average(vals), do: sum(vals) / count(vals)
end
```

Here are some simple properties we could test:

- The sum of a list containing a single value should be that value.
- The count function should never return a negative number.
- If we multiply the results returned by count and average, it should equal the result returned by sum (allowing for a little rounding).

To test the properties, the framework needs to generate large numbers of sample values of the correct type. For the first test, for example, we need a bunch of numeric values.

That's where the property-testing libraries come in. There are a number of property-based testing libraries for Elixir (including one I wrote, called Quixir). But here we'll be using a library called StreamData.² As José Valim is one of the authors, I suspect it may well make its way into core Elixir one day.

I could write something like this:

```
check all number <- real() do
  # ...
end
```

There are two pieces of magic here. The first is the `real` function. This is a generator, which will return real numbers. This is invoked by `check all`. You might think from its name that this will try all real numbers (which would take some time), but it instead just iterates a given number of times (100 by default).

Let's code this. First, add StreamData to our list of dependencies:

```
tooling/pbt/mix.exs
defp deps do
  [
    { :stream_data, ">= 0.0.0" },
  ]
end
```

2. https://github.com/whatyouhid/stream_data

Now we can write the property tests. Here's the first:

```
tooling/pbt/test/stats_property_test.exs
defmodule StatsPropertyTest do
  use ExUnit.Case
  use ExUnitProperties

  describe "Stats on lists of ints" do
    property "single element lists are their own sum" do
      check all number <- integer() do
        assert Stats.sum([number]) == number
      end
    end
  end
end
```

You'll see this looks a lot like a regular test. We have to include use ExCheck at the top to include the property test framework.

The actual test is in the property block. It has the check all block we saw earlier. Inside this we have a test: assert Stats.sum([number]) == number.

Let's run it:

```
$ mix test test/stats_property_test.exs
.
.
.
Finished in 0.1 seconds
1 property, 0 failures
```

Let's break the test, just to see what a failure looks like:

```
check all number <- real do
  assert Stats.sum([number]) == number + 1
end

1) property Stats on lists of ints single-element lists are
   their own sum (StatsPropertyTest)
   test/stats_property_test.exs:17
   Failed with generated values (after 0 successful run(s)):

     number <- integer()
     #=> 0

   Assertion with == failed
   code: assert Stats.sum([number]) == number + 1
   left: 0
   right: 1
```

We failed, and the value of number at the time was zero.

Let's fix the test, and add tests for the other two properties.

```
tooling/pbt/test/stats_property_test.exs
property "count not negative" do
  check all l <- list_of(integer()) do
    assert Stats.count(l) >= 0
  end
end

property "single element lists are their own sum" do
  check all number <- integer() do
    assert Stats.sum([number]) == number
  end
end

property "sum equals average times count" do
  check all l <- list_of(integer()) do
    assert_in_delta(
      Stats.sum(l),
      Stats.count(l)*Stats.average(l),
      1.0e-6
    )
  end
end
```

The two new tests use a different generator: `list(int)` generates a number of lists, each containing zero or more ints.

Running this code is surprising—it fails!

```
$ mix test test/stats_property_test.exs
....
1) property Stats on lists of ints sum equals average times
   count (StatsPropertyTest)
   test/stats_property_test.exs:27
   ** (ExUnitProperties.Error) failed with generated values
   (after 40 successful run(s)):

▶  l <- list_of(integer())
▶  #=> []
▶  ** (ArithmetricError) bad argument in arithmetic expression
   code: check all l <- list_of(integer()) do
     ...
Finished in 0.1 seconds
5 properties, 1 failure

Randomized with seed 947482
```

The exception shows we failed with an `ArithmetricError`, and the value that caused the failure was `l = []`, the empty list. That's because we were trying to find the average of an empty list. This means our code will be dividing the sum (0) by the list size (0), and dividing by zero is an error.

This is cool. The property tests explored the range of possible input values, and found one that causes our code to fail.

Arguably, this is a bug in our Stats module. But let's treat it instead as a boundary condition that the tests should avoid. We can do this in two ways.

First, we can tell the property test to skip generated values that fail to meet a condition. We do this with the `nonempty` function:

```
tooling/pbt/test/stats_property_test.exs
property "sum equals average times count (nonempty)" do
  check all l <- list_of(integer()) |> nonempty do
    assert_in_delta(
      Stats.sum(l),
      Stats.count(l)*Stats.average(l),
      1.0e-6
    )
  end
end
```

Now, whenever the generator returns an empty list, the `nonempty` function will filter it out. This is just one example of StreamData filters. A number are predefined, and you can also write your own.

A second approach is to prevent the generator from creating empty lists in the first place. This uses the `min_length` option:

```
tooling/pbt/test/stats_property_test.exs
property "sum equals average times count (min_length)" do
  check all l <- list_of(integer(), min_length: 1) do
    assert_in_delta(
      Stats.sum(l),
      Stats.count(l)*Stats.average(l),
      1.0e-6
    )
  end
end
```

Digging Deeper

In case you're interested in exploring property-based testing, the documentation for ExUnitProperties has some examples and references.³

The StreamData module is designed to be used on its own—it's not just for testing. If you find yourself needing to generate streams of values that meet some criteria, it might be your library of choice.

3. https://hexdocs.pm/stream_data/ExUnitProperties.html

Test Coverage

Some people believe that if there are any lines of application code that haven't been exercised by a test, the code is incomplete. (I'm not one of them.) These folks use *test coverage* tools to check for untested code.

Here we'll use excoveralls to see where to add tests for the Issues app.⁴ (Another good coverage tool is coverex.)⁵

All of the work to add the tool to our project takes place in mix.exs.

First, we add the dependency:

```
tooling/issues/mix.exs
defp deps do
  [
    {:httpoison,   "~> 0.9"},
    {:poison,      "~> 2.2"},
    {:ex_doc,      "~> 0.12"},
    {:earmark,     "~> 1.0",  override: true},
    ▶  {:excoveralls, "~> 0.5.5", only: :test}
  ]
end
```

Then, in the project section, we integrate the various coveralls commands into mix, and force them to run in the test environment:

```
tooling/issues/mix.exs
def project do
  [
    app:           :issues,
    version:       "0.0.1",
    name:          "Issues",
    source_url:    "https://github.com/pragdave/issues",
    escript:        escript_config(),
    build_embedded: Mix.env == :prod,
    start_permanent: Mix.env == :prod,
    ▶  test_coverage: [tool: ExCoveralls],
    ▶  preferred_cli_env: [
    ▶    "coveralls": :test,
    ▶    "coveralls.detail": :test,
    ▶    "coveralls.post": :test,
    ▶    "coveralls.html": :test
    ▶  ],
    deps:          deps()
  ]
end
```

4. <https://github.com/parrot/excoveralls>

5. <https://github.com/alfert/coverex>

After a quick mix deps.get, you can run your first coverage report:

```
$ mix coveralls
.....
Finished in 0.1 seconds
5 doctests, 8 tests, 0 failures

Randomized with seed 5441
-----
COV    FILE                      LINES  RELEVANT  MISSED
 0.0% lib/issues.ex                5        0        0
46.7% lib/issues/cli.ex            73       15        8
 0.0% lib/issues/github_issues.ex  46        6        6
100.0% lib/issues/table_formatter.ex 109      15        0
[TOTAL] 61.1%
-----
```

It runs the tests first, and then reports on the files in our application.

We have no tests for issues.ex. As this is basically a boilerplate no-op, that's not surprising. We wrote some tests for cli.ex, but could do better. The github_issues.ex file is not being tested. But, saving the best for last, we have 100% coverage in the table formatter (because we used it as an example of doc testing).

excoveralls can produce detailed reports to the console (mix coveralls.detail) and as an HTML file (mix coveralls.html). The latter generates the file cover/excoveralls.html, as shown in the following figure.

FILE	COV	LINES	RELEVANT	MISSED
lib/issues.ex	0.0%	5	0	0
lib/issues/cli.ex	46.7%	73	15	8
lib/issues/github_issues.ex	0.0%	46	6	6
lib/issues/table_formatter.ex	100.0%	109	15	0
[TOTAL]	61.1%			

Finally, excoveralls works with a number of continuous integration systems. See its GitHub page for details.

Code Dependencies

The mix tool is smart when it compiles your project. It analyzes the dependencies between your source files, and only recompiles a file when it has changed or a file it depends on has changed. As a developer, you can also access this dependency information, gaining valuable insights into your code. You do this with the mix xref commands.

mix xref unreachable

List functions that are unknown at the time they are called.

mix xref warnings

List warnings associated with dependencies (for example, calls to unknown functions).

mix xref callers Mod | Mod.func | Mod.func/arity

List the callers to a module or function:

```
$ mix xref callers Logger
mix xref callers Logger
web/controllers/page_controller.ex:1: Logger.bare_log/3
web/controllers/page_controller.ex:1: Logger.debug/1
lib/webapp/endpoint.ex:1: Logger.bare_log/3
lib/webapp/endpoint.ex:1: Logger.error/1
```

mix xref graph

Show the dependency tree for the application:

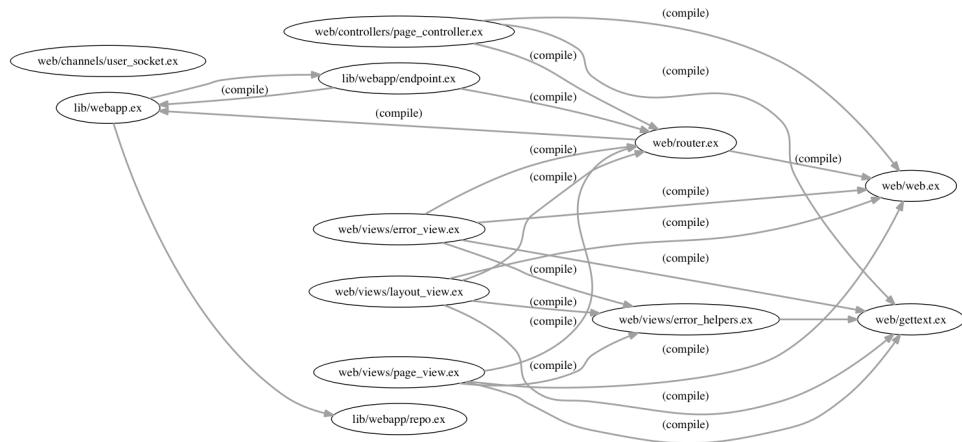
```
$ mix xref graph
lib/webapp.ex
└── lib/webapp/endpoint.ex
    ├── lib/webapp.ex (compile)
    │   └── web/router.ex (compile)
    │       └── lib/webapp.ex (compile)
    └── web/web.ex (compile)
lib/webapp/repo.ex
```

You can produce a circles-and-arrows picture of dependencies using dot.⁶

```
$ mix xref graph --format dot
$ dot -Grankdir=LR -Epenwidth=2 -Ecolor="#a0a0a0" \
    -Tpng xref_graph.dot -o xref_graph.png
```

6. <http://www.graphviz.org/>

This produces something like this:



Server Monitoring

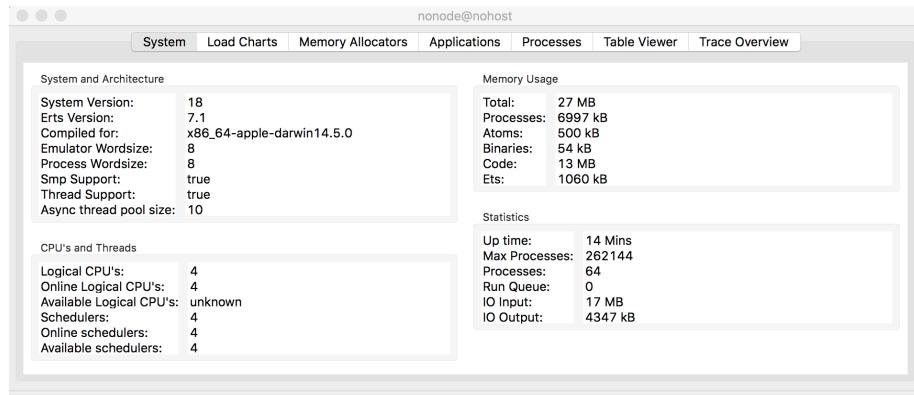
As you might expect from a platform that has been running demanding and critical applications for 20 years, Erlang has great server-monitoring tools.

One of the easiest to use is already baked in. Inside IEx, run

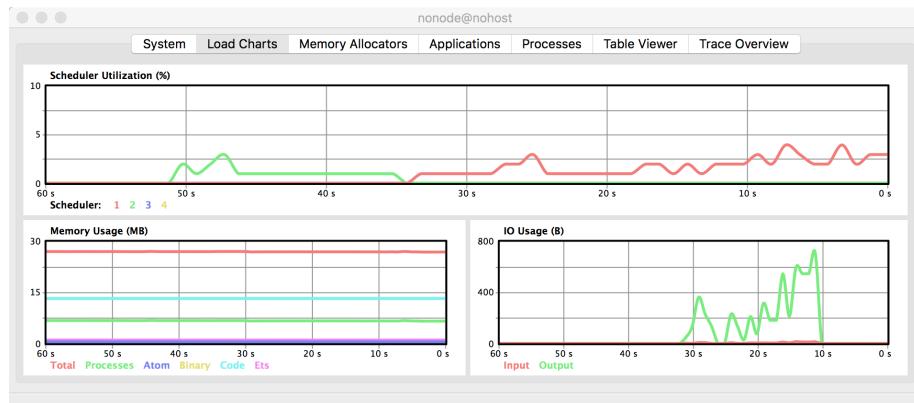
```
iex> :observer.start()
```

Use this to get insight into...

- Basic system information:



- Dynamic charts of load:



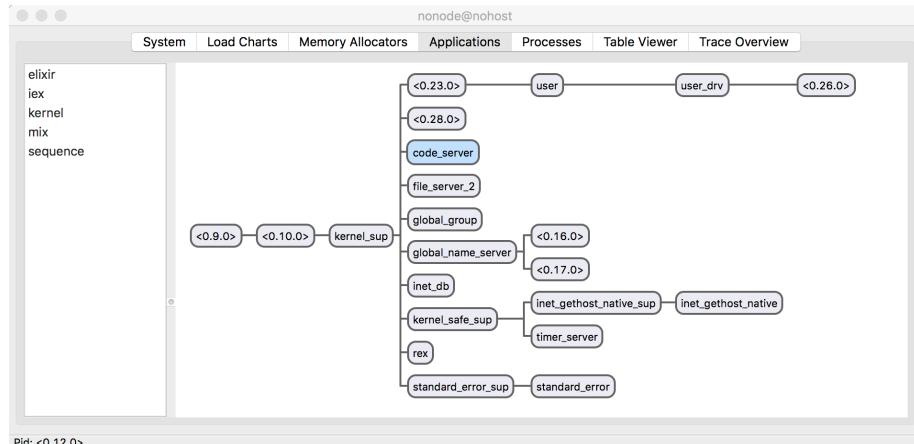
- Information and contents of Erlang ETS tables:

Table Name	Table Id	Objects	Size (kB)	Owner Pid	Owner Name
Elixir.IEx.Config		1	2	<0.43.0>	
elixir_config		11	3	<0.36.0>	
elixir_modules		0	2	<0.39.0>	elixir_code_server

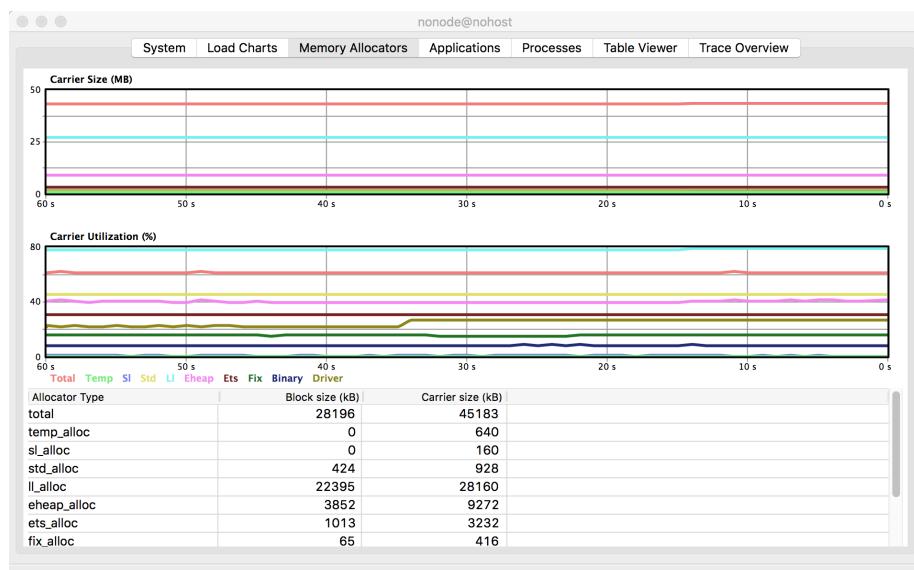
- Running processes:

Pid	Name or Initial Func	Reds	Memory	MsgQ	Current Function
<0.80.0>	wxe_server:init/1	134697	27664	0	gen_server:loop/6
<0.3.0>	erl_prim_loader	48972	122304	0	erl_prim_loader:loop/3
<0.89.0>	erlang:apply/2	36060	42344	0	observer_pro_wx:table_holder/1
<0.100.0>	gen:init_it/6	11654	42408	0	wx_object:loop/6
<0.88.0>	gen:init_it/6	6666	16904	0	wx_object:loop/6
<0.99.0>	gen:init_it/6	5653	16824	0	wx_object:loop/6
<0.12.0>	code_server	3852	263960	0	code_server:loop/1
<0.83.0>	gen:init_it/6	2596	122216	0	wx_object:loop/6
<0.82.0>	erlang:apply/2	1672	2720	0	timer:sleep/1
<0.79.0>	observer	1594	14056	0	wx_object:loop/6
<0.77.0>	erlang:apply/2	261	21680	0	Elixir.IEx.Evaluator:loop/3
<0.87.0>	gen:init_it/6	111	11936	0	wx_object:loop/6
<0.24.0>	user_drv	92	21720	0	user_drv:server_loop/6
<0.101.0>	erlang:apply/2	92	2752	0	io:execute_request/2
<0.84.0>	timer_server	87	7056	0	gen_server:loop/6

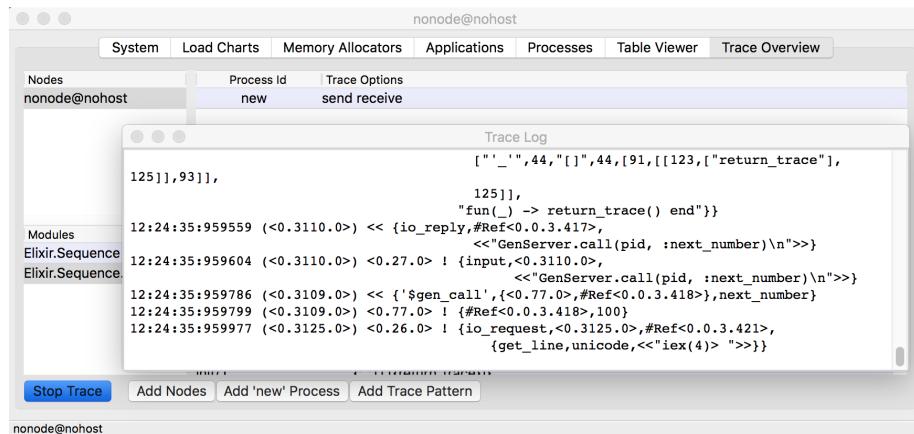
- Running applications:



- Memory allocation:



- And tracing of function calls, messages, and events:



For application-level monitoring, you might want to look at Elixometer from Pinterest.⁷

Source-Code Formatting

This is the section where I get into trouble.

The Elixir core team wanted to standardize the format of source code that was submitted to them for inclusion in the various Elixir core projects. Rather than beat people up and reject pull requests, they made it easy for submitters by including a source-code formatting tool in Elixir 1.6. This tool is pretty smart—it knows not just the syntax of Elixir but also the parse tree, meaning that it will often move things between lines, drop commas, add parentheses, and so on.

This magic is done using the `mix format` command. It can format single files, directory trees, and whole projects (see `mix help format` for information). This formatting replaces the files it touches, so you might want to make sure you're checked in before running it.

Let's have a look at some before-and-after formatting:

If we feed it code that looks like a dog's dinner:

```
def no_vowels string
  do
    string |>
      String.replace(~r/[aeiou]/, "*")
  end
```

7. <https://github.com/pinterest/elixometer>

```
def separator(column_widths) do
  map_join(column_widths, "-+-", fn width ->
    List.duplicate("-", width)
  end)
  end
```

the formatter tidies it nicely:

```
def no_vowels(string) do
  string
  |> String.replace(~r/[aeiou]/, "*")
end

def separator(column_widths) do
  map_join(column_widths, "-+-", fn width ->
    List.duplicate("-", width)
  end)
end
```

It is also pretty smart about multiline constructs:

```
@names [
  Doc, Grumpy, Happy,
  Sleepy, Bashful, Sneezy,
  Dopey
]
```

This produces:

```
@names [
  Doc,
  Grumpy,
  Happy,
  Sleepy,
  Bashful,
  Sneezy,
  Dopey
]
```

Here, because the original split onto a new line, the formatted result was normalized into one element per line.

There's lots to like about the formatter. But I personally don't use it, because it destroys some elements of layout I think are important.

For example, I like to line things up vertically. I find this much, much easier to read and to maintain. Most editors have support for this:

```
options = %{
  style: "light",
  background: "green"
}
```

and:

```
name      = "Alphabet"
url       = "https://abc.xyz"
entry_count = 10
```

Unfortunately, the Elixir formatter goes to some trouble to remove that extra space, producing:

```
options = %{
  style: "light",
  background: "green"
}

name = "Alphabet"
url = "https://abc.xyz"
entry_count = 10
```

Then there's the contentious trailing comma.

When I list things out, I put a comma after each item and put each item on a new line. This makes it easier to move things around, add new items, sort the list, and so on. Each line is just like the other: the first and last lines are not special.

```
plugins = [
  Format,
  Index,
  Print,
]
```

The formatter thinks this is silly, and removes the comma.

```
plugins = [
  Format,
  Index,
  Print
]
```

And finally, there's the trailing comment. I rarely use comments inside a block of code. When I do, and if it is short, I add it to the end of the line:

```
def format(template,  # a binary in eex format
          bindings,        # the bindings to use
          options) do      # :verbose | :narrow
  # ...
end
```

I know this is a bad example, but even so the formatter makes some unfortunate decisions:

```
# a binary in eex format
def format(
    template,
    # the bindings to use
    bindings,
    # :verbose | :narrow
    options
) do
  # ...
end
```

The upshot? If you like what it does, or if you're submitting code to a project that requires it, use the formatter.

Inevitably, There's More

Elixir is lucky when it comes to tooling, both because it inherits a wealth of tools from Erlang and because the Elixir community values great tools and develops them to fill any gaps. Keep up with the tools people use, and you'll find yourself developing faster, and with more confidence.

Now let's look at concurrent programming, a key strength of Elixir.

Part II

Concurrent Programming

You want to write concurrent programs. That's probably why you're reading this book.

Let's look at Elixir's actor-based concurrency model. Then we'll dig into OTP, the Erlang management architecture that helps you create applications that are highly scalable and very reliable.

All your objects want to grow up to be actors.

Reply Retweet Favorite More

1:16 AM - 16 Jun 13 from Limoux, Aude

Chad Fowler @chadfowler
@aeden sadly, most just end up waiting tables

Details Reply Retweet Favorite More

CHAPTER 15

Working with Multiple Processes

One of Elixir's key features is the idea of packaging code into small chunks that can be run independently and concurrently.

If you've come from a conventional programming language, this may worry you. Concurrent programming is "known" to be difficult, and there's a performance penalty to pay when you create lots of processes.

Elixir doesn't have these issues, thanks to the architecture of the Erlang VM on which it runs.

Elixir uses the *actor* model of concurrency. An actor is an independent process that shares nothing with any other process. You can spawn new processes, send them messages, and receive messages back. And that's it (apart from some details about error handling and monitoring, which we cover later).

In the past, you may have had to use threads or operating system processes to achieve concurrency. Each time, you probably felt you were opening Pandora's box—there was so much that could go wrong. But that worry just evaporates in Elixir. In fact, Elixir developers are so comfortable creating new processes, they'll often do it at times when you'd have created an *object* in a language such as Java.

One more thing—when we talk about processes in Elixir, we are not talking about native operating-system processes. These are too slow and bulky. Instead, Elixir uses process support in Erlang. These processes will run across all your CPUs (just like native processes), but they have very little overhead. As we'll cover a bit later, it's very easy to create hundreds of thousands of Elixir processes on even a modest computer.

A Simple Process

Here's a module that defines a function we'd like to run as a separate process:

```
spawn/spawn-basic.ex
defmodule SpawnBasic do
  def greet do
    IO.puts "Hello"
  end
end
```

Yup, that's it. There's nothing special—it's just regular code.

Let's fire up IEx and play:

```
iex> c("spawn-basic.ex")
[SpawnBasic]
```

First let's call it as a regular function:

```
iex> SpawnBasic.greet
Hello
:ok
```

Now let's run it in a separate process:

```
iex> spawn(SpawnBasic, :greet, [])
Hello
#PID<0.42.0>
```

The `spawn` function kicks off a new process. It comes in many forms, but the two simplest ones let you run an anonymous function and run a named function in a module, passing a list of arguments. (We used the latter here.)

The `spawn` returns a *process identifier*, normally called a PID. This uniquely identifies the process it creates. (This identifier could be unique among all processes in the world, but here it's just unique in our application.)

When we call `spawn`, it creates a new process to run the code we specify. We don't know exactly when it will execute—only that it is eligible to run.

In this example, we can see that our function ran and output "Hello" prior to IEx reporting the PID returned by `spawn`. But you can't rely on this. Instead you'll use messages to synchronize your processes' activity.

Sending Messages Between Processes

Let's rewrite our example to use messages. The top level will send `greet` a message containing a string, and the `greet` function will respond with a greeting containing that message.

In Elixir we send a message using the `send` function. It takes a PID and the message to send (an Elixir value, which we also call a *term*) on the right. You can send anything you want, but most Elixir developers seem to use atoms and tuples.

We wait for messages using `receive`. In a way, this acts the same as `case`, with the message body as the parameter. Inside the block associated with the `receive` call, you can specify any number of patterns and associated actions. Just as with `case`, the action associated with the first pattern that matches the function is run.

Here's the updated version of our `greet` function.

```
spawn/spawn1.exs
defmodule Spawn1 do
  def greet do
    receive do
      {sender, msg} ->
        send sender, { :ok, "Hello, #{msg}" }
    end
  end
end

# here's a client
pid = spawn(Spawn1, :greet, [])
send pid, {self(), "World!"}

receive do
  {:ok, message} ->
    IO.puts message
end
```

The function uses `receive` to wait for a message, and then matches the message in the block. In this case, the only pattern is a two-element tuple, where the first element is the original sender's PID and the second is the message. In the corresponding action, we use `send sender, ...` to send a formatted string back to the original message sender. We package that string into a tuple, with `:ok` as its first element.

Outside the module, we call `spawn` to create a process, and send it a tuple:

```
send pid, { self, "World!" }
```

The function `self` returns its caller's PID. Here we use it to pass our PID to the `greet` function so it will know where to send the response.

We then wait for a response. Notice that we do a pattern match on `{:ok, message}`, extracting the second element of the tuple, which contains the actual text.

We can run this in IEx:

```
iex> c("spawn1.exs")
Hello, World!
[Spawn1]
```

Very cool. The text was sent, and greet responded with the full greeting.

Handling Multiple Messages

Let's try sending a second message.

```
spawn/spawn2.exs
defmodule Spawn2 do
  def greet do
    receive do
      {sender, msg} ->
        send sender, { :ok, "Hello, #{msg}" }
    end
  end
end

# here's a client
pid = spawn(Spawn2, :greet, [])
send pid, {self(), "World!"}

receive do
  {:ok, message} ->
    IO.puts message
end

send pid, {self(), "Kermit!"}
receive do
  {:ok, message} ->
    IO.puts message
end
```

Run it in IEx:

```
iex> c("spawn2.exs")
Hello, World!
.... just sits there ....
```

The first message is sent back, but the second is nowhere to be seen. What's worse, IEx just hangs, and we have to use ^C (the Control-C key sequence) to get out of it.

That's because our greet function handles only a single message. Once it has processed the receive, it exits. As a result, the second message we send it is never processed. The second receive at the top level then just hangs, waiting for a response that will never come.

Let's at least fix the hanging part. We can tell receive that we want to time out if a response is not received in so many milliseconds. This uses a pseudo-pattern called after.

```
spawn/spawn3.exs
defmodule Spawn3 do
  def greet do
    receive do
      {sender, msg} ->
        send sender, { :ok, "Hello, #{msg}" }
    end
  end
end

# here's a client
pid = spawn(Spawn3, :greet, [])
send pid, {self(), "World!"}
receive do
  {:ok, message} ->
    IO.puts message
end

send pid, {self(), "Kermit!"}
receive do
  {:ok, message} ->
    IO.puts message
  > after 500 ->
    IO.puts "The greeter has gone away"
end

iex> c("spawn3.exs")
Hello, World!
... short pause ...
The greeter has gone away
[Spawn3]
```

But how would we make our greet function handle multiple messages? Our natural reaction is to make it loop, doing a receive on each iteration. Elixir doesn't have loops, but it does have recursion.

```
spawn/spawn4.exs
defmodule Spawn4 do
  def greet do
    receive do
      {sender, msg} ->
        send sender, { :ok, "Hello, #{msg}" }
    > greet()
    end
  end
end
```

```
# here's a client
pid = spawn(Spawn4, :greet, [])
send pid, {self(), "World!"}
receive do
  {:ok, message} ->
    IO.puts message
end

send pid, {self(), "Kermit!"}
receive do
  {:ok, message} ->
    IO.puts message
  after 500 ->
    IO.puts "The greeter has gone away"
end
```

Run this, and both messages are processed:

```
iex> c("spawn4.exs")
Hello, World!
Hello, Kermit!
[Spawn4]
```

Recursion, Looping, and the Stack

The greet function might have worried you a little. Every time it receives a message, it ends up calling itself. In many languages, that adds a new frame to the stack. After a large number of messages, you might run out of memory.

This doesn't happen in Elixir, as it implements *tail-call optimization*. If the last thing a function does is call itself, there's no need to make the call. Instead, the runtime simply jumps back to the start of the function. If the recursive call has arguments, then these replace the original parameters. But beware—the recursive call *must* be the very last thing executed. For example, the following code is not tail recursive:

```
def factorial(0), do: 1
def factorial(n), do: n * factorial(n-1)
```

While the recursive call is physically the last thing in the function, it is not the last thing executed. The function has to multiply the value it returns by n .

To make it tail recursive, we need to move the multiplication into the recursive call, and this means adding an accumulator:

```
spawn/fact_tr.exs
defmodule TailRecursive do
  def factorial(n), do: _fact(n, 1)
  defp _fact(0, acc), do: acc
  defp _fact(n, acc), do: _fact(n-1, acc*n)
end
```

Process Overhead

At the start of the chapter, I somewhat cavalierly said Elixir processes were very low overhead. Now it's time to back that up. Let's write some code that creates n processes. The first will send a number to the second. It will increment that number and pass it to the third. This will continue until we get to the last process, which will pass the number back to the top level.

```
spawn/chain.exs
Line 1  defmodule Chain do
-    def counter(next_pid) do
-        receive do
-            n ->
-                send next_pid, n + 1
-        end
-    end
-
-    def create_processes(n) do
10     code_to_run = fn (_,send_to) ->
-        spawn(Chain, :counter, [send_to])
-    end
-
-    last = Enum.reduce(1..n, self(), code_to_run)
15     send(last, 0)      # start the count by sending a zero to the last process
-
-    receive do          # and wait for the result to come back to us
-        final_answer when is_integer(final_answer) ->
-            "Result is #{inspect(final_answer)}"
20     end
-    end
-
-    def run(n) do
25     :timer.tc(Chain, :create_processes, [n])
-        |> IO.inspect
-    end
-  end
```

The counter function on line 2 is the code that will be run in separate processes. It is passed the PID of the next process in the chain. When it receives a number, it increments it and sends it on to that next process.

The create_processes function is probably the densest piece of Elixir we've encountered so far. Let's break it down.

It is passed the number of processes to create. Each process has to be passed the PID of the previous process so that it knows who to send the updated number to.

The code that creates each process is defined in a one-line anonymous function, which is assigned to the variable `code_to_run`. The function takes two parameters because we're passing it to `Enum.reduce` on line 14.

The `reduce` call will iterate over the range `1..n`. Each time around, it will pass an accumulator as the second parameter to its function. We set the initial value of that accumulator to `self`, our PID.

In the function, we spawn a new process that runs the `counter` function, using the third parameter of `spawn` to pass in the accumulator's current value (initially `self`). The value `spawn` returns is the PID of the newly created process, which becomes the accumulator's value for the next iteration.

Putting it another way, each time we spawn a new process, we pass it the previous process's PID in the `send_to` parameter.

The value that the `reduce` function returns is the accumulator's final value, which is the PID of the last process created.

On the next line we set the ball rolling by passing 0 to the last process. The process increments the value and so passes 1 to the second-to-last process. This goes on until the very first process we created passes the result back to us. We use the `receive` block to capture this, and format the result into a nice message.

Our `receive` block contains a new feature. We've already seen how guard clauses can constrain pattern matching and function calling. The same guard clauses can be used to qualify the pattern in a `receive` block.

Why do we need this, though? It turns out there's a bug in some versions of Elixir.¹ When you compile and run a program using `iex -S mix`, a residual message is left lying around from the compilation process (it records a process's termination). We ignore that message by telling the `receive` clause that we're interested only in simple integers.

The `run` function starts the whole thing off. It uses a built-in Erlang library, `tc`, which can time a function's execution. We pass it the module, name, and parameters, and it responds with a tuple. The first element is the execution time in microseconds and the second is the result the function returns.

We'll run this code from the command line rather than from IEx. (You'll see why in a second.) These results are on my 2011 MacBook Air (2.13 GHz Core 2 Duo and 4 GB of RAM).

1. <https://github.com/elixir-lang/elixir/issues/1050>

```
$ elixir -r chain.exs -e "Chain.run(10)"
{4015, "Result is 10"}
```

We asked it to run 10 processes, and it came back in 4 ms. The answer looks correct. Let's try 100 processes.

```
$ elixir -r chain.exs -e "Chain.run(100)"
{4562, "Result is 100"}
```

Only a small increase in the time. There's probably some startup latency on the first process creation. Onward! Let's try 1,000.

```
$ elixir -r chain.exs -e "Chain.run(1_000)"
{8458, "Result is 1000"}
```

Now 10,000.

```
$ elixir -r chain.exs -e "Chain.run(10_000)"
{66769, "Result is 10000"}
```

Ten thousand processes created and executed in 66 ms. Let's try for 400,000.

```
$ elixir -r chain.exs -e "Chain.run(400_000)"
=ERROR REPORT==== 25-Apr-2013::15:16:14 ===
Too many processes
** (SystemLimitError) a system limit has been reached
```

It looks like the virtual machine won't support 400,000 processes. Fortunately, this is not a hard limit—we just bumped into a default value. We can increase this using the VM's +P parameter. We pass this parameter to the VM using the --erl parameter to elixir. (This is why I chose to run from the command line.)

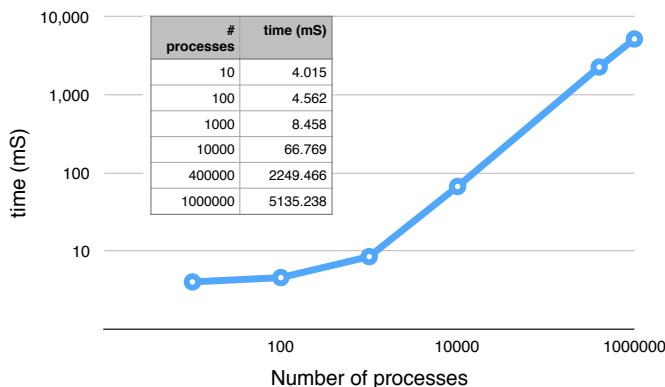
```
$ elixir --erl "+P 1000000" -r chain.exs -e "Chain.run(400_000)"
{2249466, "Result is 400000"}
```

One last run, this time with 1,000,000 processes.

```
$ elixir --erl "+P 1000000" -r chain.exs -e "Chain.run(1_000_000)"
{5135238, "Result is 1000000"}
```

We ran a million processes (sequentially) in just over 5 seconds. And, as the [graph on page 206](#) shows, the time per process was pretty much linear once we overcame the startup time.

This kind of performance is stunning, and it changes the way we design code. We can now create hundreds of little helper processes. And each process can contain its own state—in a way, processes in Elixir are like objects in an object-oriented system (but they're more self-contained).



Your Turn

► *Exercise: WorkingWithMultipleProcesses-1*

Run this code on your machine. See if you get comparable results.

► *Exercise: WorkingWithMultipleProcesses-2*

Write a program that spawns two processes and then passes each a unique token (for example, “fred” and “betty”). Have them send the tokens back.

- Is the order in which the replies are received deterministic in theory?
In practice?
- If either answer is no, how could you make it so?

When Processes Die

Who gets told when a process dies? By default, no one. Obviously the VM knows and can report it to the console, but your code will be oblivious unless you explicitly tell Elixir you want to get involved. Here’s the default case: we spawn a function that uses the Erlang timer library to sleep for 500 ms. It then exits with a status of :boom. The code that spawns it sits in a receive. If it receives a message, it reports that fact; otherwise, after one second it lets us know that nothing happened.

```
spawn/link1.exs
defmodule Link1 do
  import :timer, only: [ sleep: 1 ]

  def sad_function do
    sleep 500
    exit(:boom)
  end
end
```

```

def run do
  spawn(Link1, :sad_function, [])
  receive do
    msg ->
      IO.puts "MESSAGE RECEIVED: #{inspect msg}"
  after 1000 ->
    IO.puts "Nothing happened as far as I am concerned"
  end
end
end

Link1.run

```

(Think about how you'd have written this in your old programming language.)

We can run this from the console:

```
$ elixir -r link1.exs
Nothing happened as far as I am concerned
```

The top level got no notification when the spawned process exited.

Linking Two Processes

If we want two processes to share in each other's pain, we can *link* them. When processes are linked, each can receive information when the other exits. The `spawn_link` call spawns a process and links it to the caller in one operation.

```

spawn/link2.exs
defmodule Link2 do
  import :timer, only: [ sleep: 1 ]

  def sad_function do
    sleep 500
    exit(:boom)
  end

  def run do
    > spawn_link(Link2, :sad_function, [])
    receive do
      msg ->
        IO.puts "MESSAGE RECEIVED: #{inspect msg}"
    after 1000 ->
      IO.puts "Nothing happened as far as I am concerned"
    end
  end
end

Link2.run

```

The runtime reports the abnormal termination:

```
$ elixir -r link2.exs
** (EXIT from #PID<0.73.0>) :boom
```

So our child process died, and it killed the entire application. That's the default behavior of linked processes—when one exits abnormally, it kills the other.

What if you want to handle the death of another process? Well, you probably don't want to do this. Elixir uses the OTP framework for constructing process trees, and OTP includes the concept of process supervision. An incredible amount of effort has been spent getting this right, so I recommend using it most of the time. (We cover this in [Chapter 18, OTP: Supervisors, on page 247](#).)

However, you can tell Elixir to convert the exit signals from a linked process into a message you can handle. Do this by *trapping the exit*.

```
spawn/link3.exs
defmodule Link3 do
  import :timer, only: [ sleep: 1 ]

  def sad_function do
    sleep 500
    exit(:boom)
  end
  def run do
    Process.flag(:trap_exit, true)
    spawn_link(Link3, :sad_function, [])
    receive do
      msg ->
        IO.puts "MESSAGE RECEIVED: #{inspect msg}"
    after 1000 ->
      IO.puts "Nothing happened as far as I am concerned"
    end
  end
end

Link3.run
```

This time we see an :EXIT message when the spawned process terminates:

```
$ elixir -r link3.exs
MESSAGE RECEIVED: {:EXIT, #PID<0.78.0>, :boom}
```

It doesn't matter why a process exits—it may simply finish processing, it may explicitly exit, or it may raise an exception—the same :EXIT message is received. Following an error, however, it contains details of what went wrong.

Monitoring a Process

Linking joins the calling process and another process—each receives notifications about the other. By contrast, *monitoring* lets a process spawn another and be notified of its termination, but without the reverse notification—it is one-way only.

When you monitor a process, you receive a :DOWN message when it exits or fails, or if it doesn't exist.

You can use `spawn_monitor` to turn on monitoring when you spawn a process, or you can use `Process.monitor` to monitor an existing process. However, if you use `Process.monitor` (or link to an existing process), there is a potential race condition—if the other process dies before your monitor call completes, you may not receive a notification. The `spawn_link` and `spawn_monitor` versions are atomic, however, so you'll always catch a failure.

```
spawn/monitor1.exs
defmodule Monitor1 do
  import :timer, only: [ sleep: 1 ]

  def sad_function do
    sleep 500
    exit(:boom)
  end
  def run do
    res = spawn_monitor(Monitor1, :sad_function, [])
    IO.puts inspect res
    receive do
      msg ->
        IO.puts "MESSAGE RECEIVED: #{inspect msg}"
    after 1000 ->
      IO.puts "Nothing happened as far as I am concerned"
    end
  end
end

Monitor1.run
```

Run it, and the results are similar to the `spawn_link` version:

```
$ elixir -r monitor1.exs
{#PID<0.78.0>, #Reference<0.1328...>}
MESSAGE RECEIVED: {:DOWN, #Reference<0.1328...>, :process,
  #PID<0.78.0>, :boom}
```

(The Reference record in the message is the identity of the monitor that was created. The `spawn_monitor` call also returns it, along with the PID.)

So, when do you use links and when should you choose monitors?

It depends on your processes' semantics. If the intent is that a failure in one process should terminate another, then you need links. If instead you need to know when some other process exits for any reason, choose monitors.

Your Turn

The Erlang function `timer.sleep(time_in_ms)` suspends the current process for a given time. You might want to use it to force some scenarios in the following exercises. The key with the exercises is to get used to the different reports you'll see when you're developing code.

► *Exercise: WorkingWithMultipleProcesses-3*

Use `spawn_link` to start a process, and have that process send a message to the parent and then exit immediately. Meanwhile, sleep for 500 ms in the parent, then receive as many messages as are waiting. Trace what you receive. Does it matter that you weren't waiting for the notification from the child when it exited?

► *Exercise: WorkingWithMultipleProcesses-4*

Do the same, but have the child raise an exception. What difference do you see in the tracing?

► *Exercise: WorkingWithMultipleProcesses-5*

Repeat the two, changing `spawn_link` to `spawn_monitor`.

Parallel Map—The “Hello, World” of Erlang

Devin Torres reminded me that every book in the Erlang space must, by law, include a parallel map function. Regular map returns the list that results from applying a function to each element of a collection. The parallel version does the same, but it applies the function to each element in a separate process.

```
spawn/pmap.exs
defmodule Parallel do
  def pmap(collection, fun) do
    me = self()
    collection
    |> Enum.map(fn (elem) ->
      spawn_link fn -> (send me, { self(), fun.(elem) }) end
      end)
    |> Enum.map(fn (pid) ->
      receive do { ^pid, result } -> result end
      end)
    end
  end
```

Our method contains two transformations (look for the `|>` operator). The first transformation maps `collection` into a list of PIDs, where each PID in the list runs the given function on an individual list element. If the collection contains 1,000 items, we'll run 1,000 processes.

The second transformation converts the list of PIDs into the results returned by the processes corresponding to each PID in the list. Note how it uses `^pid` in the receive block to get the result for each PID in turn. Without this we'd get back the results in random order.

But does it work?

```
iex> c("pmap.exs")
[Parallel]
iex> Parallel.pmap 1..10, &(&1 * &1)
[1,4,9,16,25,36,49,64,81,100]
```

That's pretty sweet, but it gets better, as we'll cover [when we look at tasks and agents on page 293](#).

Your Turn

► [Exercise: WorkingWithMultipleProcesses-6](#)

In the pmap code, I assigned the value of `self` to the variable `me` at the top of the method and then used `me` as the target of the message returned by the spawned processes. Why use a separate variable here?

► [Exercise: WorkingWithMultipleProcesses-7](#)

Change the `^pid` in `pmap` to `_pid`. This means the receive block will take responses in the order the processes send them. Now run the code again. Do you see any difference in the output? If you're like me, you don't, but the program clearly contains a bug. Are you scared by this? Can you find a way to reveal the problem (perhaps by passing in a different function, by sleeping, or by increasing the number of processes)? Change it back to `^pid` and make sure the order is now correct.

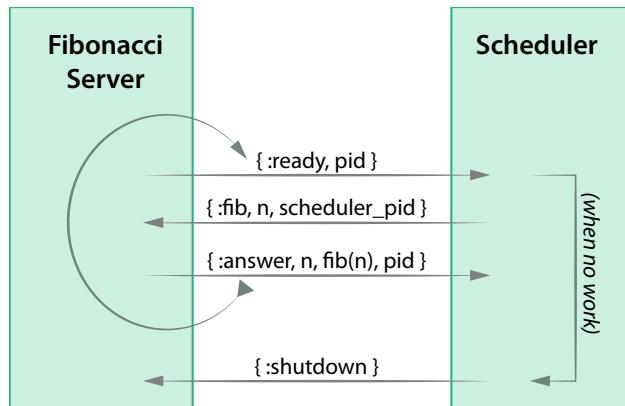
A Fibonacci Server

Let's round out this chapter with an example program. Its task is to calculate $fib(n)$ for a list of n , where $fib(n)$ is the n^{th} Fibonacci number. (The Fibonacci sequence starts 0, 1. Each subsequent number is the sum of the preceding two numbers in the sequence.)² I chose this not because it is something we all do every day, but because the naive calculation of Fibonacci numbers 10 through 37 takes a measurable number of seconds on typical computers.

The twist is that we'll write our program to calculate different Fibonacci numbers in parallel. To do this, we'll write a trivial server process that does

2. http://en.wikipedia.org/wiki/Fibonacci_number

the calculation, and a scheduler that assigns work to a calculation process when it becomes free. The following diagram shows the message flow.



When the calculator is ready for the next number, it sends a `:ready` message to the scheduler. If there is still work to do, the scheduler sends it to the calculator in a `:fib` message; otherwise it sends the calculator a `:shutdown`. When a calculator receives a `:fib` message, it calculates the given Fibonacci number and returns it in an `:answer`. If it gets a `:shutdown`, it simply exits.

Here's the Fibonacci calculator module:

```

spawn/fib.exs
defmodule FibSolver do
  def fib(scheduler) do
    send scheduler, { :ready, self() }
    receive do
      { :fib, n, client } ->
        send client, { :answer, n, fib_calc(n), self() }
        fib(scheduler)
      { :shutdown } ->
        exit(:normal)
    end
  end
  # very inefficient, deliberately
  defp fib_calc(0), do: 0
  defp fib_calc(1), do: 1
  defp fib_calc(n), do: fib_calc(n-1) + fib_calc(n-2)
end

```

The public API is the `fib` function, which takes the scheduler PID. When it starts, it sends a `:ready` message to the scheduler and waits for a message back.

If it gets a :fib message, it calculates the answer and sends it back to the client. It then loops by calling itself recursively. This will send another :ready message, telling the client it is ready for more work.

If it gets a :shutdown it simply exits.

The Task Scheduler

The scheduler is a little more complex, as it is designed to handle both a varying number of server processes and an unknown amount of work.

```
spawn/fib.exs
defmodule Scheduler do
  def run(num_processes, module, func, to_calculate) do
    (1..num_processes)
    |> Enum.map(fn(_) -> spawn(module, func, [self()]) end)
    |> schedule_processes(to_calculate, [])
  end

  defp schedule_processes(processes, queue, results) do
    receive do
      {:ready, pid} when length(queue) > 0 ->
        [next | tail] = queue
        send pid, {:fib, next, self()}
        schedule_processes(processes, tail, results)

      {:ready, pid} ->
        send pid, {:shutdown}
        if length(processes) > 1 do
          schedule_processes(List.delete(processes, pid), queue, results)
        else
          Enum.sort(results, fn {n1,_}, {n2,_} -> n1 <= n2 end)
        end

      {:answer, number, result, _pid} ->
        schedule_processes(processes, queue, [ {number, result} | results ])
    end
  end
end
```

The public API for the scheduler is the `run` function. It receives the number of processes to spawn, the module and function to spawn, and a list of things to process. The scheduler is pleasantly ignorant of the actual task being performed.

Let's emphasize that last point. Our scheduler knows nothing about Fibonacci numbers. Exactly the same code will happily manage processes working on DNA sequencing or password cracking.

The run function spawns the correct number of processes and records their PIDs. It then calls the workhorse function, schedule_processes.

This function is basically a receive loop. If it gets a :ready message from a server, it sees if there is more work in the queue. If there is, it passes the next number to the calculator and then recurses with one fewer number in the queue.

If the work queue is empty when it receives a :ready message, it sends a shutdown to the server. If this is the last process, then we're done and it sorts the accumulated results. If it isn't the last process, it removes the process from the list of processes and recurses to handle another message.

Finally, if it gets an :answer message, it records the answer in the result accumulator and recurses to handle the next message.

We drive the scheduler with the following code:

```
spawn/fib.exs
to_process = List.duplicate(37, 20)

Enum.each 1..10, fn num_processes ->
  {time, result} = :timer.tc(
    Scheduler, :run,
    [num_processes, FibSolver, :fib, to_process]
  )

  if num_processes == 1 do
    IO.puts inspect result
    IO.puts "\n# time (s)"
  end
  :io.format "~2B ~.2f~n", [num_processes, time/1000000.0]
end
```

The to_process list contains the numbers we'll be passing to our fib servers. In our case, we give it the same number, 37, 20 times. The intent here is to load each of our processors.

We run the code a total of 10 times, varying the number of spawned processes from 1 to 10. We use :timer.tc to determine the elapsed time of each iteration, reporting the result in seconds. The first time around the loop, we also display the numbers we calculated.

```
$ elixir fib.exs
[{:ok, 24157817}, {:ok, 24157817}, {:ok, 24157817}, ...]

# time (s)
1    21.22
2    11.24
3    7.99
4    5.89
5    5.95
```

```

6    6.40
7    6.00
8    5.92
9    5.84
10   5.85

```

Cody Russell kindly ran this for me on his four-core system. He saw a dramatic reduction in elapsed time when we increase the concurrency from one to two, small decreases until we hit four processes, then fairly flat performance after that. The Activity Monitor showed a consistent 380% CPU use once the concurrency got above 4. (If you want to see similar results on systems with more cores, you'll need to increase the number of entries in the `to_process` list.)

Your Turn

► *Exercise: WorkingWithMultipleProcesses-8*

Run the Fibonacci code on your machine. Do you get comparable timings? If your machine has multiple cores and/or processors, do you see improvements in the timing as we increase the application's concurrency?

► *Exercise: WorkingWithMultipleProcesses-9*

Take this scheduler code and update it to let you run a function that finds the number of times the word “cat” appears in each file in a given directory. Run one server process per file. The function `File.ls!` returns the names of files in a directory, and `File.read!` reads the contents of a file as a binary. Can you write it as a more generalized scheduler?

Run your code on a directory with a reasonable number of files (maybe around 100) so you can experiment with the effects of concurrency.

Agents—A Teaser

Our Fibonacci code is really inefficient. To calculate `fib(5)`, we calculate this:

```

fib(5)
= fib(4)                                + fib(3)
= fib(3)      + fib(2)      + fib(2)      + fib(1)
= fib(2)      + fib(1) + fib(1) + fib(0) + fib(1) + fib(0) + fib(1)
= fib(1) + fib(0) + fib(1) + fib(1) + fib(0) + fib(1) + fib(0) + fib(1)

```

Look at all that duplication. If only we could cache the intermediate values.

As you know, Elixir modules are basically buckets of functions—they cannot hold state. But processes can hold state. And Elixir comes with a library module called `Agent` that makes it easy to wrap a process containing state in a nice module interface. Don't worry about the details of the code that follows—we

cover [agents and tasks on page 293](#). For now, just see how processes are among the tools we use to add persistence to Elixir code. (This code comes from a mailing-list post by José Valim, written in response to some ugly code I wrote.)

```
spawn/fib_agent.exs
defmodule FibAgent do
  def start_link do
    Agent.start_link(fn -> %{ 0 => 0, 1 => 1 } end)
  end

  def fib(pid, n) when n >= 0 do
    Agent.get_and_update(pid, &do_fib(&l, n))
  end

  defp do_fib(cache, n) do
    case cache[n] do
      nil ->
        { n_1, cache } = do_fib(cache, n-1)
        result           = n_1 + cache[n-2]
        { result, Map.put(cache, n, result) }

      cached_value ->
        { cached_value, cache }
    end
  end
end

{:ok, agent} = FibAgent.start_link()
IO.puts FibAgent.fib(agent, 2000)
```

Let's run it:

```
$ elixir fib_agent.exs
42246963333923048787067256023414827825798528402506810980102801373143085843701
30707224123599639141511088446087538909603607640194711643596029271983312598737
3262535580260699158591522949245390499872225679531698287448247299226390183371
67780606070116154978867198798583114688708762645973690867228840236544222952433
47964480139515349562972087652656069529806499841977448720155612802665404554171
717881930324025204312082516817125
```

If we'd tried to calculate `fib(2000)` using the noncached version, the sun would grow to engulf the Earth while we were waiting for it to finish.

Thinking in Processes

If you first started programming with procedural languages and then moved to an object-oriented style, you'll have experienced a period of dislocation as you tried to get your head to think in terms of objects.

The same will be happening now as you start to think of your work in terms of processes. Just about every decent Elixir program will have many, many

processes, and by and large they'll be just as easy to create and manage as the objects were in object-oriented programming. But learning to think that way takes awhile. Stick with it.

So far we've been running our processes in the same VM. But if we're planning on taking over the world, we need to be able to scale. And that means running on more than one machine.

The abstraction for this is the *node*, and that's the subject of the next chapter.

Nodes—The Key to Distributing Services

There's nothing mysterious about a *node*. It is simply a running Erlang VM. Throughout this book we've been running our code on a node.

The Erlang VM, called *Beam*, is more than a simple interpreter. It's like its own little operating system running on top of your host operating system. It handles its own events, process scheduling, memory, naming services, and interprocess communication. In addition to all that, a node can connect to other nodes—in the same computer, across a LAN, or across the Internet—and provide many of the same services across these connections that it provides to the processes it hosts locally.

Naming Nodes

So far we haven't needed to give our node a name—we've had only one. If we ask Elixir what the current node is called, it'll give us a made-up name:

```
iex> Node.self  
:nonode@nohost
```

We can set the name of a node when we start it. With IEx, use either the `--name` or `--sname` option. The former sets a fully qualified name:

```
$ iex --name wibble@light-boy.local  
iex(wibble@light-boy.local)> Node.self  
:"wibble@light-boy.local"
```

The latter sets a short name:

```
$ iex --sname wobble  
iex(wobble@light-boy)> Node.self  
:"wobble@light-boy"
```

The name that's returned is an atom—it's in quotes because it contains characters not allowed in a literal atom.

Note that in both cases the IEx prompt contains the node's name along with my machine's name (`light-boy`).

If You Run OS X

Apple did something strange a while back—the local hostname is resolved only if you have particular sharing services enabled. If they aren't enabled, then you can't access your computer using its name. This means that Elixir can't find it when using the `--sname` option.

The simplest fix, which is a bit of a hack, is to add your machine's name to your `/etc/hosts` file.



First find the name:

```
$ scutil --get LocalHostName
«your-computer's-name»
```

Then edit `/etc/hosts` (you'll need to use sudo) and add this line:

```
127.0.0.1   «your-computer's-name»
```

Now I want to show you what happens when we have two nodes running. The easiest way to do this is to open two terminal windows and run a node in each. To represent these windows in the book, I'll show them stacked vertically.

Let's run a node called `node_one` in the top window and `node_two` in the bottom one. We'll then use the Elixir Node module's `list` function to display a list of known nodes, then connect from one to the other.

Window #1

```
$ iex --sname node_one
iex(node_one@light-boy)>
```

Window #2

```
$ iex --sname node_two
iex(node_two@light-boy)> Node.list
[]
iex(node_two@light-boy)> Node.connect :"node_one@light-boy"
true
iex(node_two@light-boy)> Node.list
[:"node_one@light-boy"]
```

Initially, `node_two` doesn't know about any other nodes. But after we connect to `node_one` (notice that we pass an atom containing that node's name), the list shows the other node. And if we go back to `node one`, it will now know about `node two`.

```
iex(node_one@light-boy)> Node.list
[:"node_two@light-boy"]
```

Now that we have two nodes, we can try running some code. On node one, let's create an anonymous function that outputs the current node name.

```
iex(node_one@light-boy)> func = fn -> IO.inspect Node.self end
#Function<erl_eval.20.82930912>
```

We can run this with the `spawn` function.

```
iex(node_one@light-boy)> spawn(func)
#PID<0.59.0>
node_one@light-boy
```

But `spawn` also lets us specify a node name. The process will be spawned on that node.

```
iex(node_one@light-boy)> Node.spawn(:"node_one@light-boy", func)
#PID<0.57.0>
node_one@light-boy
iex(node_one@light-boy)> Node.spawn(:"node_two@light-boy", func)
#PID<7393.48.0>
node_two@light-boy
```

We're running on node one. When we tell `spawn` to run on `node_one@light-boy`, we see two lines of output. The first is the PID `spawn` returns, and the second is the value of `Node.self` that the function writes.

The second `spawn` is where it gets interesting. We pass it the name of node two and the same function we used the first time. Again we get two lines of output. The first is the PID and the second is the node name. Notice the PID's contents. The first field in a PID is the node number. When running on a local node, it's zero. But here we're running on a remote node, so that field has a positive value (7393). Then look at the function's output. It reports that it is running on node two. I think that's pretty cool.

You may have been expecting the output from the second `spawn` to appear in the lower window. After all, the code runs on node two. But it was created on node one, so it inherits its process hierarchy from node one. Part of that hierarchy is something called the *group leader*, which (among other things) determines where `IO.puts` sends its output. So in a way, what we're seeing is doubly impressive. We start on node one, run a process on node two, and when the process outputs something, it appears back on node one.

Your Turn

► *Exercise: Nodes-1*

Set up two terminal windows, and go to a different directory in each. Then start up a named node in each. In one window, write a function that lists the contents of the current directory.

```
fun = fn -> IO.puts(Enum.join(File.ls!, ", ")) end
```

Run it twice, once on each node.

Nodes, Cookies, and Security

Although this is cool, it might also ring some alarm bells. If you can run arbitrary code on any node, then anyone with a publicly accessible node has just handed over his machine to any random hacker.

But that's not the case. Before a node will let another connect, it checks that the remote node has permission. It does that by comparing that node's *cookie* with its own cookie. A cookie is just an arbitrary string (ideally fairly long and very random). As an administrator of a distributed Elixir system, you need to create a cookie and then make sure all nodes use it.

If you are running the `iex` or `elixir` commands, you can pass in the cookie using the `--cookie` option.

```
$ iex --sname one --cookie chocolate-chip
iex(one@light-boy)> Node.get_cookie
:"chocolate-chip"
```

If we repeat our two-node experiment and explicitly set the cookie names to be different, what happens?

Window #1

```
$ iex --sname node_one --cookie cookie-one
iex(node_one@light-boy)> Node.connect :"node_two@light-boy"
false
```

Window #2

```
$ iex --sname node_two --cookie cookie-two
iex(node_two@light-boy)>
=ERROR REPORT==== 27-Apr-2013::21:27:43 ===
** Connection attempt from disallowed node 'node_one@light-boy' **
```

The node that attempts to connect receives `false`, indicating the connection was not made. And the node that it tried to connect to logs an error describing the attempt.

But why does it succeed when we don't specify a cookie? When Erlang starts, it looks for an `.erlang.cookie` file in your home directory. If that file doesn't exist, Erlang creates it and stores a random string in it. It uses that string as the cookie for any node the user starts. That way, all nodes you start on a particular machine are automatically given access to each other.

Be careful when connecting nodes over a public network—the cookie is transmitted in plain text.

Naming Your Processes

Although a PID is displayed as three numbers, it contains just two fields; the first number is the node ID and the next two numbers are the low and high bits of the process ID. When you run a process on your current node, its node ID will always be zero. However, when you export a PID to another node, the node ID is set to the number of the node on which the process lives.

That works well once a system is up and running and everything is knitted together. If you want to register a callback process on one node and an event-generating process on another, just give the callback PID to the generator.

But how can the callback find the generator in the first place? One way is for the generator to register its PID, giving it a name. The callback on the other node can look up the generator by name, using the PID that comes back to send messages to it.

Here's an example. Let's write a simple server that sends a notification about every 2 seconds. To receive the notification, a client has to register with the server. And we'll arrange things so that clients on different nodes can register.

While we're at it, we'll do a little packaging so that to start the server you run `Ticker.start`, and to start the client you run `Client.start`. We'll also add an API `Ticker.register` to register a client with the server.

Here's the server code:

```
nodes/ticker.ex
defmodule Ticker do
  @interval 2000  # 2 seconds
  @name      :ticker

  def start do
    pid = spawn(__MODULE__, :generator, [[]])
    :global.register_name(@name, pid)
  end
end
```

```

def register(client_pid) do
  send :global.whereis_name(@name), { :register, client_pid }
end

def generator(clients) do
  receive do
    { :register, pid } ->
      IO.puts "registering #{inspect pid}"
      generator([pid|clients])
  after
    @interval ->
      IO.puts "tick"
      Enum.each clients, fn client ->
        send client, { :tick }
  end
  generator(clients)
end
end
end

```

We define a start function that spawns the server process. It then uses :global.register_name to register the PID of this server under the name :ticker.

Clients who want to register to receive ticks call the register function. This function sends a message to the Ticker server, asking it to add those clients to its list. Clients could have done this directly by sending the :register message to the server process. Instead, we give them an interface function that hides the registration details. This helps decouple the client from the server and gives us more flexibility to change things in the future.

Before we look at the actual tick process, let's stop to consider the start and register functions. These are not part of the tick process—they are simply chunks of code in the Ticker module. This means they can be called directly wherever we have the module loaded—no message passing required. This is a common pattern; we have a module that is responsible both for spawning a process and for providing the external interface to that process.

Back to the code. The last function, generator, is the spawned process. It waits for two events. When it gets a tuple containing :register and a PID, it adds the PID to the list of clients and recurses. Alternatively, it may time out after 2 seconds, in which case it sends a {:tick} message to all registered clients.

(This code has no error handling and no means of terminating the process. I just wanted to illustrate passing PIDs and messages between nodes.) The client code is simple:

```

nodes/ticker.ex
defmodule Client do

```

```

def start do
  pid = spawn(__MODULE__, :receiver, [])
  Ticker.register(pid)
end

def receiver do
  receive do
    { :tick } ->
      IO.puts "tock in client"
      receiver()
  end
end
end

```

It spawns a receiver to handle the incoming ticks, and passes the receiver's PID to the server as an argument to the register function. Again, it's worth noting that this function call is local—it runs on the same node as the client. However, inside the Ticker.register function, it locates the node containing the server and sends it a message. As our client's PID is sent to the server, it becomes an external PID, pointing back to the client's node.

The spawned client process simply loops, writing a cheery message to the console whenever it receives a tick message.

Let's run it. We'll start up our two nodes. We'll call Ticker.start on node one. Then we'll call Client.start on both node one and node two.

Window #1

```

nodes % iex --sname one
iex(one@light-boy)> c("ticker.ex")
[Client,Ticker]
iex(one@light-boy)> Node.connect :"two@light-boy"
true
iex(one@light-boy)> Ticker.start
:yes
tick
tick
iex(one@light-boy)> Client.start
registering #PID<0.59.0>
{:register,#PID<0.59.0>}
tick
tock in client
tick
tock in client
tick
tock in client
tick
tock in client
:
:
```

Window #2

```
nodes % iex --sname two
iex(two@light-boy)> c("ticker.ex")
[Client,Ticker]
iex(two@light-boy)> Client.start
{:register,#PID<0.53.0>}
tock in client
tock in client
tock in client
:
:
```

To stop this, you'll need to exit IEx on both nodes.

When to Name Processes

When you name something, you are recording some global state. And as we all know, global state can be troublesome. What if two processes try to register the same name, for example?

The runtime has some tricks to help us. In particular, we can list the names our application will register in the app's mix.exs file. (We'll cover how when we look at [packaging an application on page 278](#).) However, the general rule is to register your process names when your application starts.

Your Turn

► [Exercise: Nodes-2](#)

When I introduced the interval server, I said it sent a tick “about every 2 seconds.” But in the receive loop, it has an explicit timeout of 2,000 ms. Why did I say “about” when it looks as if the time should be pretty accurate?

► [Exercise: Nodes-3](#)

Alter the code so that successive ticks are sent to each registered client (so the first goes to the first client, the second to the next client, and so on). Once the last client receives a tick, the process starts back at the first. The solution should deal with new clients being added at any time.

Input, Output, PIDs, and Nodes

Input and output in the Erlang VM are performed using I/O servers. These are simply Erlang processes that implement a low-level message interface. You never have to deal with this interface directly (which is a good thing, as it is complex). Instead, you use the various Elixir and Erlang I/O libraries and let them do the heavy lifting.

In Elixir you identify an open file or device by the PID of its I/O server. And these PIDs behave like all other PIDs—you can, for example, send them between nodes. If you look at the implementation of Elixir’s `IO.puts` function, you’ll see

```
def puts(device \\ group_leader(), item) do
  erl_dev = map_dev(device)
  :io.put_chars erl_dev, [to_iodata(item), ?\n]
end
```

(To see the source of an Elixir library module, view the online documentation at <http://elixir-lang.org/docs/>, navigate to the function in question, and click the *Source* link.)

The default device it uses is returned by the function `:erlang.group_leader`. (The `group_leader` function is imported from the `:erlang` module at the top of the `IO` module.) This will be the PID of an I/O server.

So, bring up two terminal windows and start a different named node in each. Connect to node one from node two, and register the PID returned by `group_leader` under a global name (we use `:two`).

	Window #1
	<pre>\$ iex --sname one iex(one@light-boy) ></pre>
	Window #2
	<pre>\$ iex --sname two iex(two@light-boy) > Node.connect(:one@light-boy) true iex(two@light-boy) > :global.register_name(:two, :erlang.group_leader) :yes</pre>

Note that once we’ve registered the PID, we can access it from the other node. And once we’ve done that, we can pass it to `IO.puts`; the output appears in the other terminal window.

	Window #1
	<pre>iex(one@light-boy) > two = :global.whereis_name :two #PID<7419.30.0> iex(one@light-boy) > IO.puts(two, "Hello") :ok iex(one@light-boy) > IO.puts(two, "World!") :ok</pre>
	Window #2
	<pre>Hello World iex(two@light-boy) ></pre>

Your Turn

► *Exercise: Nodes-4*

The ticker process in this chapter is a central server that sends events to registered clients. Reimplement this as a ring of clients. A client sends a tick to the next client in the ring. After 2 seconds, *that* client sends a tick to *its* next client.

When thinking about how to add clients to the ring, remember to deal with the case where a client's receive loop times out just as you're adding a new process. What does this say about who has to be responsible for updating the links?

Nodes Are the Basis of Distribution

We've seen how we can create and interlink a number of Erlang virtual machines, potentially communicating across a network. This is important, both to allow your application to scale and to increase reliability. Running all your code on one machine is like having all your eggs in one basket. Unless you're writing a mobile omelet app, this is probably not a good idea.

It's easy to write concurrent applications with Elixir. But writing code that follows the happy path is a lot easier than writing bullet-proof, scalable, and hot-swappable world-beating apps. For that, you're going to need some help.

In the worlds of Elixir and Erlang, that help is called OTP, and it is the subject of the next few chapters.

OTP: Servers

If you've been following Elixir or Erlang, you've probably come across OTP. It is often hyped as the answer to all high-availability distributed-application woes. It isn't, but it certainly solves many problems that you'd otherwise need to solve yourself, including application discovery, failure detection and management, hot code swapping, and server structure.

First, the obligatory one-paragraph history. OTP stands for the *Open Telecom Platform*, but the full name is largely of historical interest and everyone just says *OTP*. It was initially used to build telephone exchanges and switches. But these devices have the same characteristics we want from any large online application, so OTP is now a general-purpose tool for developing and managing large systems.

OTP is actually a bundle that includes Erlang, a database (wonderfully called *Mnesia*), and an innumerable number of libraries. It also defines a structure for your applications. But, as with all large, complex frameworks, there is a lot to learn. In this book we'll focus on the essentials and I'll point you toward other information sources.

We've been using OTP all along—mix, the Elixir compiler, and even our issue tracker followed OTP conventions. But that use was implicit. Now we'll make it explicit and start writing servers using OTP.

Some OTP Definitions

OTP defines systems in terms of hierarchies of *applications*. An application consists of one or more processes. These processes follow one of a small number of OTP conventions, called *behaviors*. There is a behavior used for general-purpose servers, one for implementing event handlers, and one for finite-state machines. Each implementation of one of these behaviors will run

in its own process (and may have additional associated processes). In this chapter we'll be implementing the *server* behavior, called *GenServer*.

A special behavior, called *supervisor*, monitors the health of these processes and implements strategies for restarting them if needed.

We'll take a look at these components from the bottom up—this chapter will cover servers, the next will explore supervisors, and finally we'll implement applications.

An OTP Server

When we wrote our Fibonacci server in the [previous chapter, on page 211](#), we had to do all the message handling ourselves. It wasn't difficult, but it was tedious. Our scheduler also had to keep track of three pieces of state information: the queue of numbers to process, the results generated so far, and the list of active PIDs.

Most servers have a similar set of needs, so OTP provides libraries that do all the low-level work for us.

When we write an OTP server, we write a module containing one or more callback functions with standard names. OTP will invoke the appropriate callback to handle a particular situation. For example, when someone sends a request to our server, OTP will call our `handle_call` function, passing in the request, the caller, and the current server state. Our function responds by returning a tuple containing an action to take, the return value for the request, and an updated state.

State and the Single Server

Think back to our recursive Fibonacci code. Where did it keep all the intermediate results as it worked? It passed them to itself, recursively, as parameters. In fact, all three of its parameters were used for state information.

Now think about servers. They use recursion to loop, handling one request on each call. So they can also pass state to themselves as a parameter in this recursive call. And that's one of the things OTP manages for us. Our handler functions get passed the current state (as their last parameter), and they return (among other things) a potentially updated state. Whatever state a function returns is the state that will be passed to the next request handler.

Our First OTP Server

Let's write what is possibly the simplest OTP server. You pass it a number when you start it up, and that becomes the current state of the server. When

you call it with a `:next_number` request, it returns that current state to the caller, and at the same time increments the state, ready for the next call. Basically, each time you call it you get an updated sequence number.

Create a New Project Using Mix

Start by creating a new mix project in your work directory. We'll call it `sequence`.

```
$ mix new sequence
* creating README.md
* creating .formatter.exs
* creating .gitignore
* creating mix.exs
* creating config
* creating config/config.exs
* creating lib
* creating lib/sequence.ex
* creating test
* creating test/test_helper.exs
* creating test/sequence_test.exs
```

Create the Basic Sequence Server

Now we'll create `Sequence.Server`, our server module. Move into the `sequence` directory, and create a subdirectory under `lib/` also called `sequence`.

```
$ cd sequence
$ mkdir lib/sequence
```

Add the file `server.ex` to `lib/sequence/`:

```
otp-server/1/sequence/lib/sequence/server.ex
Line 1 defmodule Sequence.Server do
-   use GenServer
-
-   def init(initial_number) do
5     { :ok, initial_number }
-   end
-
-   def handle_call(:next_number, _from, current_number) do
-     { :reply, current_number, current_number + 1 }
10  end
- end
```

The first thing to note is line 2. The `use` line effectively adds the OTP `GenServer` behavior to our module. This is what lets it handle all the callbacks. It also means we don't have to define every callback in our module—the behavior defines defaults for all but one of them.

The exception is the `init/1` function, defined on line 4. You can think of `init` as being like the constructor in an object-oriented language: A constructor takes

values and creates the object's initial state, and init takes some initial value and uses it to construct the state of the server. This state is returned as the second element of the `{:ok, state}` tuple. In our case, we use the init function to set the initial value of our counter.

When a client calls our server, GenServer invokes its `handle_call` function. This function receives three parameters:

1. The information the client passed to the call
2. The PID of the client
3. The server state

Your implementation of the function should perform the actions associated with the first parameter, and may update the state (the third parameter). When the `handle_call` function exits, it must return the state (updated or not).

The initial state of a GenServer is set by the return value of the `init` function.

Our implementation is simple: we return a tuple to OTP.

```
{ :reply, current_number, current_number+1 }
```

The `reply` element tells OTP to reply to the client, passing back the value that is the second element. Finally, the tuple's third element defines the new state. This will be passed as the last parameter to `handle_call` the next time it is invoked.

Fire Up Our Server Manually

We can play with our server in IEx. Open it in the project's main directory, remembering the `-S mix` option.

```
$ iex -S mix
iex> { :ok, pid } = GenServer.start_link(Sequence.Server, 100)
{:ok,#PID<0.71.0>}
iex> GenServer.call(pid, :next_number)
100
iex> GenServer.call(pid, :next_number)
101
iex> GenServer.call(pid, :next_number)
102
```

We're using two functions from the Elixir GenServer module. The `start_link` function behaves like the `spawn_link` function we used in the previous chapter. It asks GenServer to start a new process and link to us (so we'll get notifications if it fails). We pass in the module to run as a server: the initial state (100 in this

case). We could also pass GenServer options as a third parameter, but the defaults work fine here.

We get back a status (`:ok`) and the server's PID. The `call` function takes this PID and calls the `handle_call` function in the server. The `call`'s second parameter is passed as the first argument to `handle_call`.

In our case, the only value we need to pass is the identity of the action we want to perform, `:next_number`. If you look at the definition of `handle_call` in the server, you'll see that its first parameter is `:next_number`. When Elixir invokes the function, it pattern-matches the argument in the call with this first parameter in the function. A server can support multiple actions by implementing multiple `handle_call` functions with different first parameters.

If you want to pass more than one thing in the call to a server, pass a tuple. For example, our server might need a function to reset the count to a given value. We could define the handler as

```
def handle_call({:set_number, new_number}, _from, _current_number) do
  { :reply, new_number, new_number }
end
```

and call it with

```
iex> GenServer.call(pid, {:set_number, 999})
999
```

Similarly, a handler can return multiple values by packaging them into a tuple or list.

```
def handle_call({:factors, number}, _, _) do
  { :reply, { :factors_of, number, factors(number)}, [] }
end
```

Your Turn

► *Exercise: OTP-Servers-1*

You're going to start creating a server that implements a stack. The call that initializes your stack will pass in a list of the initial stack contents.

For now, implement only the `pop` interface. It's acceptable for your server to crash if someone tries to pop from an empty stack.

For example, if initialized with `[5, "cat", 9]`, successive calls to `pop` will return `5`, `"cat"`, and `9`.

One-Way Calls

The `call` function calls a server and waits for a reply. But sometimes you won't want to wait because there is no reply coming back. In those circumstances, use the `GenServer cast` function. (Think of it as casting your request into the sea of servers.)

Just like `call` is passed to `handle_call` in the server, `cast` is sent to `handle_cast`. Because there's no response possible, the `handle_cast` function takes only two parameters: the call argument and the current state. And because it doesn't want to send a reply, it will return the tuple `{:noreply, new_state}`.

Let's modify our sequence server to support an `:increment_number` function. We'll treat this as a cast, so it simply sets the new state and returns.

```
otp-server/1/sequence/lib/sequence/server.ex
defmodule Sequence.Server do
  use GenServer

  def init(initial_number) do
    { :ok, initial_number }
  end

  def handle_call(:next_number, _from, current_number) do
    {:reply, current_number, current_number + 1}
  end

  ▶  def handle_cast({:increment_number, delta}, current_number) do
  ▶    { :noreply, current_number + delta}
  ▶  end
end
```

Notice that the cast handler takes a tuple as its first parameter. The first element is `:increment_number`, and is used by pattern matching to select the handlers to run. The second element of the tuple is the delta to add to our state. The function simply returns a tuple, where the state is the previous state plus this number.

To call this from our IEx session, we first have to recompile our source. The `r` command takes a module name and recompiles the file containing that module.

```
iex> r Sequence.Server
.../sequence/lib/sequence/server.ex:2: redefining module Sequence.Server
{Sequence.Server, [Sequence.Server]}
```

Even though we've recompiled the code, the old version is still running. The VM doesn't hot-swap code until you explicitly access it by module name. So, to try our new functionality we'll create a new server. When it starts, it will pick up the latest version of the code.

```
iex> {:ok, pid} = GenServer.start_link(Sequence.Server, 100)
{:ok,#PID<0.60.0>}
iex> GenServer.call(pid, :next_number)
100
iex> GenServer.call(pid, :next_number)
101
iex> GenServer.cast(pid, {:increment_number, 200})
:ok
iex> GenServer.call(pid, :next_number)
302
```

Tracing a Server's Execution

The third parameter to start_link is a set of options. A useful one during development is the debug trace, which logs message activity to the console.

We enable tracing using the debug option:

```
➤ iex> {:ok,pid} = GenServer.start_link(Sequence.Server, 100, [debug: [:trace]])
{:ok,#PID<0.68.0>}
iex> GenServer.call(pid, :next_number)
*DBG* <0.68.0> got call next_number from <0.25.0>
*DBG* <0.68.0> sent 100 to <0.25.0>, new state 101
100
iex> GenServer.call(pid, :next_number)
*DBG* <0.68.0> got call next_number from <0.25.0>
*DBG* <0.68.0> sent 101 to <0.25.0>, new state 102
101
```

See how it traces the incoming call and the response we send back. A nice touch is that it also shows the next state.

We can also include :statistics in the debug list to ask a server to keep some basic statistics:

```
➤ iex> {:ok,pid} = GenServer.start_link(Sequence.Server, 100, [debug: [:statistics]])
{:ok,#PID<0.69.0>}
iex> GenServer.call(pid, :next_number)
100
iex> GenServer.call(pid, :next_number)
101
iex> :sys.statistics pid, :get
{:ok,
 [
  start_time: {{2017, 12, 23}, {14, 6, 7}},
  current_time: {{2017, 12, 23}, {14, 6, 24}},
  reductions: 36,
  messages_in: 2,
  messages_out: 0
 ]}
```

Most of the fields should be fairly obvious. Timestamps are given as `{y,m,d},{h,m,s}` tuples. The reductions value is a measure of the amount of work the server does. It is used in process scheduling as a way of making sure all processes get a fair share of the available CPU.

The Erlang `sys` module is your interface to the world of *system messages*. These are sent in the background between processes—they’re a bit like the backchatter in a multiplayer video game. While two players are engaged in an attack (their real work), they can also be sending each other background messages: “Where are you?,” “Stop moving,” and so on.

The list associated with the `debug` parameter you give to `GenServer` is simply the names of functions to call in the `sys` module. If you say `[debug: [:trace, :statistics]]`, then those functions will be called in `sys`, passing in the server’s PID. Look at the documentation for `sys` to see what’s available.¹

This also means you can turn things on and off *after* you have started a server. For example, you can enable tracing on an existing server using the following:

```
iex> :sys.trace pid, true
:ok
iex> GenServer.call(pid, :next_number)
*DBG* <0.69.0> got call next_number from <0.25.0>
*DBG* <0.69.0> sent 105 to <0.25.0>, new state 106
105
iex> :sys.trace pid, false
:ok
iex> GenServer.call(pid, :next_number)
106
```

`get_status` is another useful `sys` function:

```
iex> :sys.get_status pid
{:status, #PID<0.134.0>, {:module, :gen_server},
 [
  [
   "$initial_call": {Sequence.Server, :init, 1},
   "$ancestors": [#PID<0.118.0>, #PID<0.57.0>]
  ],
  :running,
  #PID<0.118.0>,
  [statistics: {{2017, 12, 23}, {14, 11, 13}}, {:reductions, 14}, 3, 0],
  [
   header: 'Status for generic server <0.134.0>',
   data: [
```

1. <http://www.erlang.org/documentation/doc-5.8.3/lib/stdlib-1.17.3/doc/html/sys.html>

```

        {'Status', :running},
        {'Parent', #PID<0.118.0>},
        {'Logged events', []}
    ],
    data: [{('State', 103)}]
}

```

This is the default formatting of the status message GenServer provides. You have the option to change the data: part to a more application-specific message by defining a format_status function. This receives an option describing why the function was called, as well as a list containing the server's process dictionary and the current state. (Note that in the code that follows, the string *State* in the response is in single quotes.)

```
otp-server/1/sequence/lib/sequence/server.ex
def format_status(_reason, [ _pdict, state ]) do
  [data: [{('State', "My current state is '#{inspect state}', and I'm happy")}]]
end
```

If we ask for the status in IEx, we get the new message (after restarting the server):

```
iex> :sys.get_status pid
{:status, #PID<0.124.0>, {:module, :gen_server},
 [
  [
   "$initial_call": {Sequence.Server, :init, 1},
   "$ancestors": [#PID<0.118.0>, #PID<0.57.0>]
  ],
  :running,
  #PID<0.118.0>,
  [statistics: {{2017, 12, 23}, {14, 6, 7}}, {:reductions, 14}, 2, 0],
  [
   header: 'Status for generic server <0.124.0>',
   data: [
     {'Status', :running},
     {'Parent', #PID<0.118.0>},
     {'Logged events', []}
   ],
   data: [{('State', "My current state is '102', and I'm happy")}]
  ]
 ]}
```

Your Turn

► *Exercise: OTP-Servers-2*

Extend your stack server with a push interface that adds a single value to the top of the stack. This will be implemented as a cast.

Experiment in IEx with pushing and popping values.

GenServer Callbacks

GenServer is an OTP protocol. OTP works by assuming that your module defines a number of callback functions (six, in the case of a GenServer). If you were writing a GenServer in Erlang, your code would have to contain implementations of all six.

When you add the line ‘use GenServer’ to a module, Elixir creates default implementations of these six callback functions. All we have to do is override the ones where we add our own application-specific behavior. Our examples so far have used the three callbacks ‘init’, ‘handle_call’, and ‘handle_cast’. Here’s a full list:

`init(start_arguments)`

Called by GenServer when starting a new server. The parameter is the second argument passed to `start_link`. It should return `{:ok, state}` on success, or `{:stop, reason}` if the server could not be started.

You can specify an optional timeout using `{:ok, state, timeout}`, in which case GenServer sends the process a `:timeout` message whenever no message is received in a span of `timeout` ms. (The message is passed to the `handle_info` function.)

The default GenServer implementation sets the server state to the argument you pass.

`handle_call(request, from, state)`

Invoked when a client uses `GenServer.call(pid, request)`. The `from` parameter is a tuple containing the PID of the client and a unique tag. The `state` parameter is the server state.

On success it returns `{:reply, result, new_state}`. The [list that follows this one, on page 239](#), shows other valid responses.

The default implementation stops the server with a `:bad_call` error, so you’ll need to implement `handle_call` for every call request type your server implements.

`handle_cast(request, state)`

Called in response to `GenServer.cast(pid, request)`.

A successful response is `{:noreply, new_state}`. It can also return `{:stop, reason, new_state}`.

The default implementation stops the server with a `:bad_cast` error.

`handle_info(info, state)`

Called to handle incoming messages that are not call or cast requests. For example, timeout messages are handled here. So are termination messages from any linked processes. In addition, messages sent to the PID using send (so they bypass GenServer) will be routed to this function.

`terminate(reason, state)`

Called when the server is about to be terminated. However, as we'll discuss in the next chapter, once we add supervision to our servers, we don't have to worry about this.

`code_change(from_version, state, extra)`

Updates a running server without stopping the system. However, the new version of the server may represent its state differently from the old version. The `code_change` callback is invoked to change from the old state format to the new.

`format_status(reason, [pdict, state])`

Used to customize the state display of the server. The conventional response is `[data: [{ 'State', state_info }]]`.

The call and cast handlers return standardized responses. Some of these responses can contain an optional `:hibernate` or `timeout` parameter. If `:hibernate` is returned, the server state is removed from memory but is recovered on the next request. This saves memory at the expense of some CPU. The `timeout` option can be the atom `:infinite` (which is the default) or a number. If the latter, a `:timeout` message is sent if the server is idle for the specified number of milliseconds.

The first two responses are common between call and cast.

```
{ :noreply, new_state [ , :hibernate | timeout ] }
```

```
{ :stop, reason, new_state }
```

Signal that the server is to terminate.

Only `handle_call` can use the last two.

```
{ :reply, response, new_state [ , :hibernate | timeout ] }
```

Send response to the client.

```
{ :stop, reason, reply, new_state }
```

Send the response and signal that the server is to terminate.

Naming a Process

The idea of referencing processes by their PIDs gets old quickly. Fortunately, there are a number of alternatives.

The simplest is local naming. We assign a name that is unique for all OTP processes on our node, and then we use that name instead of the PID whenever we reference it. To create a locally named process, we use the name: option when we start the server:

```
> iex> { :ok, pid } = GenServer.start_link(Sequence.Server, 100, name: :seq)
{:ok,#PID<0.58.0>}
iex> GenServer.call(:seq, :next_number)
100
iex> GenServer.call(:seq, :next_number)
101
iex> :sys.get_status :seq
{:status, #PID<0.69.0>, {:module, :gen_server},
 [{"$ancestors": [#PID<0.58.0>],
   "$initial_call": {Sequence.Server, :init, 1}],
   :running, #PID<0.58.0>, [],
   [header: 'Status for generic server seq',
    data: [{"Status", :running},
           {'Parent', #PID<0.58.0>},
           {'Logged events', []}],
    data: [{"State", "My current state is '102', and I'm happy"}]}]}
```

Tidying Up the Interface

As we left it, our server works but is ugly to use. Our callers have to make explicit GenServer calls, and they have to know the registered name for our server process. We can do better. Let's wrap this interface in a set of three functions in our server module: start_link, next_number, and increment_number. The first of these calls the GenServer start_link method. As we'll see in the next chapter, the name start_link is a convention. start_link must return the correct status values to OTP; as our code simply delegates to the GenServer module, this is taken care of.

Following the definition of start_link, the next two functions are the external API to issue call and cast requests to the running server process.

We'll also use the name of the module as our server's registered local name (hence the name: __MODULE__ when we start it, and the __MODULE__ parameter when we use call or cast).

```

otp-server/2/sequence/lib/sequence/server.ex
defmodule Sequence.Server do
  use GenServer

#####
# External API

➤ def start_link(current_number) do
  GenServer.start_link(__MODULE__, current_number, name: __MODULE__)
end

➤ def next_number do
  GenServer.call __MODULE__, :next_number
end

➤ def increment_number(delta) do
  GenServer.cast __MODULE__, {:increment_number, delta}
end

#####
# GenServer implementation

def init(initial_number) do
  { :ok, initial_number }
end

def handle_call(:next_number, _from, current_number) do
  { :reply, current_number, current_number+1 }
end

def handle_cast({:increment_number, delta}, current_number) do
  { :noreply, current_number + delta }
end

def format_status(_reason, [ _pdict, state ]) do
  [data: [ {'State', "My current state is '#{:inspect state}'", and I'm happy"} ] ]
end
end

```

When we run this code in IEx, it's a lot cleaner:

```

$ iex -S mix
iex> Sequence.Server.start_link 123
{:ok,#PID<0.57.0>}
iex> Sequence.Server.next_number
123
iex> Sequence.Server.next_number
124
iex> Sequence.Server.increment_number 100
:ok
iex> Sequence.Server.next_number
225

```

This is the pattern that just about everyone uses in the Elixir world.

I have a different view. However, it certainly isn't mainstream, so feel free to skip the next section if you want to avoid becoming tainted by my heresies.

Making Our Server into a Component

Earlier I said that what Elixir calls an application, most people would call a component or a service. That's certainly what our sequence server is: a free-standing chunk of code that enjoyed generating successive numbers.

Despite being the canonical way of writing this, I don't like my implementation. It puts three things into a single source file:

- The API
- The logic of our service (adding one)
- The implementation of that logic in a server

Have another look at the code [on page 241](#).

If you didn't know what it did, how would you find out? Where's the code that does the component's logic? (The image gives you a hint.) It isn't obvious, and this is just a trivial service. Imagine working with a really complex one, with lots of logic.

```

defmodule Sequence.Server do
  use GenServer
  ...
  # External API
  def start_link(current_number) do
    GenServer.start_link __MODULE__, current_number, name: __MODULE__
  end
  ...
  def next_number do
    GenServer.call __MODULE__, :next_number
  end
  ...
  def increment_number(delta) do
    GenServer.cast __MODULE__, {:increment_number, delta}
  end
  ...
  # GenServer implementation
  def init(initial_number) do
    {:ok, initial_number}
  end
  ...
  def handle_call(:next_number, _from, current_number) do
    {:reply, current_number + 1, current_number}
  end
  ...
  def handle_cast({:increment_number, delta}, current_number) do
    {:noreply, current_number + delta}
  end
  ...
  def format_status(reason, [{:dict, state}]) do
    [data: ["State", "My current state is \"#{inspect state}\", and I'm happy"]]
  end
  ...
end

```

That's why I'm experimenting with splitting the API, implementation, and server into three separate files.

We'll start afresh:

```

$ mix new sequence
$ cd sequence
$ mkdir lib/sequence
$ touch lib/sequence/impl.ex lib/sequence/server.ex
$ tree
├── README.md
├── config
│   └── config.exs
└── lib
    ├── sequence
    │   ├── impl.ex
    │   └── server.ex
    └── sequence.ex
├── mix.exs
└── test
    ├── sequence_test.exs
    └── test_helper.exs

```

We'll put the API in the top-level `lib/sequence.ex` module, and the implementation and server in the two lower-level modules.

The API is the public face of our component. It is simply the top half of the previous server module:

```
otp-server/3/sequence/lib/sequence.ex
defmodule Sequence do
  @server Sequence.Server
  def start_link(current_number) do
    GenServer.start_link(@server, current_number, name: @server)
  end
  def next_number do
    GenServer.call(@server, :next_number)
  end
  def increment_number(delta) do
    GenServer.cast(@server, {:increment_number, delta})
  end
end
```

This forwards calls on to the server implementation:

```
otp-server/3/sequence/lib/sequence/server.ex
defmodule Sequence.Server do
  use GenServer
  alias Sequence.Impl
  def init(initial_number) do
    { :ok, initial_number }
  end
  def handle_call(:next_number, _from, current_number) do
    { :reply, current_number, Impl.next(current_number) }
  end
  def handle_cast({:increment_number, delta}, current_number) do
    { :noreply, Impl.increment(current_number, delta) }
  end
  def format_status(_reason, [ _pdict, state ]) do
    [ data: [ 'State', "My current state is '#{:inspect state}', and I'm happy" ] ]
  end
end
```

Unlike the previous server, this code contains no “business logic” (which in our case is adding either 1 or some delta to our state). Instead, it uses the implementation module to do this:

```
otp-server/3/sequence/lib/sequence/impl.ex
defmodule Sequence.Impl do
  def next(number),           do: number + 1
  def increment(number, delta), do: number + delta
end
```

Now, you're probably looking at this and thinking, "all that work just to implement a counter?" And you'd be right. But this chapter isn't about implementing a counter. It's all about implementing real-world servers. Because Elixir makes it easy to bundle all the code for a server into one module, most people do, and then they end up with some fairly highly coupled (and hard-to-test) code.

So think of this example, but with some real, complex business logic. Imagine how you might write it.

You'd probably start with an API, just to see what it would look like. Then you might want to write and test some of the business logic. So go into the implementation module and do just that. What's more, test that code directly as you write it: no need to run it inside a server for that.

As you progress with the implementation, you may learn things that require changes to the overall API. Feel free.

Then, when you're feeling good about the code, add a server module and have the API use it.

The thing I like about this approach is that it leaves me a pure implementation of the actual logic on my component, independent of whether I choose to deploy it as a server. I can use (and test) the logic either as direct function calls or indirectly via a server.

Your Turn

► *Exercise: OTP-Servers-3*

Give your stack server process a name, and make sure it is accessible by that name in IEx.

► *Exercise: OTP-Servers-4*

Add the API to your stack module (the functions that wrap the GenServer calls).

► *Exercise: OTP-Servers-5*

Implement the terminate callback in your stack handler. Use IO.puts to report the arguments it receives.

Try various ways of terminating your server. For example, popping an empty stack will raise an exception. You might add code that calls `System.halt(n)` if the push handler receives a number less than 10. (This will let you generate different return codes.) Use your imagination to try different scenarios.

An OTP GenServer is just a regular Elixir process in which the message handling has been abstracted out. The GenServer behavior defines a message loop internally and maintains a state variable. That message loop then calls out to various functions that we define in our server module: `handle_call`, `handle_cast`, and so on.

We also saw that GenServer provides fairly detailed tracing of the messages received and responses sent by our server modules.

Finally, we wrapped our message-based API in module functions, which gives our users a cleaner interface and decouples them from our implementation.

But we still have an issue if our server crashes. We'll deal with this in the next chapter, when we look at supervisors.

OTP: Supervisors

I've said it a few times now: the *Elixir way* says not to worry much about code that crashes; instead, make sure the overall application keeps running.

This might sound contradictory, but really it is not.

Think of a typical application. If an unhandled error causes an exception to be raised, the application stops. Nothing else gets done until it is restarted. If it's a server handling multiple requests, they all might be lost.

The issue here is that one error takes the whole application down.

But imagine that instead your application consists of hundreds or thousands of processes, each handling just a small part of a request. If one of those crashes, everything else carries on. You might lose the work it's doing, but you can design your applications to minimize even that risk. And when that process gets restarted, you're back running at 100%.

In the Elixir and OTP worlds, *supervisors* perform all of this process monitoring and restarting.

Supervisors and Workers

An Elixir supervisor has just one purpose—it manages one or more processes. (As we'll discuss later, these processes can be workers or other supervisors.)

At its simplest, a supervisor is a process that uses the OTP supervisor behavior. It is given a list of processes to monitor and is told what to do if a process dies, and how to prevent restart loops (when a process is restarted, dies, gets restarted, dies, and so on).

To do this, the supervisor uses the Erlang VM's process-linking and -monitoring facilities. We talked about these when we covered [spawn on page 207](#).

You can write supervisors as separate modules, but the Elixir style is to include them inline. The easiest way to get started is to create your project with the `--sup` flag. Let's do this for our sequence server.

```
➤ $ mix new --sup sequence
* creating README.md
* creating .formatter.exs
* creating .gitignore
* creating mix.exs
* creating config
* creating config/config.exs
* creating lib
* creating lib/sequence.ex
* creating lib/sequence/application.ex
* creating test
* creating test/test_helper.exs
* creating test/sequence_test.exs
```

The only apparent difference is the appearance of the file `lib/sequence/application.ex`. Let's have a look inside (I stripped out some comments...):

```
defmodule Sequence.Application do
  @moduledoc false

  use Application

  def start(_type, _args) do
    children = [
      # {Sequence.Worker, arg},
    ]

    opts = [strategy: :one_for_one, name: Sequence.Supervisor]
    Supervisor.start_link(children, opts)
  end
end
```

Our `start` function now creates a supervisor for our application. All we need to do is tell it what we want supervised. Copy the second version of the `Sequence.Server` module¹ from the last chapter into the `lib/sequence` folder. Then uncomment and change the line in the `child_list` to reference this module:

```
otp-supervisor/1/sequence/lib/sequence/application.ex
def start(_type, _args) do
  children = [
    {Sequence.Server, 123},
  ]

  opts = [strategy: :one_for_one, name: Sequence.Supervisor]
  Supervisor.start_link(children, opts)
end
```

1. <http://media.pragprog.com/titles/elixir16/code/otp-server/2/sequence/lib/sequence/server.ex>

Let's look at what's going to happen:

- When our application starts, the start function is called.
- It creates a list of child server modules. In our case, there's just one, the Sequence.Server. Along with the module name, we specify an argument to be passed to the server when we start it.
- We call Supervisor.start_link, passing it the list of child specifications and a set of options. This creates a supervisor process.
- Now our supervisor process calls the start_link function for each of its managed children. In our case, this is the function in Sequence.Server. This code is unchanged—it calls GenServer.start_link to create a GenServer process.

Now we're up and running. Let's try it:

```
$ iex -S mix
Compiling 2 files (.ex)
Generated sequence app
iex> Sequence.Server.increment_number 3
:ok
iex> Sequence.Server.next_number
126
```

So far, so good. But the key thing with a supervisor is that it is supposed to manage our worker process. If it dies, for example, we want it to be restarted. Let's try that. If we pass something that isn't a number to increment_number, the process should die trying to add it to the current number.

```
iex(3)> Sequence.Server.increment_number "cat"
:ok
iex(4)> 14:22:06.269 [error] GenServer Sequence.Server terminating
Last message: {:gen_cast, {increment_number, "cat"}}
State: [data: [{State, "My current state is '127', and I'm happy"}]]
** (exit) an exception was raised:
  ** (ArithmetricError) bad argument in arithmetic expression
      (sequence) lib/sequence/server.ex:27: Sequence.Server.handle_cast/2
      (stdlib) gen_server.erl:599: :gen_server.handle_msg/5
      (stdlib) proc_lib.erl:239: :proc_lib.init_p_do_apply/3
iex(4)> Sequence.Server.next_number
123
iex(5)> Sequence.Server.next_number
124
```

We get a wonderful error report that shows us the exception, along with a stack trace from the process. We can also see the message we sent that triggered the problem.

But when we then ask our server for a number, it responds as if nothing had happened. The supervisor restarted our process for us.

This is excellent, but there's a problem. The supervisor restarted our sequence process with the initial parameters we passed in, and the numbers started again from 123. A reincarnated process has no memory of its past lives, and no state is retained across a crash.

Your Turn

► *Exercise: OTP-Supervisors-1*

Add a supervisor to your stack application. Use IEx to make sure it starts the server correctly. Use the server normally, and then crash it (try popping from an empty stack). Did it restart? What were the stack contents after the restart?

Managing Process State Across Restarts

Some servers are effectively stateless. If we had a server that calculated the factors of numbers or responded to network requests with the current time, we could simply restart it and let it run.

But our server is not stateless—it needs to remember the current number so it can generate an increasing sequence.

All of the approaches to this involve storing the state outside of the process. Let's choose a simple option—we'll write a separate process that can store and retrieve a value. We'll call it our *stash*. The sequence server can store its current number in the stash whenever it terminates, and then we can recover the number when we restart.

Now, we have to think about lifetimes. Our sequence server should be fairly robust, but we've already found one thing that crashes it. In actuarial terms, it isn't the fittest process in the scheduler queue. But our stash process must be more robust—it must outlive the sequence server, at the very least.

We have to do two things to make this happen. First, we make it as simple as possible. The fewer moving parts in a chunk of code, the less likely it is to go wrong.

Second, we have to arrange things so that the supervisor isolates it from failures in the sequence server.

Let's do the first of these things now. We'll create a trivial server whose entire purpose is to store a single value.

```
otp-supervisor/2/sequence/lib/sequence/stash.ex
defmodule Sequence.Stash do
  use GenServer

  @me __MODULE__

  def start_link(initial_number) do
    GenServer.start_link(__MODULE__, initial_number, name: @me)
  end

  def get() do
    GenServer.call(@me, { :get })
  end

  def update(new_number) do
    GenServer.cast(@me, { :update, new_number })
  end

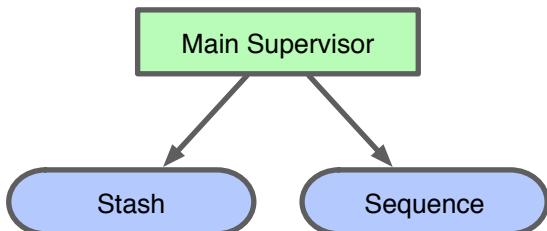
  # Server implementation

  def init(initial_number) do
    { :ok, initial_number }
  end

  def handle_call({ :get }, _from, current_number ) do
    { :reply, current_number, current_number }
  end

  def handle_cast({ :update, new_number }, _current_number) do
    { :noreply, new_number }
  end
end
```

Now we want to supervise it. It'll be running alongside the sequence server:



This is the first time we've had two servers supervised together. So now we have to face a question: what happens if just one of them crashes? The answer depends on the *supervision strategy* we have chosen.

`:one_for_one`

if a server dies, the supervisor will restart it. This is the default strategy.

`:one_for_all`

if a server dies, all the remaining servers are first terminated, and then the servers are all restarted.

:rest_for_one

if a server dies, the servers that follow it in the list of children are terminated, and then the dying server and those that were terminated are restarted.

Right now we only have two servers. The stash server is supposed to be eternal, while the sequence server might crash and need restarting. Because of this, we know that we can't use :one_for_all.

This leaves two choices, and both would work. If we use :one_for_one, a failing sequence server will restart, and the stash will not be affected. If we use :rest_for_one, the same thing will happen, but only if the sequence server follows the stash in the list of children.

Which to choose? I vote for :rest_for_one, not because it has any different behavior to :one_for_one, but because I feel it expresses my intent better. A :rest_for_one supervision strategy explicitly says, “this server depends on the health of previous servers in the list.”

Let's add the stash and update the supervision strategy in our supervisor startup code:

```
otp-supervisor/2/sequence/lib/sequence/application.ex
defmodule Sequence.Application do
  @moduledoc false
  use Application
  def start(_type, _args) do
    >> children = [
    >>   { Sequence.Stash, 123 },
    >>   { Sequence.Server, nil },
    >>
    >>   opts = [strategy: :rest_for_one, name: Sequence.Supervisor]
    >>   Supervisor.start_link(children, opts)
  end
end
```

Finally, we need to change our sequence server to use this stash. For now, we'll have it set its state to the current value in the stash when it starts, and have it store the value back into the stash if it crashes.

Setting the initial state simply means fetching the current value from the stash in our sequence server's init function. Handling the sequence server exiting involves writing another callback, terminate. Here's the full code:

```
otp-supervisor/2/sequence/lib/sequence/server.ex
defmodule Sequence.Server do
  use GenServer

#####
# External API

➤ def start_link(_) do
  GenServer.start_link(__MODULE__, nil, name: __MODULE__)
end

➤ def next_number do
  GenServer.call __MODULE__, :next_number
end

➤ def increment_number(delta) do
  GenServer.cast __MODULE__, {:increment_number, delta}
end

#####
# GenServer implementation

def init(_) do
  { :ok, Sequence.Stash.get() }
end

def handle_call(:next_number, _from, current_number) do
  { :reply, current_number, current_number+1 }
end

def handle_cast({:increment_number, delta}, current_number) do
  { :noreply, current_number + delta }
end

➤ def terminate(_reason, current_number) do
  Sequence.Stash.update(current_number)
end

end
```

Let's try this change in IEx:

```
$ iex -S mix
iex> Sequence.Server.next_number
123
iex> Sequence.Server.next_number
124
iex> Sequence.Server.next_number
125
iex> Sequence.Server.increment_number "cat"
:ok
iex>
12:15:48.424 [error] GenServer Sequence.Server terminating
** (ArithmetricError) bad argument in arithmetic expression
(sequence) lib/sequence/server.ex:39: Sequence.Server.handle_cast/2
```

```
Last message: {:"$gen_cast", {:increment_number, "cat"}}
State: 126
iex> Sequence.Server.next_number
126
iex> Sequence.Server.next_number
127
iex>
```

How cool is that? The server code crashed, but was then restarted automatically. And, in the process, the state was stored away in the stash and then recovered—the sequence continued uninterrupted.

Your Turn

► *Exercise: OTP-Supervisors-2*

Rework your stack server to use a supervision tree with a separate stash process to hold the state. Verify that it works and that when you crash the server state is retained across a restart.

Simplifying the Stash

The sole job of our stash module is to store a value. When we look at agents, we'll see that they're a perfect fit for this, and we'll be able to simplify our code.

Worker Restart Options

So far we've looked at supervision from the point of view of the supervisor. In particular, we've seen how the supervision strategy tells the supervisor how to deal with the death of a child process.

There's a second level of configuration that applies to individual workers. The most commonly used of these is the :restart option.

Previously we said that a supervisor strategy (such as :one_for_all) is invoked when a worker dies. That's not strictly true. Instead, the strategy is invoked when a worker needs restarting. And the conditions when a worker should be restarted are dictated by its restart: option:

:permanent

This worker should always be running—it is permanent. This means that the supervision strategy will be applied whenever this worker terminates, for whatever reason.

:temporary

This worker should never be restarted, so the supervision strategy is never applied if this worker dies.

`:transient`

It is expected that this worker will at some point terminate normally, and this termination should not result in a restart. However, should this worker die abnormally, then it should be restarted by running the supervision strategy.

The simplest way to specify the restart option for a worker is in the worker module. You add it to the use GenServer (or use Supervisor) line:

```
defmodule Convolver do
  use GenServer, restart: :transient
  # ...
```

A Little More Detail

You don't have to know the information in this section first time around, but as someone always asks....

We've seen that you start a supervisor by passing it a list of children. Just what is that list?

At the very lowest level, it is a list of *child specifications*. A child spec is an Elixir map. It describes which function to call to start the worker, how to shut the worker down, the restart strategy, the worker type, and any modules apart from the main module that form part of the worker.

You can create a child spec map using the Supervisor.child_spec/2 function.

At the next level up, you can specify a worker by giving its module name (or a tuple containing the module and the initial arguments). In this case, the supervisor assumes you've implemented a child_spec function in that module, and calls that function to get the specification.

Going up one more level, when you add the line use GenServer to a server module, Elixir will define a default child_spec function in that module. This function by default returns a map that tells the supervisor that the start function will be start_link and that the restart strategy will be :permanent. You can override these defaults with the options you give use GenServer.

In practice, the option you'll change the most will be :restart. Although :permanent is a good default for long-running servers, it won't work for servers that do a job and then exit. These types of servers should have a restart value of :transient.

Supervisors Are the Heart of Reliability

Think about our previous example; it was both trivial and profound. It was trivial because there are many ways of achieving some kind of fault tolerance

with a library that returns successive numbers. But it was profound because it is a concrete representation of the idea of building rings of confidence in our code. The outer ring, where our code interacts with the world, should be as reliable as we can make it. But within that ring there are other, nested rings. And in those rings, things can be less than perfect. The trick is to ensure that the code in each ring knows how to deal with failures of the code in the next ring down.

And that's where supervisors come into play. In this chapter we've seen only a small fraction of supervisors' capabilities. They have different strategies for dealing with the termination of a child, different ways of terminating children, and different ways of restarting them. There's plenty of information online about using OTP supervisors.

But the real power of supervisors is that they exist. The fact that you use them to manage your workers means you are forced to think about reliability and state as you design your application. And that discipline leads to applications with very high availability—in [Programming Erlang \(2nd edition\)](#) [Arm13], Joe Armstrong says OTP has been used to build systems with 99.9999999% reliability. That's nine nines. And that ain't bad.

(In case you were wondering, that equates to a complete application outage of roughly 1 second every 30 years. I don't know how you'd even measure that, which makes me a little suspicious....)

There's one more level in our lightning tour of OTP—the application. But before we look at that, let's use what we learned so far and build some real-world code.

A More Complex Example

When I first used Elixir, I struggled with how to organize my applications. When should I use servers? How do supervisors fit in? Even basic questions such as “how many applications should I write?” made me nervous.

Frankly, I’m still thinking about these questions, three years later, in the same way I’m still thinking about object-oriented system design after 30 years of doing it. But, on the Elixir front, I have come up with an approach that helps me think through these issues.

It isn’t rocket science. I just ask myself these five questions:

- What is the environment and what are its constraints?
- What are the obvious focal points?
- What are the runtime characteristics?
- What do I protect from errors?
- How do I get this thing running?

What I’m going to show in this chapter is just my ad hoc approach. Please don’t take it as any kind of “methodology.” But if, like me, you sometimes feel overwhelmed when designing a new Elixir system, these steps might help.

Let’s write a simple application to show you what I mean.

Introduction to Duper

I have loads of duplicate files littering my computers. In an effort to tame this, let’s write a duplicate-file finder. Well call it *Duper* (so we can later create a paid version called *SuperDuper*). It’ll work by scanning all the files in a directory tree, calculating a hash for each. If two files have the same hash, we’ll report them as duplicates.

Let’s start asking the questions.

Q1: What is the environment and what are its constraints?

We're going to run this on a typical computer. It'll have roughly two orders of magnitude more file storage than main memory. Files will range in size from 10^0 to 10^{10} bytes, and there will be roughly 10^7 of them.

What this means:

We need to allow for the fact that although we have to load files into memory to determine their hashes, it is possible we won't have enough memory to load the largest files in whole. We definitely will not be able to process all the files at once.

It also means that our design will need to cater to both big and small files. Big files will take more time to read into memory than small files, and they will also take longer to hash.

Q2: What are the focal points?

A focal point represents a responsibility of the application. By considering the focal points now, we can start to reduce coupling in the application as a whole: each focal point can be tightly coupled internally but loosely coupled to the others. This coupling can be both structural (for example, the representation of data) and temporal (for example, the sequence in which things will happen).

In Duper, we can fairly easily identify some key focal points:

- We need to have a place where we collect results. We are calculating a hash value for each file, so this results store will need to hold all of these. As we are looking for duplicate hashes, it would make sense for this to be some kind of key/value store internally, where the key is the hash and the value is a list of all files with that hash. However, this is an implementation detail, and the implementation shouldn't leak through our API.
- We need to have something that can traverse the filesystem, returning each path just once.
- We need to have something that can take a path and calculate the hash of the corresponding file. Because an individual file may be too big to fit in memory, we'll have to read it in chunks, calculating the hash incrementally.
- Because we know we will need to be processing multiple files concurrently in order to maximize our use of the CPU and IO bandwidth, we'll need something that orchestrates the overall process.

This list may well change as we start to write code, but it's good enough to get us to the next step.

What this means:

At the very least, each focus we identify is an Elixir module. My experience is that it's wise to assume that most if not all are going to end up being servers. I'm also coming to believe that many should even be separate Elixir applications, but that's not something I'm going to dig into here.

Our code will be structured into four servers. Although we could do it with fewer, using four means we can have specific characteristics for each. The four are as follows:

- The *Results*. This is the most important server, as it holds the results of the scanning in memory. We need it to be reliable, so we won't put much code in it.
- The *PathFinder*. This is responsible for returning the paths to each file in the directory tree, one at a time.
- The *Worker*. This asks the PathFinder for a path, calculates the hash of the resulting file's contents, and passes the result to the gatherer.
- The *Gatherer*. This is the server that both starts the ball rolling and determines when things have completed. When they do, it fetches the results and reports on them.

Q3: What are the runtime characteristics?

Our application is going to spend the vast majority of its time in the workers, as this is where we read the files and calculate the hash values. Our goal is to keep both the processors and the IO bus as busy as possible in order to maximize performance.

If we have just one worker, then it would read a file, hash it, read the next, hash it, and so on. We'd alternate between being IO bound and CPU bound. This doesn't come close to maximizing our performance.

On the other hand, if we had one worker for each file, then they could be reading and hashing at the same time. However, we'd run out of memory on our machine, as we'd effectively be trying to load our filesystem into memory.

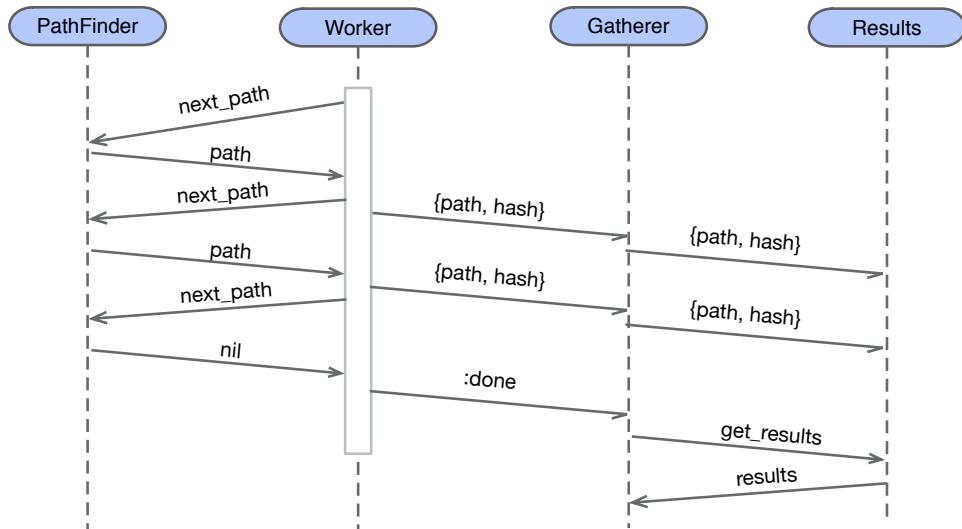
The sweet spot lies in between.

One approach is to create n workers, and then divide the work equally between them. This is the typical push model: plan the work up front and let it execute. The problem with this approach is that it assumes that each file is about the same size. If that's not the case (and on my machine it certainly isn't), then

it would be possible to give one worker mostly small files and another mostly large files. The first would finish early, and would then sit idle while the second chewed through its workload.

The approach I prefer in this scenario is what I call *hungry consumer*. It's a pull model, where each worker asks for the next thing to do, processes it, and then asks for more work. In this scheme a worker that has a small file to process will get it done quickly, then ask for more work. One with a bigger file will take more time. There'll never be an idle worker until we get to the very last files.

The following sequence diagram shows how messages flow in this system. Notice that we have a mixture of synchronous messaging (the pairs of arrows going in opposite directions) and asynchronous messaging.



Q4: What do I protect from errors?

This is where we get to be pragmatic!

In an ideal world, nothing will fail, and everything should be protected from errors.

The real world is different—we can only do so much protecting. How much do we need? It depends. If we're writing software for a pacemaker, then I'd suggest the vast majority of the implementation effort should go into error protection. If we're writing a duplicate-file finder, then not so much.

We can assume that running our finder across 500 GB of files will take minutes, not seconds. That means that if we have a failure reading one file, we

don't want to stop the whole application and lose the work done so far—it is good enough to continue ignoring that file.

What this means:

We want to protect the accumulating results, but we're not as worried about the individual workers—workers can simply restart on failure and process the next file.

Q5. How do I get this thing running?

Sequential programs are easy to start: you just run them. Applications that have many moving parts are more complex: you have to make sure that the various servers are started so that if server A needs to call server B, then B is running before A makes that call.

How should our four servers be started?

The sequence diagram tells us most of the answer.

- Worker depends on PathFinder and Gatherer.
- Gatherer depends on Results.
- Pathfinder and Results depend on nothing.

What this means:

Pathfinder and Results should be started first, followed by Gatherer, and then by the workers.

In terms of implementation, this is fairly straightforward. We simply list the servers in this order as children of a supervisor, and the supervisor will make sure each child is running before starting the next.

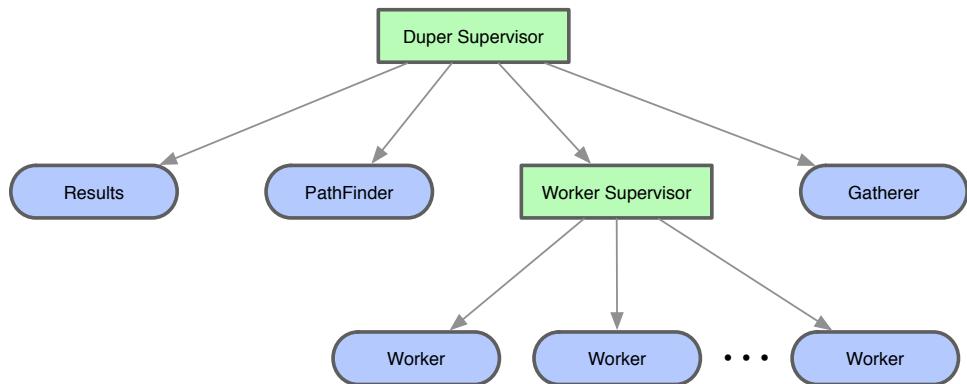
But we also know that we will have multiple workers. Rather than try to find a way of creating them up at the top level, we'll add something new: a subsupervisor. This subsupervisor is responsible for just the pool of workers.

Adding this supervisor opens up an interesting possibility. If all the children of a supervisor are the same, then that supervisor can be used to create them dynamically. Our gatherer server could create a pool of workers when it kicks off the application. This would let us experiment with the effect the number of workers has on elapsed time.

Our new dependency structure looks like this:

- Worker depends on PathFinder and Gatherer.
- Gatherer depends on Results and the worker supervisor.
- Pathfinder and Results depend on nothing.

This gives us a supervision structure that looks like this:



Of course, all this is just theory. Let's start getting some feedback by writing code.

The Duper Application

We'll start by creating a supervised application:

```
$ mix new --sup duper
$ cd duper
$ git init
$ git add .
$ git commit -a -m 'raw application'
```

Time to start writing servers.

The Results Server

The results server wraps an Elixir map. When it starts, it sets its state to an empty map. The keys of this map are hash values, and the values are the list of one or more paths whose files have that hash.

The server provides two API calls: one to add a hash/path pair to the map, the second to retrieve entries that have more than one path in the value (as these are two duplicate files).

This is similar to the code we wrote for the sequence stash:

```
duper/1/duper/lib/duper/results.ex
defmodule Duper.Results do
  use GenServer
  @me __MODULE__
```

```

# API

def start_link(_) do
  GenServer.start_link(__MODULE__, :no_args, name: @me)
end

def add_hash_for(path, hash) do
  GenServer.cast(@me, { :add, path, hash })
end

def find_duplicates() do
  GenServer.call(@me, :find_duplicates)
end

# Server

def init(:no_args) do
  { :ok, %{} }
end

def handle_cast({ :add, path, hash }, results) do
  results =
    Map.update(
      results,
      hash,
      [ path ],
      fn existing -> # if not found, store this value
        [ path | existing ]
      end)
  { :noreply, results }
end

def handle_call(:find_duplicates, _from, results) do
{
  :reply,
  hashes_with_more_than_one_path(results),
  results
}
end

defp hashes_with_more_than_one_path(results) do
  results
  |> Enum.filter(fn { _hash, paths } -> length(paths) > 1 end)
  |> Enum.map(&elem(&1, 1))
end
end

```

The only mild magic in this code is the use of `Map.update/4`. This wonderful function takes a map, a key, an initial value, and a function. If the key is not present in the map, then a new map is returned with that key and initial value added. If the key is present, then the corresponding value is passed to the function, and whatever the function returns becomes the updated value

in the returned map. In our case, we're using it to create a single-element path list the first time a hash is encountered, and then to add paths to that list on duplicates.

We'll add this server to the list of top-level children in application.ex.

```
def start(_type, _args) do
  children = [
    Duper.Results,
  ]
  opts = [strategy: :one_for_one, name: Duper.Supervisor]
  Supervisor.start_link(children, opts)
end
```

This code is easy to test:

```
duper/1/duper/test/duper/results_test.exs
defmodule Duper.ResultsTest do
  use ExUnit.Case
  alias Duper.Results

  test "can add entries to the results" do
    Results.add_hash_for("path1", 123)
    Results.add_hash_for("path2", 456)
    Results.add_hash_for("path3", 123)
    Results.add_hash_for("path4", 789)
    Results.add_hash_for("path5", 456)
    Results.add_hash_for("path6", 999)

    duplicates = Results.find_duplicates()

    assert length(duplicates) == 2
    assert ~w{path3 path1} in duplicates
    assert ~w{path5 path2} in duplicates
  end
end

$ mix test
...
```

Finished in 0.05 seconds
1 doctest, 2 tests, 0 failures

Structuring Tests



I like the directory structure of my tests to follow the same structure as the code it is testing. Because results.ex is in lib/duper/results.ex, I put the test in a subdirectory of test, also called duper.

The PathFinder Server

Our next server is responsible for returning all the file paths in a filesystem tree, one at a time.

Elixir doesn't have a filesystem-traversal API built in, so we look on 'hex.pm' and find `dir_walker`,¹ which we just need to wrap in a trivial GenServer whose state is the directory walker's PID. So we add the dependency to our `mix.exs` file:

```
duper/1/duper/mix.exs
defp deps do
  [
    {dir_walker, "~> 0.0.7"},
  ]
end
```

and code the server in `lib/duper/path_finder.ex`:

```
duper/1/duper/lib/duper/path_finder.ex
defmodule Duper.PathFinder do
  use GenServer

  @me PathFinder

  def start_link(root) do
    GenServer.start_link(__MODULE__, root, name: @me)
  end

  def next_path() do
    GenServer.call(@me, :next_path)
  end

  def init(path) do
    DirWalker.start_link(path)
  end

  def handle_call(:next_path, _from, dir_walker) do
    path = case DirWalker.next(dir_walker) do
      [path] -> path
      other    -> other
    end

    { :reply, path, dir_walker }
  end
end
```

Finally we add the pathfinder server to our application's list of children:

1. https://hex.pm/packages/dir_walker

```
def start(_type, _args) do
  children = [
    Duper.Results,
    { Duper.PathFinder, "." },
  ]

  opts = [strategy: :one_for_one, name: Duper.Supervisor]
  Supervisor.start_link(children, opts)
end
```

Notice we used a tuple to specify the PathFinder server. That's because it requires us to pass in the root of the tree to be searched as a parameter. Here, I'm using the current working directory, “.”, which will work well for playing with the code.

The Worker Supervisor

When we investigated [how we'd get things started, on page 261](#), we realized that we'd need a different supervisor for our workers. This supervisor will only manage the worker servers, and it will let us add servers dynamically, after the application has started.

The simplest way to do this is to use a *DynamicSupervisor*. This type of supervisor allows you to create an arbitrary number of workers at runtime. (A DynamicSupervisor encapsulates what used to be the `:simple_one_for_one` strategy in regular supervisors. You can still do it the old way, but DynamicSupervisors let you express your intent better.)

Let's create the supervisor (in `lib/duper/lib/duper/worker_supervisor.ex`) and then see how it works.

```
duper/1/duper/lib/duper/worker_supervisor.ex
defmodule Duper.WorkerSupervisor do
  use DynamicSupervisor

  @me WorkerSupervisor

  def start_link(_) do
    DynamicSupervisor.start_link(__MODULE__, :no_args, name: @me)
  end

  def init(:no_args) do
    DynamicSupervisor.init(strategy: :one_for_one)
  end

  def add_worker() do
    {:ok, _pid} = DynamicSupervisor.start_child(@me, Duper.Worker)
  end
end
```

The supervisor is just a regular Elixir module. It starts with `use DynamicSupervisor`, which gives it its super(visor) powers.

The `start_link` function works the same in a supervisor as it does in a GenServer: it is called to start the server containing the supervisor code. Inside this server, Elixir automatically calls the `init` callback, which in turn initializes the supervisor code itself. This initialization receives the supervisor options. In the case of a dynamic supervisor, this can only be strategy: `one_for_one`.

Later, we can call the `add_worker` function. This calls the supervisor, telling it to add another child based on the child specification we pass. In this case, we tell it to start `Duper.Worker`. A new server is created for each call, and these servers run in parallel. As a result, each time `add_worker` is called, a new `Duper.Worker` instance is spawned.

Note: One side effect of the fact that the same module is run in multiple child servers is we can't give the children a name in their `start_link` function. If we did, then there'd be multiple servers with the same name, which Elixir doesn't allow.

Let's remember to add the supervisor to the list of top-level children.

```
def start(_type, _args) do
  children = [
    Duper.Results,
    { Duper.PathFinder, "." },
    Duper.WorkerSupervisor,
  ]
  opts = [strategy: :one_for_one, name: Duper.Supervisor]
  Supervisor.start_link(children, opts)
end
```

Thinking About Supervision Strategies

Whenever I add a child to a supervisor's list, I stop and think about the supervision strategy: how do I want the failure of a child managed by this supervisor to affect the other children?

If the results server fails, then all is lost, and we have to restart everything. The same applies to the pathfinder: although we could in theory work out how far into the folder structure we were if it crashed, and restart from there, in practice that would be difficult, so for now we treat a failure of the pathfinder as a failure of the application.

What about the worker supervisor? Here we have to be careful. The worker supervisor handles the actual worker processes. If one of these fails, the worker supervisor simply restarts it and the application continues. But the

failure of a worker does not mean that the worker supervisor itself has failed. In fact, in the very unlikely event that the worker supervisor fails, it's probably best to assume we can't continue and stop the application.

So, for all three children we have on this top-level supervisor, a failure means the application should stop. The strategy that enforces this is `:one_for_all`, so we change our code accordingly.

```
def start(_type, _args) do
  children = [
    Duper.Results,
    { Duper.PathFinder, "." },
    Duper.WorkerSupervisor,
  ]
  > opts = [strategy: :one_for_all, name: Duper.Supervisor]
  Supervisor.start_link(children, opts)
end
```

That's all we're going to do on the worker side of things for now. Let's write the gatherer, and then circle back and implement the actual worker.

The Gatherer Server

Looking at the [sequence diagram on page 260](#), we can see that the gatherer is invoked by the workers. Each worker can tell the gatherer that it has run out of work (by calling `done()`) or it can give it the results of hashing a file.

The gatherer has one more function, not shown in the diagram. It is responsible for starting the workers, and it is responsible for determining when the application has finished processing the files.

To do this, it maintains a simple state: the number of worker servers that are currently running.

Knowing that, we can write most of the gatherer server:

```
duper/1/duper/lib/duper/gatherer.ex
defmodule Duper.Gatherer do
  use GenServer

  @me Gatherer

  # api

  def start_link(worker_count) do
    GenServer.start_link(__MODULE__, worker_count, name: @me)
  end

  def done() do
    GenServer.cast(@me, :done)
  end
```

```

def result(path, hash) do
  GenServer.cast(@me, { :result, path, hash })
end

# server

def init(worker_count) do
  { :ok, worker_count }
end

def handle_cast(:done, _worker_count = 1) do
  report_results()
  System.halt(0)
end

def handle_cast(:done, worker_count) do
  { :noreply, worker_count - 1 }
end

def handle_cast({:result, path, hash}, worker_count) do
  Duper.Results.add_hash_for(path, hash)
  { :noreply, worker_count }
end

defp report_results() do
  IO.puts "Results:\n"
  Duper.Results.find_duplicates()
  |> Enum.each(&IO.inspect/1)
end
end

```

See how the implementation of :done keeps track of the number of running workers? As each signals it is done the count is decremented, until the last :done is received, where we report the results and exit.

We *could* try something like this:

```

def init(worker_count) do
  1..worker_count
  |> Enum.each(fn _ -> Duper.WorkerSupervisor.add_worker() end)
  { :ok, worker_count }
end

```

However, this won't work, and it won't work in a fairly ugly way.

Remember when we described how supervisors started children in order? They wait for each child to initialize itself before starting the next.

In the preceding code, we're still initializing the gatherer when we start adding workers. The workers may start running before the initialization of the gatherer finishes. In this case, messages they send to it may well get lost. If you're only traversing a small filesystem tree, it is even possible that a worker might

signal :done before the gatherer is ready, in which case the system will hang, because it will never know that it has finished.

The answer to this mess is to follow a simple rule: when you're initializing a server, don't interact with anything that uses that server.

So how do we get the workers running? The answer will be familiar to anyone who has written JavaScript—we arrange for a callback into our gatherer server after initialization is complete:

```
def init(worker_count) do
  Process.send_after(self(), :kickoff, 0)
  { :ok, worker_count }
end

def handle_info(:kickoff, worker_count) do
  1..worker_count
  |> Enum.each(fn _ -> Duper.WorkerSupervisor.add_worker() end)
  { :noreply, worker_count }
end
```

Here the init function uses send_after to tell the runtime to queue a message to this server immediately (that is, after waiting 0 ms). When the init function exits, the server is then free to pick up this message, which triggers the handle_info callback, and the workers get started.

So, now that the gatherer code is ready, we just have to remember to start it:

```
duper/1/duper/lib/duper/application.ex
def start(_type, _args) do
  children = [
    Duper.Results,
    { Duper.PathFinder, "/Users/dave/Pictures" },
    Duper.WorkerSupervisor,
    { Duper.Gatherer, 1 },
  ]
  opts = [strategy: :one_for_all, name: Duper.Supervisor]
  Supervisor.start_link(children, opts)
end
```

What About the Workers?

Referring back one last time to [the sequence diagram on page 260](#), we can see that the workers are a little strange: they have no incoming API. All they do is ask for a path, compute the hash of the corresponding file, and send the hash to the gatherer. At some point, there are no paths left, so they then send a :done notification to the gatherer instead.

Here's the code:

```
duper/1/duper/lib/duper/worker.ex
defmodule Duper.Worker do
  use GenServer, restart: :transient

  def start_link(_) do
    GenServer.start_link(__MODULE__, :no_args)
  end

  def init(:no_args) do
    Process.send_after(self(), :do_one_file, 0)
    { :ok, nil }
  end

  def handle_info(:do_one_file, _) do
    Duper.PathFinder.next_path()
    |> add_result()
  end

  defp add_result(nil) do
    Duper.Gatherer.done()
    { :stop, :normal, nil }
  end

  defp add_result(path) do
    Duper.Gatherer.result(path, hash_of_file_at(path))
    send(self(), :do_one_file)
    { :noreply, nil }
  end

  defp hash_of_file_at(path) do
    File.stream!(path, [], 1024*1024)
    |> Enum.reduce(
      :crypto.hash_init(:md5),
      fn (block, hash) ->
        :crypto.hash_update(hash, block)
      end)
    |> :crypto.hash_final()
  end
end
```

Notice we use the same trick in the init() function to call back into ourselves, invoking handle_info(:do_one_file_). This function asks the pathfinder for the next file, and then passes the returned value to add_result().

If the pathfinder returns nil, it has run out of files, so we tell the gatherer that we're done. Otherwise, we call a private function to calculate the hash of the file contents, pass the path and the hash to the gatherer, and then send ourselves another :do_one_file message, causing the whole process to repeat.

Why Can't We Just Write a Looping Function?

We can implement a loop in Elixir using a recursive function call. But the worker server doesn't do this. Instead, it sends itself a message and then exits after processing each file.



The reason is that the Elixir runtime won't let any one invocation of a server hog the CPU forever. Instead it sets a timeout on each call or cast into a GenServer (by default 5 seconds). If the call or cast handler has not returned in that time, the runtime assumes something has gone wrong and terminates the server.

Processing a million files in a loop will take more than 5 seconds. So we instead just process one file per entry into the server, and then queue up another message to process the next on a fresh entry. The result: no timeouts.

One other thing to note—we flagged this server as being transient:

```
use GenServer, restart: :transient
```

This means that the supervisor will not restart it if it terminates normally, but will restart it if it fails.

But Does It Work?

Let's see where we are. We've implemented four GenServers and two supervisors. When the application starts, it will start the top-level supervisor, which in turn starts Results, PathFinder, WorkerSupervisor, and Gatherer.

When Gatherer starts (and it will start last), it tells the worker supervisor to start a number of workers. When each worker starts, it gets a path to process from PathFinder, hashes the corresponding file, and passes the result to Gatherer, which stores the path and the hash in the Results server. When there are no more files to process, each worker sends a :done message to the gatherer. When the last worker is done, the gatherer reports the results.

Everything seems to be wired up. Let's try it:

```
$ mix run
Compiling 7 files (.ex)
Generated duper app
$
```

Hmm...that's strange. No output.

The first time this happened to me, I wasted most of a day working it out. And the problem is obvious once you know what's happening.

The mix run command runs your application. Once it has it running, mix exits: mission accomplished.

But your application never finished; it just got started and mix went away. We have to tell mix not to exit.

```
$ mix run --no-halt
```

Results:

```
[ "./_build/dev/lib/dir_walker/.compile.elixir_scm",
  "./_build/test/lib/dir_walker/.compile.elixir_scm"]
[ "./_build/dev/lib/dir_walker/.compile.elixir",
  "./_build/test/lib/dir_walker/.compile.elixir"]
[ "./_build/dev/lib/dir_walker/.compile.xref",
  "./_build/dev/lib/duper/.compile.xref",
  "./_build/test/lib/dir_walker/.compile.xref"]
[ "./deps/dir_walker/.fetch",
  "./_build/dev/lib/dir_walker/.compile.lock",
  "./_build/dev/lib/dir_walker/.compile.fetch",
  "./_build/test/lib/dir_walker/.compile.lock",
  "./_build/test/lib/dir_walker/.compile.fetch"]
[ "./_build/dev/lib/dir_walker/ebin/dir_walker.app",
  "./_build/test/lib/dir_walker/ebin/dir_walker.app"]$
```

Much better. Even inside our Elixir project we have duplicated files, mostly between the test and dev environments.

Let's Play with Timing

Our lib/duper/application.ex file contains parameters that tell the app where to search and how many workers to use when searching. (We'll see in the next chapter how to move those values out of code and onto the command line.)

Let's change these parameters. My ~/Pictures folder used 30 GB to store about 6,000 old pictures from when I used iPhoto. Let's look for duplicates in that folder with one worker, two workers, and so on, recording elapsed time.

Here are the parameters for using a single worker:

```
children = [
  Duper.Results,
  { Duper.PathFinder,      "/Users/dave/Pictures" },
  Duper.WorkerSupervisor,
  { Duper.Gatherer,        1 },
]
```

Run it:

```
$ time mix run --no-halt >dups
 87.57 real      58.81 user      23.44 sys
$ wc -l dups
 1869 dups
```

We found 1,900-odd duplicated photos in about 88 seconds. The Elixir runtime used about 98% of one of my cores during this process.

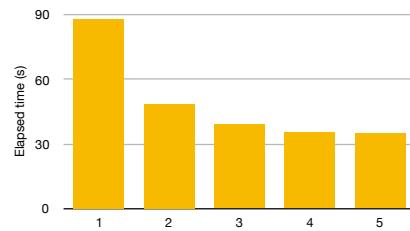
Let's try with two workers. Alter application.ex, and run this:

```
$ time mix run --no-halt >dups
 48.58 real      58.33 user      17.98 sys
```

Nice! It ran almost twice as fast. It means that I'm successfully overlapping the IO and the hashing.

To cut a long story short, here are the results for 1..5, 10, and 50 workers.

# workers	elapsed (S)	user (S)	sys (S)
1	87.57	58.81	23.44
2	48.58	58.33	17.98
3	39.11	70.32	25.89
4	35.55	72.11	27.86
5	34.66	72.80	28.09
10	35.07	72.10	29.26
50	35.70	70.47	32.36



As my machine has only two processors (four cores, but two are just hyper-threading), that's about as good as I could expect.

Planning Your Elixir Application

This book is about thinking differently. We started by thinking about the code we write, and how a function style with immutable data forces us to think in terms of transformations.

In the last few chapters we've come across another dimension of this: thinking about how we structure our application. Our code is no longer monolithic. Instead we think about independent, interacting servers. (You might even call them services.)

This shift in thinking is a difficult one, because it involves both conceptual adjustments and practical deployment issues. It is something you'll become more comfortable with over time. But remember to ask yourself the five questions, and your path should be easier:

- What is the environment and what are its constraints?
- What are the obvious focal points?
- What are the runtime characteristics?
- What do I protect from errors?
- How do I get this thing running?

Next Up

We have an application, but it doesn't really work well for us: things are hardcoded. Let's investigate a little more about what it means to be an application in the next chapter.

OTP: Applications

So far in our quick tour of Elixir and OTP we've looked at server processes and the supervisors that monitor them. There's one more stage in our journey—the application.

This Is Not Your Father's Application

Because OTP comes from the Erlang world, it uses Erlang names for things. And unfortunately, some of these names are not terribly descriptive. The name *application* is one of these. When most of us talk about applications, we think of a program we run to do something—maybe on our computer or phone, or via a web browser. An application is a self-contained whole.

But in the OTP world, that's not the case. Instead, an application is a bundle of code that comes with a descriptor. That descriptor tells the runtime what dependencies the code has, what global names it registers, and so on. In fact, an OTP application is more like a dynamic link library or a shared object than a conventional application.

It might help to see the word *application* in your head but pronounce it *component* or *service*.

For example, back when we were fetching GitHub issues using the HTTPoison library, what we actually installed was an independent application containing HTTPoison. Although it looked like we were just using a library, mix automatically loaded the HTTPoison application. When we then started it, HTTPoison in turn started a couple of other applications that it needed (SSL and Hackney), which in turn kicked off their own supervisors and workers. And all of this was transparent to us.

I've said that applications are components, but some applications are at the top of the tree and are meant to be run directly.

In this chapter we'll look at both types of application component (see what I did there?). In reality they're virtually the same, so let's cover the common ground first.

The Application Specification File

You probably noticed that every now and then mix will talk about a file called *name.app*, where *name* is your application's name.

This file is called an *application specification* and is used to define your application to the runtime environment. Mix creates this file automatically from the information in mix.exs combined with information it gleans from compiling your application.

When you run your application this file is consulted to get things loaded.

Your application does not need to use all the OTP functionality—this file will always be created and referred to. However, once you start using OTP supervision trees, stuff you add to mix.exs will get copied into the .app file.

Turning Our Sequence Program into an OTP Application

So, here's the good news. The application [on page 247](#) is already a full-blown OTP application. When mix created the initial project tree, it added a supervisor (which we then modified) and enough information to our mix.exs file to get the application started. In particular, it filled in the application function:

```
def application do
  [
    mod: {
      Sequence.Application, []
    },
    extra_applications: [:logger],
  ]
end
```

This says that the top-level module of our application is called Sequence. OTP assumes this module will implement a start function, and it will pass that function an empty list as a parameter.

In our previous version of the start function, we ignored the arguments and instead hard-wired the call to start_link to pass 123 to our application. Let's change that to take the value from mix.exs instead. First, change mix.exs to pass an initial value (we'll use 456):

```
def application do
  [
    mod: {
      Sequence.Application, 456
    },
    extra_applications: [:logger],
  ]
end
```

Then change the application.ex code to use this passed-in value:

```
otp-app/sequence/lib/sequence/application.ex
defmodule Sequence.Application do
  @moduledoc false
  use Application
  > def start(_type, initial_number) do
    children = [
  >   {Sequence.Stash, initial_number},
  >   {Sequence.Server, nil},
    ]
    opts = [strategy: :rest_for_one, name: Sequence.Supervisor]
    Supervisor.start_link(children, opts)
  end
end
```

We can check that this works:

```
$ iex -S mix
Compiling 5 files (.ex)
Generated sequence app

iex> Sequence.Server.next_number
456
```

Let's look at the application function again.

The mod: option tells OTP the module that is the main entry point for our app. If our app is a conventional runnable application, then it will need to start somewhere, so we'd write our kickoff function here. But even pure library applications may need to be initialized. (For example, a logging library may start a background logger process or connect to a central logging server.)

For the sequence app, we tell OTP that the Sequence module is the main entry point. OTP will call this module's start function when it starts the application. The second element of the tuple is the parameter to pass to this function. In our case, it's the initial number for the sequence.

There's a second option we'll want to add to this.

The registered: option lists the names that our application will register. We can use this to ensure each name is unique across all loaded applications in a node or cluster. In our case, the sequence server registers itself under the name Sequence.Server, so we'll update the configuration to read as follows:

```
otp-app/sequence/mix.exs
def application do
  [
    mod: {
      Sequence.Application, 456
    },
    registered: [
      Sequence.Server,
    ],
    extra_applications: [:logger],
  ]
end
```

Now that we've done the configuring in mix, we run mix compile, which both compiles the app and updates the sequence.app application specification file with information from mix.exs. (The same thing happens if we run mix using iex -S mix.)

```
$ mix compile
Compiling 5 files (.ex)
➤ Generated sequence app
```

Mix tells us it has created a sequence.app file, but where is it? You'll find it tucked away in _build/dev/lib/sequence/ebin. Although a little obscure, the directory structure under _build is compatible with Erlang's OTP way of doing things. This makes life easier when you release your code. You'll notice that the path has dev in it—this keeps things you're doing in development separate from other build products.

Let's look at the sequence.app that was generated.

```
otp-app/sequence/_build/dev/lib/sequence/ebin/sequence.app
{application,sequence,
  [{applications,[kernel,stdlib,elixir,logger]},
   {description,"sequence"},
   {modules,['Elixir.Sequence','Elixir.Sequence.Application',
             'Elixir.Sequence.Server','Elixir.Sequence.Stash']},
   {vsn,"0.1.0"},
   {mod,{['Elixir.Sequence.Application'],456}},
   {registered,['Elixir.Sequence.Server']},
   {extra_applications,[logger]}]}.
```

This file contains an Erlang tuple that defines the app. Some of the information comes from the project and application section of mix.exs. Mix also automatically added a list of the names of all the compiled modules in our app (the .beam files) and a list of the apps our app depends on (kernel, stdlib, and elixir). That's pretty smart.

More on Application Parameters

In the previous example, we passed the integer 456 to the application as an initial parameter. Although that's valid(ish), we really should have passed in a keyword list instead. That's because Elixir provides a function, Application.get_env, to retrieve these values from anywhere in our code. So we probably should have set up mix.exs with

```
def application do
  [
    mod: { Sequence, [] },
    env: [initial_number: 456],
    registered: [Sequence.Server]
  ]
end
```

and then accessed the value using get_env. We call this with the application name and the name of the environment parameter to fetch:

```
defmodule Sequence do
  use Application

  def start(_type, _args) do
    Sequence.Supervisor.start_link(Application.get_env(:sequence, :initial_number))
  end
end
```

Your call.

Supervision Is the Basis of Reliability

Let's briefly recap. In that last example, we ran our OTP sequence application using mix. Looking at just our code, we see that a supervisor process and two worker processes got started. These were knitted together so our system continued to run with no loss of state even if the worker that we talked to crashed. And any other Erlang process on this node (including IEx itself) can talk to our sequence application and enjoy its stream of freshly minted integers.

You probably noticed that the start function takes two parameters. The second corresponds to the value we specified in the mod: option in the mix.exs file (in

our case, the counter’s initial value). The first parameter specifies the status of the restart, which we’re not going to get into, because...

Your Turn

► [Exercise: OTP-Applications-1](#)

Turn your stack server into an OTP application.

► [Exercise: OTP-Applications-2](#)

So far, we haven’t written any tests for the application. Is there anything you can test? See what you can do.

Releasing Your Code

One way Erlang achieves nine-nines application availability is by having a rock-solid release-management system. Elixir makes this system easy to use.

Before we get too far, let’s talk terminology.

A *release* is a bundle that contains a particular version of your application, its dependencies, its configuration, and any metadata it requires to get running and stay running. A *deployment* is a way of getting a release into an environment where it can be used.

A *hot upgrade* is a kind of deployment that allows the release of a currently running application to be changed while that application continues to run—the upgrade happens in place with no user-detectable disruption.

In this section we’ll talk about releases and hot upgrades. We won’t dig too deeply into deployment.

Distillery—The Elixir Release Manager

Distillery is an Elixir package that makes most release tasks easy. In particular, it can take the complexity that is the source of your project, along with its dependencies, and reduce it down to a single deployable file.

Imagine you were managing the deployment of hundreds of thousands of lines of code into running telephone switches, while maintaining all the ongoing connections, providing a full audit trail, and maintaining contractual uptime guarantees. This is clearly complex. Very complex. And this is the task the Erlang folks faced, so they created tools that help.

Distillery is a layer of abstraction on top of this complexity. Normally it manages to hide it, but sometimes the lower levels leak out and you get to see how the sausage is made.

This book isn't going to get that deep. Instead, I just want to give you a feel for the process.

Before We Start

In Elixir, we version both the application code and the data it operates on. The two are independent—we might go for a dozen code releases without changing any data structures.

The code version is stored in the project dictionary in `mix.exs`. But how do we version the data? Come to think of it, where do we even define the data?

In an OTP application, all state is maintained by servers, and each server's state is independent. So it makes sense to version the app data within each server module. Perhaps a server initially holds its state in a two-element tuple. That could be version 0. Later, it is changed to hold state in a three-element tuple. That could be version 1.

We'll see the significance of this later. For now, let's just set the version of the state data in our server. We use the `@vsn` (version) directive:

```
otp-app/sequence_v0/lib/sequence/server.ex
defmodule Sequence.Server do
  use GenServer
  @vsn "0"
```

Now let's generate a release.

Your First Release

First, we have to add distillery as a project dependency. Open the sequence project's `mix.exs` file and update the `deps` function.

```
otp-app/sequence_v0/mix.exs
defp deps do
  [
    {:distillery, "~> 1.5", runtime: false},
  ]
end
```

(The `runtime: false` option tells mix that Distillery is not to be started with the running application.)

Remember to install the dependency:

```
$ mix do deps.get, deps.compile
```

Distillery makes sensible choices for the various configuration options, so for a basic app like this we're now ready to create our first release. To start off we create the release configuration:

```
$ mix release.init
```

An example config file has been placed in `rel/config.exs`, review it, make edits as needed/desired, and then run `'mix release'` to build the release

If you're following along at home, have a look at `rel.config.exs` (it's too large to show here). The defaults look good, so let's build the actual release. We'll tell it we want a production version of the release. If we don't, the release it generates will be for development, and won't be a self-contained package.

```
$ mix release --env=prod
==> Assembling release..
==> Building release sequence:0.1.0 using environment prod
==> Including ERTS 9.1 from /usr/local/Cellar/erlang/20.1/lib/erlang/erts-9.1
==> Packaging release..
==> Release successfully built!
You can run it in one of the following ways:
  Interactive: _build/dev/rel/sequence/bin/sequence console
  Foreground: _build/dev/rel/sequence/bin/sequence foreground
  Daemon: _build/dev/rel/sequence/bin/sequence start
```

Distillery got the application name and version number from your `mix.exs` file, and packaged your app into the `_build/dev/rel/` directory:

```
_build/dev/rel
└── sequence
    ├── bin                                « global scripts
    │   ├── nodetool
    │   ├── release_utils.script
    │   ├── sequence
    │   ├── sequence.bat
    │   ├── sequence_loader.sh
    │   └── start_clean.boot
    ├── erts-9.1                            « the runtime (Erlang + Elixir)
    │   ├── ...
    │   ├── elixir-1.5.2
    │   └── sequence-0.1.0                    « our compiled application
    │       └── consolidated
    │           ├── Elixir.Collectable.beam
    │           ├── Elixir.Enumerable.beam
    │           ├── Elixir.IEx.Info.beam
    │           ├── Elixir.Inspect.beam
    │           ├── Elixir.List.Chars.beam
    │           └── Elixir.String.Chars.beam
```

```

└── ebin
    ├── Elixir.Sequence.Application.beam
    ├── Elixir.Sequence.Server.beam
    ├── Elixir.Sequence.Stash.beam
    └── Elixir.Sequence.beam
        └── sequence.app

└── stdlib-3.4.2
└── releases
    └── 0.1.0      « release specific stuff
        └── « our initial release
            ├── commands
            ├── hooks
            ├── ...
            ├── libexec
            ├── ...
            ├── sequence.bat
            ├── sequence.boot
            ├── sequence.rel
            ├── sequence.script
            ├── sequence.sh
            ├── sequence.tar.gz « the packaged release
            ├── start_clean.boot
            ├── sys.config
            └── vm.args
    └── RELEASES
    └── start_erl.data

```

The most important file is `rel/sequence/releases/0.0.1/sequence.tar.gz`. It contains everything needed to run this release. This is the file we deploy to our servers.

A Toy Deployment Environment

I don't want to slow things down by having you provision a server in the cloud, so I'm going to deploy to my local machine. However, to make it a little more realistic, I'll pretend this machine is remote, and use `ssh` to do all the deploying. I'll also be creating directories and copying files manually. In practice, you'd want to automate all of this with something like Capistrano or Ansible.

We'll store the releases in a `deploy` directory. I'll put this inside my home directory—feel free to put it anywhere (writable) you want.

```
$ ssh localhost mkdir ~/deploy
```

Deploy and Run the App

Now we need to set up the initial release and its directory structure. Copy the `sequence.tar.gz` file into the `deploy` directory, and then extract its contents.

```
$ scp _build/dev/rel/sequence/releases/0.1.0/sequence.tar.gz localhost:deploy
$ ssh localhost tar -x -f ~/deploy/sequence.tar.gz -C ~/deploy
```

The app is now ready to run. The scripts in deploy/bin control it. These, in turn, delegate to scripts in the current release directory (all on the server).

Let's start an IEx console. (The ssh -t option lets us control the remote IEx with ^C and ^G.)

```
$ ssh -t localhost ~/deploy/bin/sequence console
Using /Users/dave/deploy/releases/0.0.1/sequence.sh
Interactive Elixir (1.x) - press Ctrl+C to exit (type h() ENTER for help)

iex(sequence@127.0.0.1)2> Sequence.Server.next_number
456
iex(sequence@127.0.0.1)3> Sequence.Server.next_number
457
```

(Leave this session running—we'll use it to demonstrate hot code reloading.)

A Second Release

Our marketing team ran a focus group. It seems our customers want the next_number function to return a message like “the next number is 458.”

First we'll change server.ex:

```
otp-app/sequence_v1/lib/sequence/server.ex
def next_number do
  with number = GenServer.call(__MODULE__, :next_number),
    do: "The next number is #{number}"
end
```

Then we'll bump the application's version number in mix.exs.

```
otp-app/sequence_v1/mix.exs
def project do
  [
    app: :sequence,
    version: "0.2.0",
    elixir: "~> 1.6-dev",
    start_permanent: Mix.env() == :prod,
    deps: deps()
  ]
end
```

(We don't have to change the @vsn value—the representation of the server's state is not affected by this change.)

After exhaustive testing, we decide we're ready to create a new release. Here we have a choice. If we just run mix release we'll create a whole new releasable application. To deploy it, we'd basically copy it just as we did before, then stop the old app and start the new one.

The alternative is to deploy an *upgrade release* (sometimes called a *hot upgrade*). This is code that will upgrade the application while it is still running—there should be no downtime. This is one reason why Elixir apps can achieve such high availability numbers. Let's take the upgrade path:

```
$ mix release --env=prod --upgrade
==> Assembling release..
==> Building release sequence:0.2.0 using environment prod
==> Including ERTS 9.1 from /usr/local/Cellar/erlang/20.1/lib/erlang/erts-9.1
➤ ==> Generated .appup for sequence 0.1.0 -> 0.2.0
==> Relup successfully created
==> Packaging release..
==> Release successfully built!
You can run it in one of the following ways:
  Interactive: _build/dev/rel/sequence/bin/sequence console
  Foreground: _build/dev/rel/sequence/bin/sequence foreground
  Daemon: _build/dev/rel/sequence/bin/sequence start
```

The key thing to note is the creation of the `.appup` file. This is what tells the Erlang runtime how to upgrade our running app.

Deploying an Upgrade

The deployment of the first release of an app is special: it has to create an environment for that app. With that in place, this release (and all subsequent releases) will be slightly different. We have to create a release directory on the server and copy the tarball into it. The directory will be under `deploy/releases`, and will be named the same as the release's version number.

```
$ ssh localhost mkdir deploy/releases/0.2.0
$ scp _build/dev/rel/sequence/releases/0.2.0/sequence.tar.gz \
    localhost:deploy/releases/0.2.0
```

Now let's upgrade the running code:

```
$ ssh localhost ~/deploy/bin/sequence upgrade 0.2.0
Release 0.2.0 not found, attempting to unpack releases/0.2.0/sequence.tar.gz
Unpacked successfully: "0.2.0"
Release 0.2.0 is already unpacked, now installing.
Installed Release: 0.2.0
Made release permanent: "0.2.0"
```

Head back over to the terminal session that's talking to the app. Don't restart it—just make another request:

```
iex(sequence@127.0.0.1)4> Sequence.Server.next_number
"The next number is 458"
iex(sequence@127.0.0.1)5> Sequence.Server.next_number
"The next number is 459"
```

Erlang can actually run two versions of a module at the same time. Currently executing code will continue to use the old version until that code explicitly cites the name of the module that has changed. At that point, and for that particular process, execution will swap to the new version.

This is a critical part of hot loading of code. We want to let code that is currently running continue without interruption, but the new release may not be compatible with it. So Erlang lets it run on the old release. But the next request will reference the module explicitly, and the new code will be loaded.

Here, when we say `Sequence.Server.next_number`, the reference to `Sequence.Server` triggers the reload, so the 0.2.0 release handles the next request.

What if our new release was a disaster? That's not a problem—we can always downgrade to a previous version.

```
$ ssh localhost ~/deploy/bin/sequence downgrade 0.1.0
Release 0.1.0 is already unpacked
Release 0.1.0 is marked old, switching to it.
Installed Release: 0.1.0
Made release permanent: "0.1.0"

Warning: "/Users/dave/deploy/releases/0.0.1/relup" missing (optional)

iex(sequence@127.0.0.1)6> Sequence.Server.next_number
460
iex(sequence@127.0.0.1)7> Sequence.Server.next_number
461
```

Cool. Let's go back to the current version before continuing.

```
$ ssh localhost ~/deploy/bin/sequence upgrade 0.2.0
```

Migrating Server State

Our boss calls. We're about to go for a second round of funding for our wildly successful sequence-server business, but customers have noticed a bug. We implemented `increment_number` to add a delta to the current number—a one-time change. But apparently it was instead supposed to set the difference between successive numbers we served.

Let's try the existing code in our already-running console:

```
iex(sequence@127.0.0.1)8> Sequence.Server.next_number
The next number is 462
iex(sequence@127.0.0.1)9> Sequence.Server.increment_number 10
:ok
iex(sequence@127.0.0.1)10> Sequence.Server.next_number
The next number is 472
iex(sequence@127.0.0.1)10> Sequence.Server.next_number
The next number is 473
```

Yup, we're applying the delta only once.

Well, that's an easy change to the code. We simply have to keep one extra thing in the state—a delta value. Here's the updated server code:

```
otp-app/sequence_v2/lib/sequence/server.ex
defmodule Sequence.Server do
  use GenServer
  require Logger

  @vsn "1"

  defmodule State do
    defstruct(current_number: 0, delta: 1)
  end

  #####
  # External API

  def start_link(_) do
    GenServer.start_link(__MODULE__, nil, name: __MODULE__)
  end

  def next_number do
    with number = GenServer.call(__MODULE__, :next_number),
    do: "The next number is #{number}"
  end

  def increment_number(delta) do
    GenServer.cast __MODULE__, {:increment_number, delta}
  end

  #####
  # GenServer implementation

  def init(_) do
    state = %State{ current_number: Sequence.Stash.get() }
    { :ok, state }
  end

  def handle_call(:next_number, _from, state = %{current_number: n}) do
    { :reply, n, %{state | current_number: n + state.delta} }
  end

  def handle_cast({:increment_number, delta}, state) do
    { :noreply, %{state | delta: delta } }
  end

  def terminate(_reason, current_number) do
    Sequence.Stash.update(current_number)
  end
end
```

The big change is that we made the state a struct rather than a tuple and added the delta value. We updated the increment handler to change the value of delta, and the next number handler now adds in the delta each time.

The format of the state changed, so we updated the version number (@vsn) to 1.

If we simply stop the old server and start the new one, we'll lose the state stored in the old one. But we can't just copy the state across—the old server had a single integer and the new one has a struct.

Fortunately, OTP has a callback for this. In the new server, implement the code_change function.

```
otp-app/sequence_v2/lib/sequence/server.ex
def code_change("0", old_state = current_number, _extra) do
  new_state = %State{
    current_number: current_number,
    delta:           1
  }
  Logger.info "Changing code from 0 to 1"
  Logger.info inspect(old_state)
  Logger.info inspect(new_state)
  { :ok, new_state }
end
```

The callback takes three arguments—the old version number, the old state, and an additional parameter we don't use. The callback's job is to return `{:ok, new_state}`. In our case, the new state is a struct containing the stash PID and the old current number, along with the new delta value, initialized to 1. We'll need to bump the version number in mix.exs.

```
otp-app/sequence_v2/mix.exs
def project do
  [
    app: :sequence,
    version: "0.3.0",
    elixir: "~> 1.6-dev",
    start_permanent: Mix.env() == :prod,
    deps: deps()
  ]
end
```

Time to create the new release:

```
mix release --env=prod --upgrade
==> Assembling release..
==> Building release sequence:0.3.0 using environment prod
==> Including ERTS 9.1 from /usr/local/Cellar/erlang/20.1/lib/erlang/erts-9.1
==> Generated .appup for sequence 0.2.0 -> 0.3.0
==> Relup successfully created
```

```

==> Packaging release..
==> Release successfully built!
  You can run it in one of the following ways:
    Interactive: _build/dev/rel/sequence/bin/sequence console
    Foreground: _build/dev/rel/sequence/bin/sequence foreground
    Daemon: _build/dev/rel/sequence/bin/sequence start

```

Copy it into the deployment location:

```

$ ssh localhost mkdir ~/deploy/releases/0.3.0/
$ scp _build/dev/rel/sequence/releases/0.3.0/sequence.tar.gz \
  localhost:deploy/releases/0.3.0/

```

Cross your fingers, and upgrade the app:

```

$ ssh localhost ~/deploy/bin/sequence upgrade 0.3.0
Release 0.3.0 not found, attempting to unpack releases/0.3.0/sequence.tar.gz
Unpacked successfully: "0.3.0"
Release 0.3.0 is already unpacked, now installing.
Installed Release: 0.3.0
Made release permanent: "0.3.0"

```

But the real magic happened over in the console window:

```

16:03:12.096 [info]  Changing code from 0 to 1
16:03:12.096 [info]  459
16:03:12.096 [info]  %Sequence.Server.State{current_number: 459, delta: 1}

```

That's the logging we added to our `code_change` function. We seem to have migrated the server's state into our new structure. Let's try it out:

```

iex(sequence@127.0.0.1)10> Sequence.Server.next_number
"The next number is 459"
iex(sequence@127.0.0.1)11> Sequence.Server.increment_number 10
:ok
iex(sequence@127.0.0.1)13> Sequence.Server.next_number
"The next number is 460"
iex(sequence@127.0.0.1)14> Sequence.Server.next_number
"The next number is 470"

```

That's the new behavior, running with our new state structure. We updated the code twice and migrated data once, all while the application continued to run. There was no service disruption, and no loss of data.

Plutarch records the story of a ship called *Theseus*. Over the course of many years most of the ship's structure was replaced, piece by piece. While this was happening, the ship was in continuous use. Plutarch raises the question, “Is the renovated *Theseus* the same as the original?”

Using Elixir release management, our applications can work the same way the *Theseus* did, running continuously but being updated all the time.

Is the latest application the same as the original? Who cares, as long as it's still running?

OTP Is Big—Unbelievably Big

This book barely scratches OTP's surface. But (I hope) it does introduce the major concepts and give you an idea of what's possible.

More advanced uses of OTP may include release management (including hot code-swapping), handling of distributed failover, automated scaling, and so on. But if you have an application that needs such things, you likely will already have or will soon need dedicated operations experts who know the low-level details of making OTP apps perform the way you need them to.

There is never anything simple about scaling out to the kind of size and sophistication that is possible with OTP. But now you know you can start small and get there.

However, there are ways of writing *some* OTP servers more simply, and that's the subject of the next chapter.

Your Turn

► *Exercise: OTP-Applications-3*

Our boss notices that after we applied our version-0-to-version-1 code change, the delta indeed works as specified. However, she also notices that if the server crashes, the delta is forgotten—only the current number is retained. Create a new release that stashes both values.

Tasks and Agents

This part of the book is about processes and process distribution. So far we've covered two extremes. In the first chapters we looked at the `spawn` primitive, along with message sending and receiving and multinode operations. We then looked at OTP, the 800-pound gorilla of process architecture.

Sometimes, though, we want something in the middle. We want to be able to run simple processes, either for background processing or for maintaining state. But we don't want to be bothered with the low-level details of `spawn`, `send`, and `receive`, and we really don't need the extra control that writing our own `GenServer` gives us.

Enter tasks and agents, two simple-to-use Elixir abstractions. These use OTP's features but insulate us from these details.

Tasks

An Elixir task is a function that runs in the background.

```
tasks/tasks1.exs
defmodule Fib do
  def of(0), do: 0
  def of(1), do: 1
  def of(n), do: Fib.of(n-1) + Fib.of(n-2)
end

IO.puts "Start the task"
worker = Task.async(fn -> Fib.of(20) end)
IO.puts "Do something else"
# ...
IO.puts "Wait for the task"
result = Task.await(worker)

IO.puts "The result is #{result}"
```

The call to `Task.async` creates a separate process that runs the given function. The return value of `async` is a task descriptor (actually a PID and a ref) that we'll use to identify the task later.

Once the task is running, the code continues with other work. When it wants to get the function's value, it calls `Task.await`, passing in the task descriptor. This call waits for our background task to finish and returns its value.

When we run this, we see

```
$ elixir tasks1.exs
Start the task
Do something else
Wait for the task
The result is 6765
```

We can also pass `Task.async` the name of a module and function, along with any arguments. Here are the changes:

```
tasks/tasks2.exs
worker = Task.async(Fib, :of, [20])
result = Task.await(worker)
IO.puts "The result is #{result}"
```

Tasks and Supervision

Tasks are implemented as OTP servers, which means we can add them to our application's supervision tree. We can do this in a number of ways.

First, we can link a task to a currently supervised process by calling `start_link` instead of `async`. This has less impact than you might think. If the function running in the task crashes and we use `start_link`, our process will be terminated immediately. If instead we use `async`, our process will be terminated only when we subsequently call `await` on the crashed task.

The second way to supervise tasks is to run them directly from a supervisor. Here we specify the `Task` module itself as the module to run, and pass it the function to be run in the background as a parameter.

```
children = [
  { Task, fn -> do_something_extraordinary() end }
]
Supervisor.start_link(children, strategy: :one_for_one)
```

You can take this approach a step further by moving the task's code out of the supervisor and into its own module.

```
tasks/my_app/lib/my_app/my_task.ex
defmodule MyApp.MyTask do
  use Task

  def start_link(param) do
    Task.start_link(__MODULE__, :thing_to_run, [ param ])
  end

  def thing_to_run(param) do
    IO.puts "running task with #{param}"
  end
end
```

The key thing here is `use Task`. This defines a `child_spec` function, allowing this module to be supervised:

```
children = [
{ MyApp.MyTask, 123 }
]
```

The problem with this approach is that you can't use `Task.await`, because your code is not directly calling `Task.async`.

The solution to this is to supervise the tasks dynamically. This is similar in concept to using a `:simple_one_for_one` supervisor strategy for regular servers. See the `Task` documentation for details.¹

However, before you get too carried away, remember that a simple `start_link` in an already-supervised process may well be all you need.

Agents

An agent is a background process that maintains state. This state can be accessed at different places within a process or node, or across multiple nodes.

The initial state is set by a function we pass in when we start the agent.

We can interrogate the state using `Agent.get`, passing it the agent descriptor and a function. The agent runs the function on its current state and returns the result.

We can also use `Agent.update` to change the state held by an agent. As with the `get` operator, we pass in a function. Unlike with `get`, the function's result becomes the new state.

Here's a bare-bones example. We start an agent whose state is the integer 0. We then use the identity function, `&(&1)`, to return that state. Calling `Agent.update` with `&(&1+1)` increments the state, as verified by a subsequent `get`.

1. <https://hexdocs.pm/elixir/Task.html>

```
iex> { :ok, count } = Agent.start(fn -> 0 end)
{:ok, #PID<0.69.0>}
iex> Agent.get(count, &(&1))
0
iex> Agent.update(count, &(&1+1))
:ok
iex> Agent.update(count, &(&1+1))
:ok
iex> Agent.get(count, &(&1))
2
```

In the previous example, the variable count holds the agent process's PID. We can also give agents a local or global name and access them using this name. In this case we exploit the fact that an uppercase bareword in Elixir is converted into an atom with the prefix Elixir., so when we say Sum it is actually the atom :Elixir.Sum.

```
iex> Agent.start(fn -> 1 end, name: Sum)
{:ok, #PID<0.78.0>}
iex> Agent.get(Sum, &(&1))
1
iex> Agent.update(Sum, &(&1+99))
:ok
iex> Agent.get(Sum, &(&1))
100
```

The following example shows a more typical use. The Frequency module maintains a list of word/frequency pairs in a map. The dictionary itself is stored in an agent, which is named after the module.

This is all initialized with the start_link function, which, presumably, is invoked during application initialization.

```
tasks/agent_dict.exs
defmodule Frequency do
  def start_link do
    Agent.start_link(fn -> %{} end, name: __MODULE__)
  end
  def add_word(word) do
    Agent.update(__MODULE__,
      fn map ->
        Map.update(map, word, 1, &(&1+1))
      end)
  end
  def count_for(word) do
    Agent.get(__MODULE__, fn map -> map[word] end)
  end
```

```

def words do
  Agent.get(__MODULE__, fn map -> Map.keys(map) end)
end
end

```

We can play with this code in IEx.

```

iex> c "agent_dict.exs"
[Frequency]
iex> Frequency.start_link
{:ok, #PID<0.101.0>}
iex> Frequency.add_word "dave"
:ok
iex> Frequency.words
["dave"]
iex(41)> Frequency.add_word "was"
:ok
iex> Frequency.add_word "here"
:ok
iex> Frequency.add_word "he"
:ok
iex> Frequency.add_word "was"
:ok
iex> Frequency.words
["he", "dave", "was", "here"]
iex> Frequency.count_for("dave")
1
iex> Frequency.count_for("was")
2

```

In a way, you can look at our `Frequency` module as the implementation part of a `gen_server`—using agents has simply abstracted away all the housekeeping we had to do.

A Bigger Example

Let's rewrite our anagram code to use both tasks and an agent.

We'll load words in parallel from a number of separate dictionaries. A separate task handles each dictionary. We'll use an agent to store the resulting list of words and signatures.

```

tasks/anagrams.exs
defmodule Dictionary do
  @name __MODULE__
  ##
  # External API

```

```

def start_link,
do: Agent.start_link(fn -> %{} end, name: @name)

def add_words(words),
do: Agent.update(@name, &do_add_words(&1, words))

def anagrams_of(word),
do: Agent.get(@name, &Map.get(&1, signature_of(word)))

<##>
# Internal implementation

defp do_add_words(map, words),
do: Enum.reduce(words, map, &add_one_word(&1, &2))

defp add_one_word(word, map),
do: Map.update(map, signature_of(word), [word], &[word|&1])

defp signature_of(word),
do: word |> to_charlist |> Enum.sort |> to_string

end

defmodule WordlistLoader do
  def load_from_files(file_names) do
    file_names
    |> Stream.map(fn name -> Task.async(fn -> load_task(name) end) end)
    |> Enum.map(&Task.await/1)
  end

  defp load_task(file_name) do
    File.stream!(file_name, [], :line)
    |> Enum.map(&String.trim/1)
    |> Dictionary.add_words
  end
end

```

Our four wordlist files contain the following:

list1	list2	list3	list4
angor	ester	palet	rogan
argon	estre	patel	ronga
caret	goran	pelta	steer
carte	grano	petal	stere
cater	groan	pleat	stree
crate	leapt	react	terse
creat	nagor	recta	tsere
creta	orang	reest	tepal

Let's run it:

```

$ iex anagrams.exs
iex> Dictionary.start_link
{:ok, #PID<0.66.0>}
iex> Enum.map(1..4, &"words/list#{&1}") |> WordlistLoader.load_from_files
[:ok, :ok, :ok, :ok]

```

```
iex> Dictionary.anagrams_of "organ"
["ronga", "rogan", "orang", "nagor", "groan", "grano", "goran",
 "argon", "angor"]
```

Making It Distributed

Agents and tasks run as OTP servers, so they are easy to distribute—just give our agent a globally accessible name. That's a one-line change:

```
@name {:global, __MODULE__}
```

Now we'll load our code into two separate nodes and connect them. (Remember that we have to specify names for the nodes so they can talk.)

Window #1
\$ iex --sname one anagrams_dist.exs iex(one@FasterAir)>
Window #2
\$ iex --sname two anagrams_dist.exs iex(two@FasterAir)> Node.connect :one@FasterAir true iex(two@FasterAir)> Node.list [:one@FasterAir]

We'll start the dictionary agent in node one—this is where the actual dictionary will end up. We'll then load the dictionary using both nodes one and two:

Window #1
iex(one@FasterAir)> Dictionary.start_link {:ok, #PID<0.68.0>} iex(one@FasterAir)> WordlistLoader.load_from_files(~w{words/list1 words/list2}) [:ok, :ok]
Window #2
iex(two@FasterAir)> WordlistLoader.load_from_files(~w{words/list3 words/list4}) [:ok, :ok]

Finally, we'll query the agent from both nodes:

Window #1
iex(one@FasterAir)> Dictionary.anagrams_of "argon" ["ronga", "rogan", "orang", "nagor", "groan", "grano", "goran", "argon", "angor"]
Window #2
iex(two@FasterAir)> Dictionary.anagrams_of "crate" ["recta", "react", "creta", "creat", "crate", "cater", "carte", "caret"]

Agents and Tasks, or GenServer?

When do you use agents and tasks, and when do you use a GenServer?

The answer is to use the simplest approach that works. Agents and tasks are great when you're dealing with very specific background activities, whereas GenServers (as their name suggests) are more general.

You can eliminate the need to make a decision by wrapping your agents and tasks in modules, as we did in our anagram example. That way you can always switch from the agent or task implementation to the full-blown GenServer without affecting the rest of the code base.

It's time to move on and look at some advanced Elixir.

Part III

More Advanced Elixir

Among the joys of Elixir is that it laughs at the concept of “what you see is what you get.” Instead, you can extend it in many different ways. This allows you to add layers of abstraction to your code, which makes your code easier to work with.

This part covers macros (which let you extend the language’s syntax), protocols (which let you add behaviors to existing modules), and use (which lets you add capabilities to a module). We finish with a grab-bag chapter of miscellaneous Elixir tricks and tips.

Macros and Code Evaluation

Have you ever felt frustrated that a language didn't have just the right feature for some code you were writing? Or have you found yourself repeating chunks of code that weren't amenable to factoring into functions? Or have you just wished you could program closer to your problem domain?

If so, then you'll love this chapter.

But, before we get into the details, here's a warning: macros can easily make your code harder to understand, because you're essentially rewriting parts of the language. For that reason, never use a macro when you could use a function. Let's repeat that:



Never use a macro when you could use a function.

In fact, you'll probably not write a macro in regular application code. But if you're writing a library and want to use some of the metaprogramming techniques that we show in later chapters, you'll need to know how macros work.

Implementing an if Statement

Let's imagine that Elixir didn't have an `if` statement—that all it has is `case`. Although we're prepared to abandon our old friend the `while` loop, not having an `if` statement is just too much to bear, so we set about implementing one.

We'll want to call it using something like

```
myif «condition» do
  «evaluate if true»
else
  «evaluate if false»
end
```

We know that blocks in Elixir are converted into keyword parameters, so this is equivalent to

```
myif <>condition>>,
  do: <>evaluate if true>>,
  else: <>evaluate if false>>
```

Here's a sample call:

```
My.myif 1==2, do: (IO.puts "1 == 2"), else: (IO.puts "1 != 2")
```

Let's try to implement myif as a function:

```
defmodule My do
  def myif(condition, clauses) do
    do_clause = Keyword.get(clauses, :do, nil)
    else_clause = Keyword.get(clauses, :else, nil)

    case condition do
      val when val in [false, nil]
        -> else_clause
      _otherwise
        -> do_clause
    end
  end
end
```

When we run it, we're (mildly) surprised to get the following output:

```
iex> My.myif 1==2, do: (IO.puts "1 == 2"), else: (IO.puts "1 != 2")
1 == 2
1 != 2
:ok
```

When we call the myif function, Elixir has to evaluate all of its parameters before passing them in. So both the do: and else: clauses are evaluated, and we see their output. Because IO.puts returns :ok on success, what actually gets passed to myif is

```
myif 1==2, do: :ok, else: :ok
```

This is why the final return value is :ok.

We need a way of delaying the execution of these clauses. This is where macros come in. But before we implement our myif macro, we need a little background.

Macros Inject Code

Let's pretend we're the Elixir compiler. We read a module's source top to bottom and generate a representation of the code we find. That representation is a nested Elixir tuple.

If we want to support macros, we need a way to tell the compiler that we'd like to manipulate a part of that tuple. We do that using defmacro, quote, and unquote.

In the same way that def defines a function, defmacro defines a macro. You'll see what that looks like shortly. However, the real magic starts not when we define a macro, but when we use one.

When we pass parameters to a macro, Elixir doesn't evaluate them. Instead, it passes them as tuples representing their code. We can examine this behavior using a simple macro definition that prints out its parameter.

```
macros/dumper.exs
defmodule My do
  defmacro macro(param) do
    IO.inspect param
  end
end

defmodule Test do
  require My

  # These values represent themselves
  My.macro :atom      #=> :atom
  My.macro 1          #=> 1
  My.macro 1.0        #=> 1.0
  My.macro [1,2,3]    #=> [1,2,3]
  My.macro "binaries" #=> "binaries"
  My.macro { 1, 2 }    #=> {1,2}
  My.macro do: 1      #=> [do: 1]

  # And these are represented by 3-element tuples

  My.macro { 1,2,3,4,5 }
  # => { :"{}",[line: 20],[1,2,3,4,5]}

  My.macro do: ( a = 1; a+a )
  # => [do:
  #       {:_block_,[]},
  #       [{:=,[line: 22],[{a,[line: 22],nil},1]},,
  #        {:+,[line: 22],[{a,[line: 22],nil},{a,[line: 22],nil}]}]]}

  My.macro do
    1+2
  else
    3+4
  end
  # => [do: {:+,[line: 24],[1,2]},,
  #       else: {:+,[line: 26],[3,4]}]
end
```

This shows us that atoms, numbers, lists (including keyword lists), binaries, and tuples with two elements are represented internally as themselves. All other Elixir code is represented by a three-element tuple. Right now, the internals of that representation aren't important.

Load Order

You may be wondering about the structure of the preceding code. We put the macro definition in one module, and the usage of that macro in another. And that second module included a require call.

Macros are expanded before a program executes, so the macro defined in one module must be available as Elixir is compiling another module that uses those macros. The require function tells Elixir to ensure the named module is compiled before the current one. In practice it is used to make the macros defined in one module available in another.

But the reason for the two modules is less clear. It has to do with the fact that Elixir first compiles source files and then runs them.

If we have one module per source file and we reference a module in file A from file B, Elixir will load the module from A, and everything just works. But if we have a module and the code that uses it in the same file, and the module is defined in the same scope in which we use it, Elixir will not know to load the module's code. We'll get this error:

```
** (CompileError)
.../dumper.ex:7:
module My is not loaded but was defined. This happens because you
are trying to use a module in the same context it is defined. Try
defining the module outside the context that requires it.
```

By placing the code that uses the module `My` in a separate module, we force `My` to load.

The quote Function

We've seen that when we pass parameters to a macro they are not evaluated. The language comes with a function, `quote`, that also forces code to remain in its unevaluated form. `quote` takes a block and returns the internal representation of that block. We can play with it in IEx:

```
iex> quote do: :atom
:atom
iex> quote do: 1
1
iex> quote do: 1.0
1.0
iex> quote do: [1,2,3]
[1,2,3]
iex> quote do: "binaries"
"binaries"
iex> quote do: {1,2}
{1,2}
iex> quote do: [do: 1]
[do: 1]
iex> quote do: {1,2,3,4,5}
{:[],[],[1,2,3,4,5]}
iex> quote do: (a = 1; a + a)
{:__block__, [],
 [{:=, [], [{:a, [], Elixir}, 1]}, 
  {:+, [context: Elixir, import: Kernel], 
   [{:a, [], Elixir}, {:a, [], Elixir}]}]}
iex> quote do: [ do: 1 + 2, else: 3 + 4]
[do: {:+, [context: Elixir, import: Kernel], [1, 2]}, 
 else: {:+, [context: Elixir, import: Kernel], [3, 4]}]
```

There's another way to think about quote. When we write "abc", we create a binary containing a string. The double quotes say, "interpret what follows as a string of characters and return the appropriate representation."

quote is the same: it says, "interpret the content of the block that follows as code, and return the internal representation."

Using the Representation as Code

When we extract the internal representation of some code (either via a macro parameter or using quote), we stop Elixir from adding it automatically to the tuples of code it is building during compilation—we've effectively created a free-standing island of code. How do we inject that code back into our program's internal representation?

There are two ways.

The first is our old friend the macro. Just like with a function, the value a macro returns is the last expression evaluated in that macro. That expression is expected to be a fragment of code in Elixir's internal representation. But Elixir does not return this representation to the code that invoked the macro. Instead it injects the code back into the internal representation of our program

and returns to the caller the result of *executing* that code. But that execution takes place only if needed.

We can demonstrate this in two steps. First, here's a macro that simply returns its parameter (after printing it out). The code we give it when we invoke the macro is passed as an internal representation, and when the macro returns that code, that representation is injected back into the compile tree.

```
macros/eg.exs
defmodule My do
  defmacro macro(code) do
    IO.inspect code
    code
  end
end
defmodule Test do
  require My
  My.macro(IO.puts("hello"))
end
```

When we run this, we see

```
{{:.,[line: 11],[{:_aliases_,[line: 11],[:IO]},:puts]}, [line: 11],["hello"]}
hello
```

Now we'll change that file to return a different piece of code. We use quote to generate the internal form:

```
macros/eg1.exs
defmodule My do
  defmacro macro(code) do
    IO.inspect code
    quote do: IO.puts "Different code"
  end
end
defmodule Test do
  require My
  My.macro(IO.puts("hello"))
end
```

This generates

```
{{:.,[line: 11],[{:_aliases_,[line: 11],[:IO]},:puts]}, [line: 11],["hello"]}
Different code
```

Even though we passed `IO.puts("hello")` as a parameter, it was never executed by Elixir. Instead, it ran the code fragment we returned using quote.

Before we can write our version of if, we need one more trick—the ability to substitute existing code into a quoted block. There are two ways of doing this: by using the unquote function and with bindings.

The unquote Function

Let's get two things out of the way. First, we can use unquote only inside a quote block. Second, unquote is a silly name. It should really be something like inject_code_fragment.

Let's see why we need this. Here's a simple macro that tries to output the result of evaluating the code we pass it:

```
defmacro macro(code) do
  quote do
    IO.inspect(code)
  end
end
```

Unfortunately, when we run it, it reports an error:

```
** (CompileError).../eg2.ex:11: function code/0 undefined
```

Inside the quote block, Elixir is just parsing regular code, so the name code is inserted literally into the code fragment it returns. But we don't want that. We want Elixir to insert the evaluation of the code we pass in. And that's where we use unquote. It temporarily turns off quoting and simply injects a code fragment into the sequence of code being returned by quote.

```
defmodule My do
  defmacro macro(code) do
    quote do
      IO.inspect(unquote(code))
    end
  end
end
```

Inside the quote block, Elixir is busy parsing the code and generating its internal representation. But when it hits the unquote, it stops parsing and simply copies the code parameter into the generated code. After unquote, it goes back to regular parsing.

There's another way of thinking about this. Using unquote inside a quote is a way of deferring the execution of the unquoted code. It doesn't run when the quote block is parsed. Instead it runs when the code generated by the quote block is executed.

Or, we can think in terms of our quote-as-string-literal analogy. In this case, we can make a (slightly tenuous) case that unquote is a little like the interpolation we can do in strings. When we write "sum=#{1+2}", Elixir evaluates 1+2 and interpolates the result into the quoted string. When we write quote do: def unquote(name) do end, Elixir interpolates the contents of name into the code representation it is building as part of the list.

Expanding a List—unquote_splicing

Consider this code:

```
iex> Code.eval_quoted(quote do: [1,2,unquote([3,4])])
{[1,2,[3,4]],[]}
```

The list [3,4] is inserted, as a list, into the overall quoted list, resulting in [1,2,[3,4]].

If we instead wanted to insert just the elements of the list, we could use unquote_splicing.

```
iex> Code.eval_quoted(quote do: [1,2,unquote_splicing([3,4])])
{[1,2,3,4],[]}
```

Remembering that single-quoted strings are lists of characters, this means we can write

```
iex> Code.eval_quoted(quote do: [?a, ?= ,unquote_splicing('1234')])
{'a=1234',[]}
```

Back to Our myif Macro

We now have everything we need to implement an if macro.

```
macros/myif.ex
defmodule My do
  defmacro if(condition, clauses) do
    do_clause = Keyword.get(clauses, :do, nil)
    else_clause = Keyword.get(clauses, :else, nil)
    quote do
      case unquote(condition) do
        val when val in [false, nil] -> unquote(else_clause)
        _ -> unquote(do_clause)
      end
    end
  end
end
```

```
defmodule Test do
  require My
  My.if 1==2 do
    IO.puts "1 == 2"
  else
    IO.puts "1 != 2"
  end
end
```

It's worth studying this code.

The if macro receives a condition and a keyword list. The condition and any entries in the keyword list are passed as code fragments.

The macro extracts the do: and/or else: clauses from that list. It is then ready to generate the code for our if statement, so it opens a quote block. That block contains an Elixir case expression. This case expression has to evaluate the condition that is passed in, so it uses unquote to inject that condition's code as its parameter.

When Elixir executes this case statement, it evaluates the condition. At that point, case will match the first clause if the result is nil or false; otherwise it matches the second clause. When a clause matches (and only then), we want to execute the code that was passed in either the do: or else: values in the keyword list, so we use unquote again to inject that code into the case.

Your Turn

► *Exercise: MacrosAndCodeEvaluation-1*

Write a macro called myunless that implements the standard unless functionality. You're allowed to use the regular if expression in it.

► *Exercise: MacrosAndCodeEvaluation-2*

Write a macro called times_n that takes a single numeric argument. It should define a function called times_n in the caller's module that itself takes a single argument, and that multiplies that argument by n . So, calling times_n(3) should create a function called times_3, and calling times_3(4) should return 12. Here's an example of it in use:

```
defmodule Test do
  require Times
  Times.times_n(3)
  Times.times_n(4)
end

IO.puts Test.times_3(4)    #=> 12
IO.puts Test.times_4(5)    #=> 20
```

Using Bindings to Inject Values

Remember that there are two ways of injecting values into quoted blocks. One is unquote. The other is to use a binding. However, the two have different uses and different semantics.

A binding is simply a keyword list of variable names and their values. When we pass a binding to quote, the variables are set inside the body of that quote.

This is useful because macros are executed at compile time. This means they don't have access to values that are calculated at runtime.

Here's an example. The intent is to have a macro that defines a function that returns its own name:

```
defmacro mydef(name) do
  quote do
    def unquote(name)(), do: unquote(name)
  end
end
```

We try this out using something like `mydef(:some_name)`. Sure enough, that defines a function that, when called, returns `:some_name`.

Buoyed by our success, we try something more ambitious:

```
macros/macro_no_binding.exs
defmodule My do
  defmacro mydef(name) do
    quote do
      def unquote(name)(), do: unquote(name)
    end
  end
end

defmodule Test do
  require My
  [ :fred, :bert ] |> Enum.each(&My.mydef(&1))
end

IO.puts Test.fred
```

And we're rewarded with this:

```
macro_no_binding.exs:12: invalid syntax in def _@1()
```

At the time the macro is called, the each loop hasn't yet executed, so we have no valid name to pass it. This is where bindings come in:

```
macros/macro_binding.exs
defmodule My do
  defmacro mydef(name) do
    quote bind_quoted: [name: name] do
      def unquote(name)(), do: unquote(name)
    end
  end
end

defmodule Test do
  require My
  [ :fred, :bert ] |> Enum.each(&My.mydef(&1))
end

IO.puts Test.fred    #=> fred
```

Two things happen here. First, the binding makes the current value of name available inside the body of the quoted block. Second, the presence of the bind_quoted: option automatically defers the execution of the unquote calls in the body. This way, the methods are defined at runtime.

As its name implies, bind_quoted takes a quoted code fragment. Simple things such as tuples are the same as normal and quoted code, but for most values you probably want to quote them or use Macro.escape to ensure that your code fragment will be interpreted correctly.

Macros Are Hygienic

It is tempting to think of macros as some kind of textual substitution—a macro’s body is expanded as text and then compiled at the point of call. But that’s not the case. Consider this example:

```
macros/hygiene.ex
defmodule Scope do
  defmacro update_local(val) do
    local = "some value"
    result = quote do
      local = unquote(val)
      IO.puts "End of macro body, local = #{local}"
    end
    IO.puts "In macro definition, local = #{local}"
    result
  end
end
```

```
defmodule Test do
  require Scope

  local = 123
  Scope.update_local("cat")
  IO.puts "On return, local = #{local}"
end
```

Here's the result of running that code:

```
In macro definition, local = some value
End of macro body, local = cat
On return, local = 123
```

If the macro body were just substituted in at the point of call, both it and the module `Test` would share the same scope, and the macro would overwrite the variable `local`, so we'd see

```
In macro definition, local = some value
End of macro body, local = cat
On return, local = cat
```

But that isn't what happens. Instead, the macro definition has both its own scope and a scope during execution of the quoted macro body. Both are distinct from the scope within the `Test` module. The upshot is that macros will not clobber each other's variables or the variables of modules and functions that use them.

The `import` and `alias` functions are also locally scoped. See the documentation for `quote` for a full description. This also describes how to turn off hygiene for variables and how to control the stack trace's format if things go wrong while executing a macro.

Other Ways to Run Code Fragments

We can use the function `Code.eval_quoted` to evaluate code fragments, such as those returned by `quote`.

```
iex> fragment = quote do: IO.puts("hello")
{:[],[{:__aliases__,[alias: false],[:IO]},:puts]},[],["hello"]}
iex> Code.eval_quoted fragment
hello
{:ok,[]}
```

By default, the quoted fragment is hygienic, and so does not have access to variables outside its scope. Using `var!(name)`, we can disable this feature and allow a quoted block to access variables in the containing scope. In this case, we pass the binding to `eval_quoted` as a keyword list.

```
iex> fragment = quote do: IO.puts(var!(a))
{:ok, [], [{:_aliases_, [alias: false], [:IO]}, :puts]}, [],
  [{:var!, [context: Elixir, import: Kernel], [:a, [], Elixir]}]}
iex> Code.eval_quoted fragment, [a: "cat"]
cat
{:ok,[a: "cat"]}
```

`Code.string_to_quoted` converts a string containing code to its quoted form, and `Macro.to_string` converts a code fragment back into a string.

```
iex> fragment = Code.string_to_quoted("defmodule A do def b(c) do c+1 end end")
{:ok,{:defmodule,[line: 1],[:_aliases_,[line: 1],[A]],{do: {:def,[line: 1],[:b,[line: 1],{:c,[line: 1],nil}]}, {do: {:+,[line: 1],[:c,[line: 1],nil],1]}]}]}
iex> Macro.to_string(fragment)
"{:ok, defmodule(A) do\n  def(b(c)) do\n    c + 1\n  end\nend}"
```

We can also evaluate a string directly using `Code.eval_string`.

```
iex> Code.eval_string("[a, a*b, c]", [a: 2, b: 3, c: 4])
{[2,6,4],[a: 2, b: 3, c: 4]}
```

Macros and Operators

(This is definitely dangerous ground.)

We can override the unary and binary operators in Elixir using macros. To do so, we need to remove any existing definition first.

For example, the operator `+` (which adds two numbers) is defined in the Kernel module. To remove the Kernel definition and substitute our own, we'd need to do something like the following (which redefines addition to concatenate the string representation of the left and right arguments).

```
macros/operators.ex
defmodule Operators do
  defmacro a + b do
    quote do
      to_string(unquote(a)) <> to_string(unquote(b))
    end
  end
end

defmodule Test do
  IO.puts(123 + 456)           #=> "579"
  import Kernel, except: [+]
  import Operators
  IO.puts(123 + 456)           #=> "123456"
end

IO.puts(123 + 456)           #=> "579"
```

Note that the macro's definition is lexically scoped—the + operator is overridden from the point when we import the Operators module through the end of the module that imports it. We could also have done the import inside a single method, and the scoping would be just that method.

Digging Deeper

The Code and Macro modules contain the functions that manipulate the internal representation of code.

Check the source of the Kernel module for a list of the majority of the operator macros, along with macros for things such as def, defmodule, alias, and so on. If we look at the source code, we'll see the calling sequence for these. However, many of the bodies will be absent, as the macros are defined within the Elixir source.

Digging Ridiculously Deep

Here's the internal representation of a simple expression:

```
iex(1)> quote do: 1 + 2
{:+, [context: Elixir, import: Kernel], [1, 2]}
```

It's just a three-element tuple. In this particular case, the first element is the function (or macro), the second is housekeeping metadata, and the third is the arguments.

We know we can evaluate this code fragment using eval_quoted, and we can save typing by leaving off the metadata:

```
iex> Code.eval_quoted {:+, [], [1,2]}
{3,[]}
```

And now we can start to see the promise (and danger) of a homoiconic language (a language in which the internal representation is expressed in the language itself). Because code is just tuples and because we can manipulate those tuples, we have the ability to rewrite the definitions of existing functions. We can create new code on the fly, and we can do it in a safe way because we can control the scope of both the changes and the access to variables.

Next we'll look at *protocols*, a way of adding functionality to built-in code and of integrating our code into other people's modules.

Your Turn

► *Exercise: MacrosAndCodeEvaluation-3*

The Elixir test framework, ExUnit, uses some clever code-quoting tricks. For example, if you assert

```
assert 5 < 4
```

you'll get the error

```
Assertion with < failed
code: 5 < 4
lhs:  5
rhs:  4
```

The test code parsed the assertion parameter into the left-hand side, the operator, and the right-hand side.

The Elixir source code is on GitHub (at <https://github.com/elixir-lang/elixir>). The implementation of this is in the file elixir/lib/ex_unit/lib/ex_unit/assertions.ex. Spend some time reading this file, and work out how it implements this trick.

(Hard) Once you've done that, see if you can use the same technique to implement a function that takes an arbitrary arithmetic expression and returns a natural-language version.

```
explain do: 2 + 3 * 4
#=> multiply 3 and 4, then add 2
```

Linking Modules: Behavio(u)rs and use

When we wrote our OTP server, we wrote a module that started with the code

```
defmodule Sequence.Server do
  use GenServer
  ...
```

In this chapter we'll explore what lines such as `use GenServer` actually do, and how we can write modules that extend the capabilities of other modules that use them.

Behaviours

An Elixir behaviour is nothing more than a list of functions. A module that declares that it implements a particular behaviour must implement all of the associated functions. If it doesn't, Elixir will generate a compilation warning. You can think of a behaviour definition as being like an abstract base class in some object-oriented languages.

A behaviour is therefore a little like an *interface* in Java. A module uses it to declare that it implements a particular interface. For example, an OTP GenServer should implement a standard set of callbacks (`handle_call`, `handle_cast`, and so on). By declaring that our module implements that behaviour, we let the compiler validate that we have actually supplied the necessary interface. This reduces the chance of an unexpected runtime error.

Defining Behaviours

We define a behaviour with `@callback` definitions.

For example, the mix utility can fetch dependencies from various source-code control systems. Out of the box, it supports git and the local filesystem. However, the interface to the source-code control system (which mix abbreviates

internally as SCM) is defined using a behaviour, allowing new version-control systems to be added cleanly.

The behaviour is defined in the module Mix.Scm:

```
defmodule Mix.SCM do
  @moduledoc """
  This module provides helper functions and defines the behaviour
  required by any SCM used by Mix.
  """

  @type opts :: Keyword.t

  @doc """
  Returns a boolean if the dependency can be fetched or it is meant to
  be previously available in the filesystem.

  Local dependencies (i.e. non fetchable ones) are automatically
  recompiled every time the parent project is compiled.
  """

  @callback fetchable? :: boolean

  @doc """
  Returns a string representing the SCM. This is used when printing
  the dependency and not for inspection, so the amount of information
  should be concise and easy to spot.
  """

  @callback format(opts) :: String.t
  # and so on for 8 more callbacks
```

This module defines the interface that modules implementing the behaviour must support. It uses `@callback` to define the functions in the behaviour. But the syntax looks a little different. That's because we're using a minilanguage: Erlang type specifications. For example, the `fetchable?` function takes no parameters and returns a Boolean. The `format` function takes a parameter of type `opts` (which is defined near the top of the code to be a keyword list) and returns a string. There's more information on these type specifications [on page 361](#).

In addition to the type specification, we can include module- and function-level documentation with our behaviour definitions.

Declaring Behaviours

Now that we've defined the behaviour, we can declare that another module implements it using the `@behaviour` attribute. Here's the start of the Git implementation for `mix`:

```
defmodule Mix.SCM.Git do
  @behaviour Mix.SCM

  def fetchable? do
    true
  end

  def format(opts) do
    opts[:git]
  end

  # ...
end
```

The module defines each of the functions declared as callbacks in Mix.SCM. This module will compile cleanly. However, imagine we'd misspelled fetchable:

```
defmodule Mix.SCM.Git do
  @behaviour Mix.SCM

  ➤ def fetchible? do
    true
  end

  def format(opts) do
    opts[:git]
  end

  # ...
end
```

When we compile the module, we'd get this error:

```
git.ex:1: warning: undefined behaviour function fetchable?/0 (for behaviour Mix.SCM)
```

Behaviours give us a way of both documenting and enforcing the public functions that a module should implement.

Taking It Further

In the implementation of Mix.SCM for Git, we created a bunch of functions that implemented the behaviour. But those are unlikely to be the only functions in this module. And, unless you're intimately familiar with the Mix.SCM behaviour, you won't be able to tell the callback functions from the rest.

To remedy this, you can flag the callback functions with the `@impl` attribute. This takes a parameter: either `true` or the name of a behaviour (guess which one I prefer).

```
defmodule Mix.SCM.Git do
  @behaviour Mix.SCM

  def init(arg) do      # plain old function
    # ...
  end

  @impl Mix.SCM      # callback
  def fetchable? do
    true
  end

  @impl Mix.SCM      # callback
  def format(opts) do
    opts[:git]
  end
```

use and _using_

In one sense, `use` is a trivial function. You pass it a module along with an optional argument, and it invokes the function or macro `_using_` in that module, passing it the argument.

Yet this simple interface gives you a powerful extension facility. For example, in our unit tests we write `use ExUnit.Case` and we get the test macro and assertion support. When we write an OTP server, we write `use GenServer` and we get both a behaviour that documents the `gen_server` callback and default implementations of those callbacks.

Typically, the `_using_` callback will be implemented as a macro, as it will be used to invoke code in the original module.

Putting It Together—Tracing Method Calls

Let's work through a larger example. We want to write a module called `Tracer`. If we use `Tracer` in another module, entry and exit tracing will be added to any subsequently defined function. For example, given the following:

```
use/tracer.ex
defmodule Test do
  use Tracer
  def puts_sum_three(a,b,c), do: IO.inspect(a+b+c)
  def add_list(list),         do: Enum.reduce(list, 0, &(&1+&2))
end

Test.puts_sum_three(1,2,3)
Test.add_list([5,6,7,8])
```

we'd get this output:

```
=> call    puts_sum_three(1, 2, 3)
6
<== returns 6
=> call    add_list([5,6,7,8])
<== returns 26
```

My approach to writing this kind of code is to start by exploring what we have to work with, and then to generalize. The goal is to metaprogram as little as possible.

It looks as if we have to override the `def` macro, which is defined in Kernel. So let's do that and see what gets passed to `def` when we define a method.

```
use/tracer1.ex
defmodule Tracer do
  defmacro def(definition, do: _content) do
    IO.inspect definition
    quote do: {}
  end
end

defmodule Test do
  import Kernel, except: [def: 2]
  import Tracer, only:  [def: 2]

  def puts_sum_three(a,b,c), do: IO.inspect(a+b+c)
  def add_list(list),         do: Enum.reduce(list, 0, &(&1+&2))
end

Test.puts_sum_three(1,2,3)
Test.add_list([5,6,7,8])
```

This outputs

```
{:puts_sum_three, [line: 12],
 [{:a, [line: 12], nil}, {:b, [line: 12], nil}, {:c, [line: 12], nil}]}
{:add_list, [line: 13], [{:list, [line: 13], nil}]}
** (UndefinedFunctionError) undefined function: Test.puts_sum_three/3
```

The definition part of each method is a three-element tuple. The first element is the name, the second is the line on which it is defined, and the third is a list of the parameters, where each parameter is itself a tuple.

We also get an error: `puts_sum_three` is undefined. That's not surprising—we intercepted the `def` that defined it, and we didn't create the function.

You may be wondering about the form of the macro definition: `defmacro def(definition, do: _content)....` The `do:` in the parameters is not special syntax: it's a pattern match on the block passed as the function body, which is a keyword list.

You may also be wondering if we have affected the built-in `Kernel.def` macro. The answer is no. We've created another macro, also called `def`, which is defined in the scope of the `Tracer` module. In our `Test` module we tell Elixir not to import the `Kernel` version of `def` but instead to import the version from `Tracer`. Shortly, we'll make use of the fact that the original `Kernel` implementation is unaffected.

Let's see if we can define a real function given this information. That turns out to be surprisingly easy. We already have the two arguments passed to `def`. All we have to do is pass them on.

```
use/tracer2.ex
defmodule Tracer do
  defmacro def(definition, do: content) do
    quote do
      Kernel.def(unquote(definition)) do
        unquote(content)
      end
    end
  end
end

defmodule Test do
  import Kernel, except: [def: 2]
  import Tracer, only: [def: 2]

  def puts_sum_three(a,b,c), do: IO.inspect(a+b+c)
  def add_list(list),       do: Enum.reduce(list, 0, &(&1+&2))
end

Test.puts_sum_three(1,2,3)
Test.add_list([5,6,7,8])
```

When we run this, we see 6, the output from `puts_sum_three`.

Now it's time to add some tracing.

```
use/tracer3.ex
defmodule Tracer do
  def dump_args(args) do
    args |> Enum.map(&inspect/1) |> Enum.join(", ")
  end

  def dump_defn(name, args) do
    "#{name} (#{dump_args(args)})"
  end

  defmacro def(definition={name, _, args}, do: content) do
    quote do
```

```

Kernel.def(unquote(definition)) do
  IO.puts "==> call:  #{Tracer.dump_defn(unquote(name), unquote(args))}"
  result = unquote(content)
  IO.puts "<== result: #{result}"
  result
end
end
end
end

defmodule Test do
  import Kernel, except: [def: 2]
  import Tracer, only:  [def: 2]

  def puts_sum_three(a,b,c), do: IO.inspect(a+b+c)
  def add_list(list),         do: Enum.reduce(list, 0, &(&1+&2))
end

Test.puts_sum_three(1,2,3)
Test.add_list([5,6,7,8])

```

Looking good:

```

==> call:  puts_sum_three(1, 2, 3)
6
<== result: 6
==> call:  add_list([5,6,7,8])
<== result: 26

```

Let's package our Tracer module so clients only have to add `use Tracer` to their own modules. We'll implement the `_using_` callback. The tricky part here is differentiating between the two modules: Tracer and the module that uses it.

```

use/tracer4.ex
defmodule Tracer do

  def dump_args(args) do
    args |> Enum.map(&inspect/1) |> Enum.join(", ")
  end

  def dump_defn(name, args) do
    "##{name}({#{dump_args(args)}})"
  end

  defmacro def(definition={name,_,args}, do: content) do
    quote do
      Kernel.def(unquote(definition)) do
        IO.puts "==> call:  #{Tracer.dump_defn(unquote(name), unquote(args))}"
        result = unquote(content)
        IO.puts "<== result: #{result}"
        result
      end
    end
  end
end

```

```

defmacro __using__(_opts) do
  quote do
    import Kernel, except: [def: 2]
    import unquote(_MODULE_), only: [def: 2]
  end
end
end

defmodule Test do
  use Tracer
  def puts_sum_three(a,b,c), do: IO.inspect(a+b+c)
  def add_list(list),           do: Enum.reduce(list, 0, &(&1+&2))
end

Test.puts_sum_three(1,2,3)
Test.add_list([5,6,7,8])

```

Use use

Elixir behaviours are fantastic—they let you easily inject functionality into modules you write. And they’re not just for library creators—use them in your own code to cut down on duplication and boilerplate.

Although behaviours let you add to modules that you are writing, you sometimes need to extend the functionality of modules written by others—code that you can’t change. Fortunately, Elixir comes with *protocols*, the subject of the next chapter.

Your Turn

► [Exercise: LinkingModules-BehavioursAndUse-1](#)

In the body of the def macro, there’s a quote block that defines the actual method. It contains

```

IO.puts "==> call:  #{Tracer.dump_dfn(unquote(name), unquote(args))}"
result = unquote(content)
IO.puts "<== result: #{result}"

```

Why does the first call to puts have to unquote the values in its interpolation but the second call does not?

► [Exercise: LinkingModules-BehavioursAndUse-2](#)

The built-in module IO.ANSI defines functions that represent ANSI escape sequences. You can use it to build output than will display (for example) colors and bold, inverse, or underlined text (assuming the terminal supports it).

```
iex> import IO.ANSI
iex> IO.puts [ "Hello, ", white(), green_background(), "world!" ]
Hello, world!
```

Explore the module, and use it to colorize our tracing's output.

Why does passing a list of strings to IO.puts work?

► [Exercise: LinkingModules-BehavioursAndUse-3](#)

(Hard) Try adding a method definition with a guard clause to the Test module. You'll find that the tracing no longer works.

- Find out why.
- See if you can fix it.

Protocols—Polymorphic Functions

We have used the `inspect` function many times in this book. It returns a printable representation of any value as a binary (which is what we hard-core folks call strings).

But stop and think for a minute. Just how can Elixir, which doesn't have objects, know what to call to do the conversion to a binary? You can pass `inspect` anything, and Elixir somehow makes sense of it.

It *could* be done using guard clauses:

```
def inspect(value) when is_atom(value), do: ...
def inspect(value) when is_binary(value), do: ...
: : :
```

But there's a better way.

Elixir has the concept of *protocols*. A protocol is a little like the behaviours we saw in the previous chapter in that it defines the functions that must be provided to achieve something. But a behaviour is internal to a module—the module implements the behaviour. Protocols are different—you can place a protocol's implementation completely outside the module. This means you can extend modules' functionality without having to add code to them—in fact, you can extend the functionality even if you don't have the modules' source code.

Defining a Protocol

Protocol definitions are very similar to basic module definitions. They can contain module- and function-level documentation (`@moduledoc` and `@doc`), and they will contain one or more function definitions. However, these functions will not have bodies—they are there simply to declare the interface that the protocol requires.

For example, here is the definition of the `Inspect` protocol:

```
defprotocol Inspect do
  @fallback_to_any true
  def inspect(thing, opts)
end
```

Just like a module, the protocol defines one or more functions. But we implement the code separately.

Implementing a Protocol

The `defimpl` macro lets you give Elixir the implementation of a protocol for one or more types. The code that follows is the implementation of the `Inspect` protocol for PIDs and references.

```
defimpl Inspect, for: PID do
  def inspect(pid, _opts) do
    "#PID" <> IO.iodata_to_binary(pid_to_list(pid))
  end
end

defimpl Inspect, for: Reference do
  def inspect(ref, _opts) do
    '#Ref' ++ rest = :erlang.ref_to_list(ref)
    "#Reference" <> IO.iodata_to_binary(rest)
  end
end
```

Finally, the `Kernel` module implements `inspect`, which calls `Inspect.inspect` with its parameter. This means that when you call `inspect(self)`, it becomes a call to `Inspect.inspect(self)`. And because `self` is a PID, this in turn resolves to something like "`#PID<0.25.0>`".

Behind the scenes, `defimpl` puts the implementation for each protocol-and-type combination into a separate module. The protocol for `Inspect` for the `PID` type is in the module `Inspect.PID`. And because you can recompile modules, you can change the implementation of functions accessed via protocols.

```
iex> inspect self
"#PID<0.25.0>"
iex> defimpl Inspect, for: PID do
...>   def inspect(pid, _) do
...>     "#Process: " <> IO.iodata_to_binary(:erlang.pid_to_list(pid)) <> "!!"
...>   end
...> end
iex:3: redefining module Inspect.PID
{:module, Inspect.PID, <>70,79....}
iex> inspect self
"#Process: <0.25.0>!!"
```

The Available Types

You can define implementations for one or more of the following types:

Any	Atom	BitString	Float	Function
Integer	List	Map	PID	Port
Record	Reference	Tuple		

The type BitString is used in place of Binary.

The type Any is a catchall, allowing you to match an implementation with any type. Just as with function definitions, you'll want to put the implementations for specific types before an implementation for Any.

You can list multiple types on a single defimpl. For example, the following protocol can be called to determine whether a type is a collection:

```
protocols/is_collection.exs
defprotocol Collection do
  @fallback_to_any true
  def is_collection?(value)
end

defimpl Collection, for: [List, Tuple, BitString, Map] do
  def is_collection?(_), do: true
end

defimpl Collection, for: Any do
  def is_collection?(_), do: false
end

Enum.each [ 1, 1.0, [1,2], {1,2}, %{}, "cat" ], fn value ->
  IO.puts "#{inspect value}: #{Collection.is_collection?(value)}"
end
```

We write defimpl stanzas for the collection types: List, Tuple, BitString, and Map. But what about the other types? To handle those, we use the special type Any in a second defimpl. If we use Any, though, we also have to add an annotation to the protocol definition. That's what the @fallback_to_any line does.

This produces

```
1: false
1.0: false
[1,2]: true
{1,2}: true
%{}: true
"cat": true
```

Your Turn

► *Exercise: Protocols-1*

A basic Caesar cypher consists of shifting the letters in a message by a fixed offset. For an offset of 1, for example, a will become b, b will become c, and z will become a. If the offset is 13, we have the ROT13 algorithm.

Lists and binaries can be *stringlike*. Write a Caesar protocol that applies to both. It would include two functions: `encrypt(string, shift)` and `rot13(string)`.

► *Exercise: Protocols-2*

Using a list of words in your language, write a program to look for words where the result of calling `rot13(word)` is also a word in the list. (For various English word lists, look at <http://wordlist.sourceforge.net/>. The SCOWL collection looks promising, as it already has words divided by size.)

Protocols and Structs

Elixir doesn't have classes, but (perhaps surprisingly) it does have user-defined types. It pulls off this magic using structs and a few conventions.

Let's play with a simple struct. Here's the definition:

```
protocols/basic.exs
defmodule Blob do
  defstruct content: nil
end
```

And here we use it in IEx:

```
iex> c "basic.exs"
[Blob]
iex> b = %Blob{content: 123}
%Blob{content: 123}
iex> inspect b
"%Blob{content: 123}"
```

It looks for all the world as if we've created some new type, the blob. But that's only because Elixir is hiding something from us. By default, `inspect` recognizes structs. If we turn this off using the `structs: false` option, `inspect` reveals the true nature of our blob value:

```
iex> inspect b, structs: false
"%{__struct__: Blob, content: 123}"
```

A struct value is actually just a map with the key `_struct_` referencing the struct's module (`Blob` in this case) and the remaining elements containing the keys and values for this instance. The `inspect` implementation for maps checks

for this—if you ask it to inspect a map containing a key `_struct_` that references a module, it displays it as a struct.

Many built-in types in Elixir are represented as structs internally. It's instructive to try creating values and inspecting them with `structs: false`.

Built-in Protocols

Elixir comes with the following protocols:

- Enumerable and Collectable
- Inspect
- List.Chars
- String.Chars

To play with these, let's work with MIDI files.

A MIDI file consists of a sequence of variable-length frames. Each frame contains a four-character type, a 32-bit length, and then *length* bytes of data.¹

We'll define a module that represents the MIDI file content as a struct, because the struct lets us use it with protocols. The file also defines a submodule for the individual frame structure.

```
protocols/midi.exs
defmodule Midi do
  defstruct(content: <>>>)

  defmodule Frame do
    defstruct(
      type: "xxxx",
      length: 0,
      data: <>>
    )

    def to_binary(%Midi.Frame{type: type, length: length, data: data}) do
      <<
        type::binary-4,
        length::integer-32,
        data::binary
      >>
    end
  end

  def from_file(name) do
    %Midi{content: File.read!(name)}
  end
end
```

1. <https://www.csie.ntu.edu.tw/~r92092/ref/midi/>

Built-in Protocols: Enumerable and Collectable

The Enumerable protocol is the basis of all the functions in the `Enum` module—any type implementing it can be used as a collection argument to `Enum` functions.

We're going to implement Enumerable for our `Midi` structure, so we'll need to wrap the implementation in something like this:

```
defimpl Enumerable, for: Midi do
  # ...
end
```

The protocol is defined in terms of four functions:

```
defprotocol Enumerable do
  def count(collection)
  def member?(collection, value)
  def reduce(collection, acc, fun)
  def slice(collection)
end
```

`count` returns the number of elements in the collection, `member?` is truthy if the collection contains `value`, and `reduce` applies the given function to successive values in the collection and the accumulator; the value it reduces becomes the next accumulator. Finally, `slice` is used to create a subset of a collection. Perhaps surprisingly, all the `Enum` functions can be defined in terms of these four.

However, life isn't that simple. Maybe you're using `Enum.find` to find a value in a large collection. Once you've found it, you want to halt the iteration—continuing is pointless. Similarly, you may want to suspend an iteration and resume it sometime later. These two features become particularly important when we talk about streams, which let you enumerate a collection lazily.

We'll start with the most difficult function to implement, `Enumerable.reduce/3`. It is worth reading the documentation for it before we start:

```
iex> h Enumerable.reduce
      def reduce(enumerable, acc, fun)
      @spec reduce(t(), acc(), reducer()) :: result()
```

Reduces the enumerable into an element.

Most of the operations in `Enum` are implemented in terms of `reduce`. This function should apply the given `t:reducer/0` function to each item in the enumerable and proceed as expected by the returned accumulator.

See the documentation of the types `t:result/0` and `t:acc/0` for more information.

Examples

As an example, here is the implementation of reduce for lists:

```
def reduce(_,      {:halt, acc}, _fun),
  do: {:halted, acc}

def reduce(list,    {:suspend, acc}, fun),
  do: {:suspended, acc, &reduce(list, &l, fun)}

def reduce([],     {:cont, acc}, _fun),
  do: {:done, acc}

def reduce([h | t], {:cont, acc}, fun),
  do: reduce(t, fun.(h, acc), fun)
```

The first two function heads do housekeeping: they handle the cases where the enumeration has been halted or suspended. Here are the versions for our MIDI enumerator:

```
protocols/midi.exs
def _reduce(_content, {:halt, acc}, _fun) do
  {:halted, acc}
end

def _reduce(content, {:suspend, acc}, fun) do
  {:suspended, acc, &_reduce(content, &l, fun)}
end
```

The next two function heads do the actual iteration. In the list example in the documentation, you'll see the typical pattern: check for the end condition ([]]) and the recursive step [h|t].

We'll do the same with our MIDI file, but we'll use binaries instead of doing list pattern matching:

```
protocols/midi.exs
def _reduce(_content = "", {:cont, acc}, _fun) do
  {:done, acc}
end

def _reduce(<<
            type::binary-4,
            length::integer-32,
            data::binary-size(length),
            rest::binary
          >>,
          {:cont, acc},
          fun
        ) do
  frame = %Midi.Frame{type: type, length: length, data: data}
  _reduce(rest, fun.(frame, acc), fun)
end
```

See how we split out the binary content of a frame, then wrap it into a `Midi.Frame` struct before passing it back. This means that folks who use our MIDI module will only see these frame structures, and not the raw data.

Before we try this, there's one little tweak we have to make. You may have noticed that all my functions were named `_reduce`, with a leading underscore. That's because they need to work on the content of the MIDI file, and not on the structure that wraps that content. We have a single function head that implements the actual `reduce` function, and that then forwards the call on to `_reduce`:

```
protocols/midi.exs
def reduce(%Midi{content: content}, state, fun) do
  _reduce(content, state, fun)
end
```

At this point we have enough code to try it out. I've included a MIDI file in the `code/protocols` directory for you to play with (courtesy of [midiworld.com](#)).²

```
iex midi.exs
warning: function count/1 required by protocol Enumerable is not
         implemented (in module Enumerable.Midi) midi.exs:21
warning: function member?/2 required by protocol Enumerable is not
         implemented (in module Enumerable.Midi) midi.exs:21
warning: function slice/1 required by protocol Enumerable is not
         implemented (in module Enumerable.Midi) midi.exs:21
Interactive Elixir (1.6.0-rc.0) - press Ctrl+C to exit (type h() ENTER for help)
iex> midi = Midi.from_file("dueling-banjos.mid")
%Midi{
  content: <<77, 84, 104, 100, 0, 0, 0, 6, 0, 1, 0, 8, 0, 120, 77, 84, 114, 107,
          0, 0, 0, 66, 0, 255, 3, 14, 68, 117, 101, 108, 105, 110, 103, 32, 66, 97,
          110, 106, 111, 115, 0, 255, 3, 11, 68, 101, 108, 105, 118, ...>>
}
iex> Enum.take(midi, 2)
[
  %Midi.Frame{data: <<0, 1, 0, 8, 0, 120>>, length: 6, type: "MThd"},
  %Midi.Frame{
    data: <<0, 255, 3, 14, 68, 117, 101, 108, 105, 110, 103, 32, 66, 97, 110,
          106, 111, 115, 0, 255, 3, 11, 68, 101, 108, 105, 118, 101, 114, 97, 110,
          99, 101, 0, 255, 88, 4, 4, 2, 24, 8, 0, 255, 89, 2, 0, 0, ...>>,
    length: 66,
    type: "MTrk"
  }
]
```

2. [midiworld.com](#)

First, see how we get warnings because we haven't yet implemented the full Enumerable protocol. Normally this would worry me, but I happen to know that we won't be using those functions just yet.

Next, look at how we called Enum.take/2 and got back two Midi.Frame structures. That's because take/2 is defined in terms of reduce/3, and we supplied an implementation of it.

So far, so good. Let's implement count/1 next.

If the collection is countable, the count function returns a tuple containing {`:ok`, `count`}. If it isn't countable (perhaps it is being read one piece at a time from an external source), count should return {`:ok`, `_MODULE_`}.

In our case, we have the whole MIDI file available in memory, and we have a way to traverse it using reduce, so counting it easy:

```
protocols/midi.exs
def count(midi = %Midi{}) do
  frame_count = Enum.reduce(midi, 0, fn (_, count) -> count+1 end)
  { :ok, frame_count }
end
```

Let's try it:

```
iex> r Enumerable.Midi
warning: redefining module Midi (current version defined in memory)
  midi.exs:2

warning: redefining module Midi.Frame (current version defined in memory)
  midi.exs:6

warning: redefining module Enumerable.Midi (current version defined in memory)
  midi.exs:21

warning: function member?/2 required by protocol Enumerable is not
  implemented (in module Enumerable.Midi) midi.exs:21

warning: function slice/1 required by protocol Enumerable is not
  implemented (in module Enumerable.Midi) midi.exs:21

{:reloaded, Enumerable.Midi, [Midi.Frame, Midi, Enumerable.Midi]}
iex> Enum.count midi
9
```

On to member? and slice.

Technically, both of these can be implemented using reduce. But the Elixir team recognized that some types of collection have more direct ways to test membership and partition elements. For example, if you implement a set using a map, then testing to see if a key is present can be done in constant

time. If you have an array-like structure with fixed-length elements, you can split it into two in (almost) constant time.

So the implementations of both `member?` and `slice` depend on the characteristics of your collection. In our case, we don't currently have a fast way of testing for membership, nor do we have a fast way to slice a MIDI file into two. In both cases, we return an error tuple. This doesn't actually cause an error for the user; it just tells `Enumerable` to fall back to a naive algorithm.

```
protocols/midi.exs
def member?(%Midi{}, %Midi.Frame{}) do
  { :error, __MODULE__ }
end

def slice(%Midi{}) do
  { :error, __MODULE__ }
end
```

And with that, we're done. Our `Midi` type is enumerable, and you can use every function in `Enum` on it.

This gives us the ability to treat a MIDI stream as a collection of MIDI frames. But how do we assemble frames back into a MIDI stream? That's what we'll address next.

Collectable

We've already seen `Enum.into/2`. It takes something that's enumerable and creates a new collection from it:

```
iex> 1..4 |> Enum.into([])
[1, 2, 3, 4]
iex> [ {1, 2}, {"a", "b"} ] |> Enum.into(%{})
%{1 => 2, "a" => "b"}
```

The target of `Enum.into` must implement the `Collectable` protocol. This defines a single function, somewhat confusingly also called `into`. This function returns a two-element tuple. The first element is the initial value of the target collection. The second is a function to be called to add each item to the collection. (If this reminds you of the second and third parameters passed to `Enum.reduce`, that's because in a way `into` is the opposite of `reduce`.)

Let's look at the code first:

```
protocols/midi.exs
defimpl Collectable, for: Midi do
  use Bitwise
```

```
def into(%Midi{content: content}) do
  {
    content,
    fn
      acc, {:cont, frame = %Midi.Frame{}} ->
        acc <> Midi.Frame.to_binary(frame)

      acc, :done ->
        %Midi{content: acc}

      _, :halt ->
        :ok
    end
  }
end
```

It works like this:

- `Enum.into` calls the `into` function for `Midi`, passing it the target value—`Midi{content: content}` in this case.
- `Midi.into` returns a tuple. The first element is the current content of the target. This acts as the initial value for an accumulator. The second element of the tuple is a function.
- `Enum.into` then calls this function, passing it the accumulator and a command. If the command is `:done`, the iteration over the collection being injected into the MIDI stream has finished, so we return a new `Midi` structure using the accumulator as a value. If the command is `:halt`, the iteration has terminated early and nothing needs to be done.
- The real work is done when the function is passed the `{:cont, frame}` command. Here is where the `Collectable` appends the binary representation of the next frame to the accumulator.

We can call it in IEx:

```
iex> list = Enum.to_list(midi)
[
  %Midi.Frame{data: <<0, 1, 0, 8, 0, 120>>, length: 6, type: "MThd"},
  %Midi.Frame{
    data: <<0, 255, 3, 14, 68, 117, 101, 108, 105, 110, 103, 32, 66, 97, 110,
           106, 111, 115, 0, 255, 3, 11, 68, 101, 108, 105, 118, 101, 114, 97, 110,
           99, 101, 0, 255, 88, 4, 4, 2, 24, 8, 0, 255, 89, 2, 0, 0, ...>>,
    length: 66,
    type: "MTrk"
  },
  ...
]
```

```
iex> new_midi = Enum.into(list, %Midi{})
%Midi{
  content: <<77, 84, 104, 100, 0, 0, 0, 6, 0, 1, 0, 8, 0, 120, 77, 84, 114, 107,
  0, 0, 0, 66, 0, 255, 3, 14, 68, 117, 101, 108, 105, 110, 103, 32, 66, 97,
  110, 106, 111, 115, 0, 255, 3, 11, 68, 101, 108, 105, 118, ...>>
}
iex> new_midi == midi
true
iex> Enum.take(new_midi, 1)
[%Midi.Frame{data: <<0, 1, 0, 8, 0, 120>>, length: 6, type: "MThd"}]
```

Because the `into` function uses the initial value of the target collection, we can use it to append to a MIDI stream:

```
iex> midi2 = %Midi{}
%Midi{content: ""}
iex> midi2 = Enum.take(midi, 1) |> Enum.into(midi2)
%Midi{content: <<77, 84, 104, 100, 0, 0, 0, 6, 0, 1, 0, 8, 0, 120>>}
iex> midi2 = [Enum.at(midi, 3)] |> Enum.into(midi2)
%Midi{
  content: <<77, 84, 104, 100, 0, 0, 0, 6, 0, 1, 0, 8, 0, 120, 77, 84, 114, 107,
  0, 0, 8, 34, 0, 255, 33, 1, 0, 0, 193, 25, 0, 177, 7, 127, 0, 10, 100, 0,
  64, 0, 134, 24, 145, 43, 99, 22, 43, 0, 15, ...>>
}
iex> Enum.count(midi2)
2
```

Remember the Big Picture

If you think all this enumerable/collectable stuff is complicated—well, you’re correct. It is. In part that’s because these conventions allow all enumerable values to be used both eagerly and lazily. And when you’re dealing with big (or even infinite) collections, this is a big deal.

Built-in Protocols: Inspect

This is the protocol that is used to inspect a value. The rule is simple—if you can return a representation that is a valid Elixir literal, do so. Otherwise, prefix the representation with `#Typename`.

We *could* just delegate the `inspect` function to the Elixir default. (That’s what we’ve been doing so far.) But we can do better. Not surprisingly, we do that by implementing the `Inspect` protocol. We’ll do it for both the overall `Midi` type and for the individual `Midi.Frame`s.

```
protocols/midi_inspect.exs
defimpl Inspect, for: Midi do
  def inspect(%Midi{content: <<>>}, _opts) do
    "#Midi[<<empty>>]"
  end
```

```

def inspect(midi = %Midi{}, _opts) do
  content =
    Enum.map(midi, fn frame-> Kernel.inspect(frame) end)
    |> Enum.join("\n")
    "#Midi[\n#{content}\n]"
 $\langle\!\langle$ 
  end
 $\rangle\!\rangle$ 

defimpl Inspect, for: Midi.Frame do
  def inspect(%Midi.Frame{type: "MThd",
    length: 6,
    data: <<
      format::integer-16,
      tracks::integer-16,
      division::bits-16
    >>},
    _opts) do
    beats = decode(division)
    "#Midi.Header{Midi format: #{format}, tracks: #{tracks}, timing: #{beats}}"
   $\langle\!\langle$ 
  end

  def inspect(%Midi.Frame{type: "MTrk", length: length, data: data}, _opts) do
    "#Midi.Track{length: #{length}, data: #{Kernel.inspect(data)}}"
   $\langle\!\langle$ 
  end

  defp decode(<< 0::1, beats::15>>) do
    "J = #{beats}"
   $\langle\!\langle$ 
  end

  defp decode(<< 1::1, fps::7, beats::8>>) do
    "#{-fps} fps, #{beats}/frame"
   $\langle\!\langle$ 
  end
 $\rangle\!\rangle$ 
 $\langle\!\langle$ 
  end
 $\rangle\!\rangle$ 

```

Run the code in IEx:

```

iex> midi = Midi.from_file "dueling-banjos.mid"
#Midi[
#Midi.Header{Midi format: 1, tracks: 8, timing: J = 120}
#Midi.Track{length: 66, data: <<0, 255, 3, 14, 68, 117, 101, ...>>
  ...
#Midi.Track{length: 6291, data: <<0, 255, 33, 1, 0, 0, 185, ... >>}
#Midi.Track{length: 9, data: <<0, 255, 33, 1, 0, 0, 255, 47, 0>>}
]

```

I added a little bit of decoding for the header frame, but just treated the track frames as binary. You could extend this to do a full decode of each track frame too.

There's a wrinkle here. If you pass structs: false to IO.inspect (or Kernel.inspect), it never calls our inspect function. Instead, it formats it as a struct.

Better Formatting with Algebra Documents

The formatting of our MIDI stream leaves a little to be desired: there's no indentation or reasonable line wrapping.

To fix this, we use a feature called *algebra documents*. An algebra document is a tree structure that represents some data you'd like to pretty-print.³ Your job is to create the structure based on the data you want to inspect, and Elixir will then find a nice way to display it.

I'd like the inspect string to show the nesting of data, and to wrap long lines honoring that nesting.

We do this by having our inspect function return an algebra document rather than a string. In that document, we indicate places where breaks are allowed (but not required) and we show how the nesting works:

```
protocols/midi_algebra.exs
defimpl Inspect, for: Midi do
  import Inspect.Algebra

  def inspect(%Midi{content: <>>}, _opts) do
    "#Midi[«empty»]"
  end

  def inspect(midi = %Midi{}, opts) do
    open      = color("#Midi[", :map, opts)
    close     = color("]", :map, opts)
    separator = color(",", :map, opts)

    container_doc(
      open,
      Enum.to_list(midi),
      close,
      %Inspect.Opts{limit: 4},
      fn frame, _opts -> Inspect.Midi.Frame.inspect(frame, opts) end,
      separator: separator,
      break: :strict
    )
  end
end

defimpl Inspect, for: Midi.Frame do
  import Inspect.Algebra

  def inspect(
    %Midi.Frame{type: "MThd",
               length: 6,
               data: <<
```

3. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.34.2200>

```

        format::integer-16,
        tracks::integer-16,
        division::bits-16
    >>
},
opts)
do
concat(
[
nest(
concat(
[
color("#{Midi.Header}", :map, opts),
break(""),
"Midi format: #{format},",
break(" "),
"tracks: #{tracks},",
break(" "),
"timing: #{decode(division)}",
]
),
2
),
break(""),
color("}", :map, opts)
]
)
end

def inspect(%Midi.Frame{type: "MTrk", length: length, data: data}, opts) do
open      = color("#{Midi.Track}", :map, opts)
close     = color("}", :map, opts)
separator = color(",", :map, opts)
content = [
  length: length,
  data:   data
]
container_doc(
  open,
  content,
  close,
  %Inspect.Opts{limit: 15},
  fn {key, value}, opts ->
    key = color("#{key}:", :atom, opts)
    concat(key, concat(" ", to_doc(value, opts)))
  end,
  separator: separator,
  break: :strict
)
end

```

```

defp decode(<< 0::1, beats::15 >>) do
  "J = #{beats}"
end

defp decode(<< 1::1, fps::7, beats::8 >>) do
  "#{-fps} fps, #{beats}/frame"
end

defp decode(x) do
  raise inspect x
end
end

```

On a narrow terminal window, we get this output:

```

iex> Midi.from_file "dueling-banjos.mid"
#Midi[
  #Midi.Header{
    Midi format: 1,
    tracks: 8,
    timing: J = 120
  },
  #Midi.Track{
    length: 66,
    data: <<0, 255, 3, 14, 68, 117, 101, 108, 105,
          110, 103, 32, 66, ...>>
  },
  #Midi.Track{
    length: 1319,
    data: <<0, 255, 33, 1, 0, 0, 192, 105, 0, 176, 7,
          127, 0, ...>>
  },
  #Midi.Track{
    length: 2082,
    data: <<0, 255, 33, 1, 0, 0, 193, 25, 0, 177, 7,
          127, 0, ...>>
  },
  ...
]

```

For more information, see the documentation for `Inspect.Algebra`.

Built-in Protocols: List.Chars and String.Chars

The `List.Chars` protocol is used by `Kernel.to_charlist` to convert a value into a list of characters (think *single-quoted string*).

The `String.Chars` protocol is used to convert a value to a string (binary, or *double-quoted string*). This is the protocol used for string interpolation.

The protocols are implemented identically, except `List.Chars` requires that you write a `to_charlist` function, and `String.Chars` requires you write `to_string`.

Although we *could* implement a `String.Chars.to_string` for our `Midi` struct, it probably wouldn't make much sense. What would you interpolate into the string?

Protocols Are Polymorphism

When you want to write a function that behaves differently depending on the type of its arguments, you're looking at a polymorphic function. Elixir protocols give you a tidy and controlled way to implement this. Whether you're integrating your types into the existing Elixir library or creating a new library with a flexible interface, protocols let you package the behaviour in a well-documented and disciplined way. And with that, we're almost done. But when you write about a language, there are always little details that don't seem to fit anywhere. That's why the next chapter is full of odds and ends.

Your Turn

► *Exercise: Protocols-3*

Collections that implement the `Enumerable` protocol define `count`, `member?`, `reduce`, and `slice` functions. The `Enum` module uses these to implement methods such as `each`, `filter`, and `map`.

Implement your own versions of `each`, `filter`, and `map` in terms of `reduce`.

► *Exercise: Protocols-4*

In many cases, `inspect` will return a valid Elixir literal for the value being inspected. Update the `inspect` function for structs so that it returns valid Elixir code to construct a new struct equal to the value being inspected.

More Cool Stuff

Elixir is packed with features that make coding a joy. This chapter contains a smattering of them.

Writing Your Own Sigils

You know by now that you can create strings and regular-expression literals using sigils:

```
string = ~s{now is the time}
regex  = ~r{..h..}
```

Have you ever wished you could extend these sigils to add your own specific literal types? You can.

When you write a sigil such as `~s{...}`, Elixir converts it into a call to the function `sigil_s`. It passes the function two values. The first is the string between the delimiters. The second is a list containing any lowercase letters that immediately follow the closing delimiter. (This second parameter is used to pick up any options you pass to a regex literal, such as `~r/cat/if`.)

Here's the implementation of a sigil `~l` that takes a multiline string and returns a list containing each line as a separate string. We know that `~l...` is converted into a call to `sigil_l`, so we just write a simple function in the `LineSigil` module.

```
odds/line_sigil.exs
defmodule LineSigil do
  @doc """
  Implement the `~l` sigil, which takes a string containing
  multiple lines and returns a list of those lines.
```

```
## Example usage

iex> import LineSigil
nil
iex> ~l"""
...> one
...> two
...> three
...> |"""
[ "one", "two", "three" ]
"""

def sigil_l(lines, _opts) do
  lines |> String.trim_trailing |> String.split("\n")
end
end
```

We can play with this in a separate module:

```
odds/line_sigil.exs
defmodule Example do
  import LineSigil

  def lines do
    ~l"""
    line 1
    line 2
    and another line in #__MODULE__
    """
  end
end

IO.inspect Example.lines
```

This produces ["line 1","line 2","and another line in Elixir.Example"].

Because we import the `sigil_l` function inside the example module, the `~l` sigil is lexically scoped to this module. Note also that Elixir performs interpolation before passing the string to our method. That's because we used a lowercase `l`. If our sigil were `~L{...}` and the function were renamed `sigil_L`, no interpolation would be performed.

The predefined sigil functions are `sigil_C`, `sigil_c`, `sigil_R`, `sigil_r`, `sigil_S`, `sigil_s`, `sigil_W`, and `sigil_w`. If you want to override one of these, you'll need to explicitly import the `Kernel` module and use an `except` clause to exclude it.

In this example, we used the heredoc syntax (`"""`). This passes our function a multiline string with leading spaces removed. Sigil options are not supported with heredocs, so we'll switch to a regular literal syntax to play with them.

Picking Up the Options

Let's write a sigil that enables us to specify color constants. If we say `~c{red}`, we'll get `0xff0000`, the RGB representation. We'll also support the option `h` to return an HSB value, so `~c{red}h` will be `{0,100,100}`.

Here's the code:

```
odds/color.exs
defmodule ColorSigil do

  @color_map [
    rgb: [ red: 0xff0000, green: 0x00ff00, blue: 0x0000ff, # ...
           ],
    hsb: [ red: {0,100,100}, green: {120,100,100}, blue: {240,100,100}
           ]
  ]

  def sigil_c(color_name, []), do: _c(color_name, :rgb)
  def sigil_c(color_name, 'r'), do: _c(color_name, :rgb)
  def sigil_c(color_name, 'h'), do: _c(color_name, :hsb)

  defp _c(color_name, color_space) do
    @color_map[color_space][String.to_atom(color_name)]
  end

  defmacro __using__(_opts) do
    quote do
      import Kernel, except: [sigil_c: 2]
      import unquote(__MODULE__), only: [sigil_c: 2]
    end
  end
end

defmodule Example do
  use ColorSigil

  def rgb, do: IO.inspect ~c{red}
  def hsb, do: IO.inspect ~c{red}h
end

Example.rgb  #=> 16711680  (== 0xff0000)
Example.hsb  #=> {0,100,100}
```

The three clauses for the `sigil_c` function let us select the colorspace to use based on the option passed. As the single-quoted string `'r'` is actually represented by the list `[?r]`, we can use the string literal to pattern-match the options parameter.

Because I'm overriding a built-in sigil, I decided to implement a `_using_` macro that automatically removes the Kernel version and adds our own (but only in the lexical scope that calls `use` on our module).

The fact that we can write our own sigils is liberating. But misuse could lead to some pretty impenetrable code.

Your Turn

► *Exercise: MoreCoolStuff-1*

Write a sigil `~v` that parses multiple lines of comma-separated data, returning a list where each element is a row of data and each row is a list of values. Don't worry about quoting—just assume each field is separated by a comma.

For example

```
csv = ~v"""
1,2,3
cat,dog
"""
```

would generate `[["1","2","3"], ["cat","dog"]]`.

► *Exercise: MoreCoolStuff-2*

The function `Float.parse` converts leading characters of a string to a float, returning either a tuple containing the value and the rest of the string, or the atom `:error`.

Update your CSV sigil so that numbers are automatically converted:

```
csv = ~v"""
1,2,3.14
cat,dog
"""
```

should generate `[[1.0,2.0,3.14], ["cat","dog"]]`.

► *Exercise: MoreCoolStuff-3*

(Hard) Sometimes the first line of a CSV file is a list of the column names. Update your code to support this, and return the values in each row as a keyword list, using the column names as the keys. Here's an example:

```
csv = ~v"""
Item,Qty,Price
Teddy bear,4,34.95
Milk,1,2.99
Battery,6,8.00
"""
```

would generate

```
[  
  [Item: "Teddy bear", Qty: 4, Price: 34.95],  
  [Item: "Milk", Qty: 1, Price: 2.99],  
  [Item: "Battery", Qty: 6, Price: 8.00]  
]
```

Multi-app Umbrella Projects

It is unfortunate that Erlang chose to call self-contained bundles of code *apps*. In many ways, they are closer to being shared libraries. And as your projects grow, you may find yourself wanting to split your code into multiple libraries, or apps. Fortunately, mix makes this painless.

To illustrate the process, we'll create a simple Elixir evaluator. Given a set of input lines, it will return the result of evaluating each. This will be one app.

To test it, we'll need to pass in lists of lines. We've already written a trivial `~l` sigil that creates lists of lines for us, so we'll make that sigil code into a separate application.

Elixir calls these multi-app projects *umbrella projects*.

Create an Umbrella Project

We use `mix new` to create an umbrella project, passing it the `--umbrella` option.

```
$ mix new --umbrella eval  
* creating README.md  
* creating mix.exs  
* creating apps
```

Compared to a normal `mix new` project, the umbrella is pretty lightweight—just a `mix.exs` file and an `apps` directory.

Create the Subprojects

Subprojects are stored in the `apps` directory. There's nothing special about them—they are simply regular projects created using `mix new`. Let's create our two projects now:

```
$ cd eval/apps  
$ mix new line_sigil  
* creating README.md  
... and so on  
$ mix new evaluator  
* creating README.md  
... and so on  
* creating test/evaluator_test.exs
```

At this point we can try out our umbrella project. Go back to the overall project directory and try `mix compile`.

```
$ cd ..
$ mix compile
==> evaluator
Compiled lib/evaluator.ex
Generated evaluator app
==> line_sigil
Compiled lib/line_sigil.ex
Generated line_sigil app
```

Now we have an umbrella project containing two regular projects. Because there's nothing special about the subprojects, you can use all the regular `mix` commands in them. At the top level, though, you can build all the subprojects as a unit.

Making the Subproject Decision

The fact that subprojects are just regular `mix` projects means you don't have to worry about whether to start a new project using an umbrella. Simply start as a simple project. If you later discover the need for an umbrella project, create it and move your existing simple project into the `apps` directory.

The LineSigil Project

This project is trivial—just copy the `LineSigil` module from the previous section into `apps/line_sigil/lib/line_sigil.ex`. Verify it builds by running `mix compile`—in either the top-level directory or the `line_sigil` directory.

The Evaluator Project

The evaluator takes a list of strings containing Elixir expressions and evaluates them. It returns a list containing the expressions intermixed with the value of each. For example, given

```
a = 3
b = 4
a + b
```

our code will return

```
code> a = 3
value> 3
code> b = 4
value> 4
code> a + b
value> 7
```

We'll use `Code.eval_string` to execute the Elixir expressions. To have the values of variables pass from one expression to the next, we'll also need to explicitly maintain the current binding.

Here's the code:

```
odds/eval/apps/evaluator/lib/evaluator.ex
defmodule Evaluator do
  def eval(list_of_expressions) do
    { result, _final_binding } =
      Enum.reduce(list_of_expressions,
        {_result = [], _binding = binding()},
        &evaluate_with_binding/2)
    Enum.reverse result
  end
  defp evaluate_with_binding(expression, { result, binding }) do
    { next_result, new_binding } = Code.eval_string(expression, binding)
    { [ "value> #{next_result}", "code>  #{expression}" | result ], new_binding }
  end
end
```

Linking the Subprojects

Now we need to test our evaluator. It makes sense to use our `~l` sigil to create lists of expressions, so let's write our tests that way.

```
odds/eval/apps/evaluator/test/evaluator_test.exs
defmodule EvaluatorTest do
  use ExUnit.Case
  import LineSigil

  test "evaluates a basic expression" do
    input = ~l"""
    1 + 2
    """
    output = ~l"""
    code> 1 + 2
    value> 3
    """
    run_test input, output
  end

  test "variables are propagated" do
    input = ~l"""
    a = 123
    a + 1
    """
    output = ~l"""
    code> a = 123
    """
    run_test input, output
  end
end
```

```

value> 123
code> a + 1
value> 124
"""

run_test input, output
end

defp run_test(lines, output) do
  assert output == Evaluator.eval(lines)
end
end

```

But if we simply run this in the `apps/evaluator` directory, Elixir won't be able to find the `LineSigil` module, as we don't have a dependency. Instead, just run the tests from the top-level directory (the one containing the overall umbrella project). Mix will automatically load all the child apps for you.

```

$ mix test
▶ ==> evaluator
..
Finished in 0.02 seconds
2 tests, 0 failures
Randomized with seed 334706
▶ ==> line_sigil
..
Finished in 0.04 seconds
1 doctest, 1 test, 0 failures
Randomized with seed 334706

```

The first stanza of test output is for the `evaluator` tests, and the second is for `line_sigil`.

But Wait! There's More!

We've reached the end of our Elixir exploration.

This book was never intended to be exhaustive. Instead, it is intended to hit the highlights, and to give you enough information to start coding apps in Elixir yourself.

That means there's a lot more to learn, both about the language and about how to write great apps in it.

And I think that's fun. Enjoy!

Exceptions: raise and try, catch and throw

Elixir (like Erlang) takes the view that errors should normally be fatal to the processes in which they occur. A typical Elixir application's design involves many processes, which means the effects of an error will be localized. A supervisor will detect the failing process, and the restart will be handled at that level.

For that reason, you won't find much exception-handling code in Elixir programs. Exceptions are raised, but you rarely catch them.

Use exceptions for things that are exceptional—things that should never happen.

Exceptions do exist. This appendix is an overview of how to generate them and how to catch them when they occur.

Raising an Exception

You can raise an exception using the `raise` function. At its simplest, you pass it a string and it generates an exception of type `RuntimeError`.

```
iex> raise "Giving up"
** (RuntimeError) Giving up
  erl_eval.erl:572: :erl_eval.do_apply/6
```

You can also pass the type of the exception, along with other optional fields. All exceptions implement at least the `message` field.

```
iex> raise RuntimeError
** (RuntimeError) runtime error
  erl_eval.erl:572: :erl_eval.do_apply/6
iex> raise RuntimeError, message: "override message"
** (RuntimeError) override message
  erl_eval.erl:572: :erl_eval.do_apply/6
```

You can intercept exceptions using the `try` function. It takes a block of code to execute, and optional `rescue`, `catch`, and `after` clauses.

The `rescue` and `catch` clauses look a bit like the body of a `case` function—they take patterns and code to execute if the pattern matches. The subject of the pattern is the exception that was raised.

Here's an example of exception handling in action. We define a module that has a public function, `start`. It calls a different helper function depending on the value of its parameter. With 0, it runs smoothly. With 1, 2, or 3, it causes the VM to raise an error, which we catch and report.

```
exceptions/exception.ex
defmodule Boom do
  def start(n) do
    try do
      raise_error(n)
    rescue
      [ FunctionClauseError, RuntimeError ] ->
        IO.puts "no function match or runtime error"
      error in [ArithmetricError] ->
        IO.inspect error
        IO.puts "Uh-oh! Arithmetric error"
        reraise "too late, we're doomed", System.stacktrace
      other_errors ->
        IO.puts "Disaster! #{inspect other_errors}"
    after
      IO.puts "DONE!"
    end
  end

  defp raise_error(0) do
    IO.puts "No error"
  end

  defp raise_error(val = 1) do
    IO.puts "About to divide by zero"
    1 / (val-1)
  end

  defp raise_error(2) do
    IO.puts "About to call a function that doesn't exist"
    raise_error(99)
  end

  defp raise_error(3) do
    IO.puts "About to try creating a directory with no permission"
    File.mkdir!("/not_allowed")
  end
end
```

We define three different exception patterns. The first matches one of the two exceptions, `FunctionClauseError` or `RuntimeError`. The second matches an `ArithmaticError` and stores the exception value in the variable `error`. And the last clause catches *any* exception into the variable `other_error`.

We also include an `after` clause. This will always run at the end of the `try` function, regardless of whether an exception was raised.

Finally, look at the handling of `ArithmaticError`. As well as reporting the error, we call `reraise`. This raises the current exception, but lets us add a message. We also pass in the stack trace (which is actually the stack trace at the point the original exception was raised). Let's see all this in IEx:

```
iex> c("exception.ex")
[Boom]
iex> Boom.start 1
About to divide by zero
%ArithmaticError{}
Uh-oh! Arithmatic error
DONE!
** (RuntimeError) too late, we're doomed
  exception.ex:26: Boom.raise_error/1
  exception.ex:5: Boom.start/1

iex> Boom.start 2
About to call a function that doesn't exist
no function match or runtime error
DONE!
:ok

iex> Boom.start 3
About to try creating a directory with no permission
Disaster! %File.Error{action: "make directory", path: "/not_allowed",
                     reason: :eacces}

DONE!
:ok
```

catch, exit, and throw

Elixir code (and the underlying Erlang libraries) can raise a second kind of error. These are generated when a process calls `error`, `exit`, or `throw`. All three take a parameter, which is available to the catch handler.

Here's an example:

```
exceptions/catch.ex
defmodule Catch do

  def start(n) do
    try do
      incite(n)
    catch
      :exit, code  -> "Exited with code #{inspect code}"
      :throw, value -> "throw called with #{inspect value}"
      what, value   -> "Caught #{inspect what} with #{inspect value}"
    end
  end

  defp incite(1) do
    exit(:something_bad_happened)
  end

  defp incite(2) do
    throw {:animal, "wombat"}
  end

  defp incite(3) do
    :erlang.error "Oh no!"
  end
end
```

Calling the start function with 1, 2, or 3 will cause an exit, a throw, or an error to be thrown. Just to illustrate wildcard pattern matching, we handle the last case by matching any type into the variable what.

```
iex> c("catch.ex")
[Catch]
iex> Catch.start 1
"Exited with code :something_bad_happened"
iex> Catch.start 2
"throw called with {:animal,\\"wombat\\"}"
iex> Catch.start 3
"Caught :error with \"Oh no!\""
```

Defining Your Own Exceptions

Exceptions in Elixir are basically records. You can define your own exceptions by creating a module. Inside it, use defexception to define the various fields in the exception, along with their default values. Because you're creating a module, you can also add functions—often these are used to format the exception's fields into meaningful messages.

Say we're writing a library to talk to a Microsoft Kinect controller. It might want to raise an exception on various kinds of communication errors. Some of these are permanent, but others are likely to be transient and can be retried.

We'll define our exception with its (required) message field and an additional can_retry field. We'll also add a function that formats these two fields into a nice message.

```
exceptions/defexception.ex
defmodule KinectProtocolError do
  defexception message: "Kinect protocol error",
               can_retry: false

  def full_message(me) do
    "Kinect failed: #{me.message}, retriable: #{me.can_retry}"
  end
end
```

Users of our library could write code like this:

```
exceptions/defexception.ex
try do
  talk_to_kinect()
rescue
  error in [KinectProtocolError] ->
    IO.puts KinectProtocolError.full_message(error)
    if error.can_retry, do: schedule_retry()
end
```

If an exception gets raised, the code handles it and possibly retries:

```
Kinect failed: usb unplugged, retriable: true
Retrying in 10 seconds
```

Now Ignore This Appendix

The Elixir source code for the mix utility contains no exception handlers. The Elixir compiler itself contains a total of five (but it is doing some pretty funky things).

If you find yourself defining new exceptions, ask if you should be isolating the code in a separate process instead. After all, if it can go wrong, wouldn't you want to isolate it?

Type Specifications and Type Checking

When we looked at [@callback, on page 319](#), we saw that we defined callbacks in terms of their parameter types and return values. For example, we might write

```
@callback parse(uri_info :: URI.Info.t) :: URI.Info.t  
@callback default_port() :: integer
```

The terms `URI.Info.t` and `integer` are examples of type specifications. And, as José Valim pointed out to me, the cool thing is that they are implemented (by Yurii Rashkovskii) directly in the Elixir language itself—no special parsing is involved. This is a great illustration of the power of Elixir metaprogramming.

In this appendix we'll discuss how to specify types in Elixir. But before we do, there's another question to address: *Why bother?*

When Specifications Are Used

Elixir type specifications come from Erlang. It is very common to see Erlang code where every exported (public) function is preceded by a `-spec` line. This is metadata that gives type information. The following code comes from the Elixir parser (which is [currently] written in Erlang). It says the `return_error` function takes two parameters, an `integer` and any type, and never returns.

```
-spec return_error(integer(), any()) -> no_return().  
return_error(Line, Message) ->  
    throw({error, {Line, ?MODULE, Message}}).
```

One of the reasons the Erlang folks do this is to document their code. You can read it inline while reading the source, and you can also read it in the pages created by their documentation tool.

The other reason is that they have tools such as *dialyzer* that perform static analysis of Erlang code and report on some kinds of type mismatches.¹

These same benefits can apply to Elixir code. We have the `@spec` module attribute for documenting a function's type specification; in IEx we have the `s` helper for displaying specifications and the `t` helper for showing user-defined types. You can also run Erlang tools such as dialyzer on compiled Elixir `.beam` files.

However, type specifications are not currently in wide use in the Elixir world. Whether you use them is a matter of personal taste.

Specifying a Type

A type is simply a subset of all possible values in a language. For example, the type `integer` means all the possible integer values, but excludes lists, binaries, PIDs, and so on.

The basic types in Elixir are as follows: `any`, `atom`, `float`, `fun`, `integer`, `list`, `map`, `maybe_improper_list`, `none`, `pid`, `port`, `reference`, `struct`, and `tuple`.

The type `any` (and its alias, `_`) is the set of all values, and `none` is the empty set.

A literal atom or integer is the set containing just that value.

The value `nil` can be represented as `nil`.

Collection Types

A list is represented as `[type]`, where `type` is any of the basic or combined types. This notation does not signify a list of one element—it simply says that elements of the list will be of the given type. If you want to specify a nonempty list, use `[type, ...]`. As a convenience, the type `list` is an alias for `[any]`.

Binaries are represented using this syntax:

`<< >>`

An empty binary (size 0).

`<< _ :: size >>`

A sequence of `size` bits. This is called a *bitstring*.

`<< _ :: size * unit_size >>`

A sequence of `size` units, where each unit is `unit_size` bits long.

In the last two instances, `size` can be specified as `_`, in which case the binary has an arbitrary number of bits/units.

1. <http://www.erlang.org/doc/man/dialyzer.html>

The predefined type `bitstring` is equivalent to `<<:_>>`, an arbitrarily sized sequence of bits. Similarly, `binary` is defined as `<<:_*8>>`, an arbitrary sequence of 8-bit bytes.

Tuples are represented as `{ type, type, ... }` or using the type tuple, so both `{atom, integer}` and `tuple(atom, integer)` represent a tuple whose first element is an atom and whose second element is an integer.

Combining Types

The range operator `(..)` can be used with literal integers to create a type representing that range. The three built-in types, `non_neg_integer`, `pos_integer`, and `neg_integer`, represent integers that are greater than or equal to, greater than, or less than zero, respectively.

The union operator `(|)` indicates that the acceptable values are the unions of its arguments.

Parentheses may be used to group terms in a type specification.

Structures

As structures are basically maps, you could just use the `map` type for them, but doing so throws away a lot of useful information. Instead, I recommend that you define a specific type for each struct:

```
defmodule LineItem do
  defstruct sku: "", quantity: 1
  @type t :: %LineItem{sku: String.t, quantity: integer}
end
```

You can then reference this type as `LineItem.t`.

Anonymous Functions

Anonymous functions are specified using `(head -> return_type)`.

The `head` specifies the arity and possibly the types of the function parameters. Use `...` to mean an arbitrary number of arbitrarily typed arguments, or a list of types, in which case the number of types is the function's arity.

```
(... -> integer)           # Arbitrary parameters; returns an integer
(list(integer) -> integer) # Takes a list of integers and returns an integer
(() -> String.t)          # Takes no parameters and returns an Elixir string
(integer, atom -> list(atom)) # Takes an integer and an atom and returns
                                # a list of atoms
```

You can put parentheses around the head if you find it clearer:

```
( atom, float -> list )
( (atom, float) -> list )
(list(integer) -> integer)
((list(integer)) -> integer)
```

Handling Truthy Values

The type `as_boolean(T)` says that the actual value matched will be of type `T`, but the function that uses the value will treat it as a *truthy* value (anything other than `nil` or `false` is considered true). Thus the specification for the Elixir function `Enum.count` is

```
@spec count(t, (element -> as_boolean(term))) :: non_neg_integer
```

Some Examples

`integer | float`

Any number (Elixir has an alias for this).

`[{atom, any}]`

`list(atom, any)`

A list of key/value pairs. The two forms are the same.

`non_neg_integer | {:error, String.t}`

An integer greater than or equal to zero, or a tuple containing the atom `:error` and a string.

`(integer, atom -> { :pair, atom, integer })`

An anonymous function that takes an integer and an atom and returns a tuple containing the atom `:pair`, an atom, and an integer.

`<<_::_*4>>`

A sequence of 4-bit nibbles.

Defining New Types

The attribute `@type` can be used to define new types.

```
@type type_name :: type_specification
```

Elixir uses this to predefine some built-in types and aliases. Here are just some of them.

```
@type term      :: any
@type binary    :: <<_::_*8>>
@type bitstring :: <<_::_*1>>
@type boolean   :: false | true
@type byte      :: 0..255
@type char      :: 0..0xffff
```

```

@type charlist :: [ char ]
@type list :: [ any ]
@type list(t) :: [ t ]
@type number :: integer | float
@type module :: atom
@type mfa :: {module, atom, byte}
@type node :: atom
@type nonempty_charlist :: [ char ]
@type timeout :: :infinity | non_neg_integer
@type no_return :: none

```

As the list(t) entry shows, you can parameterize the types in a new definition. Simply use one or more identifiers as parameters on the left side, and use these identifiers where you'd otherwise use type names on the right. Then when you use the newly defined type, pass in actual types for each of these parameters:

```

@type variant(type_name, type) :: { :variant, type_name, type}
@spec create_string_tuple(:string, String.t) :: variant(:string, String.t)

```

As well as @type, Elixir has the @typep and @opaque module attributes. They have the same syntax as @type, and do basically the same thing. The difference is in the visibility of the result.

@typep defines a type that is local to the module that contains it—the type is private. @opaque defines a type whose name may be known outside the module but whose definition is not.

Specs for Functions and Callbacks

The @spec specifies a function's parameter count, types, and return-value type. It can appear anywhere in a module that defines the function, but by convention it sits immediately before the function definition, following any function documentation.

We've already seen the syntax:

```
@spec function_name( param1_type, ... ) :: return_type
```

Let's see some examples. These come from the built-in Dict module.

```

Line 1 @type key :: any
2 @type value :: any
3 @type keys :: [ key ]
4 @type t :: tuple | list    # `t` is the type of the collection
5
6 @spec values(t) :: [value]
7 @spec size(t) :: non_neg_integer
8 @spec has_key?(t, key) :: boolean
9 @spec update(t, key, value, (value -> value)) :: t

```

Line 6

`values` takes a collection (tuple or list) and returns a list of values (any).

Line 7

`size` takes a collection and returns an integer (≥ 0).

Line 8

`has_key?` takes a collection and a key, and returns true or false.

Line 9

`update` takes a collection, a key, a value, and a function that maps a value to a value. It returns a (new) collection.

For functions with multiple heads (or those that have default values), you can specify multiple `@spec` attributes. Here's an example from the `Enum` module:

```
@spec at(t, index) :: element | nil
@spec at(t, index, default) :: element | default
def at(collection, n, default \\ nil) when n >= 0 do
  ...
end
```

The `Enum` module also has many examples of the use of `as_boolean`:

```
@spec filter(t, (element -> as_boolean(term))) :: list
def filter(collection, fun) when is_list(collection) do
  ...
end
```

This says `filter` takes something enumerable and a function. That function maps an element to a term (which is an alias for any), and the `filter` function treats that value as being truthy. `filter` returns a list.

For more information on Elixir support for type specifications, see the guide.²

Using Dialyzer

Dialyzer analyzes code that runs on the Erlang VM, looking for potential errors. To use it with Elixir, we have to compile our source into `.beam` files and make sure that the `debug_info` compiler option is set (which it is when running `mix` in the default, development mode). Let's see how to do that by creating a trivial project with two source files.

```
$ mix new simple
...
$ cd simple
```

2. <http://elixir-lang.org/getting-started/typespecs-and-behaviours.html#types-and-specs>

Inside the project, let's create a simple function. Being lazy, I haven't implemented the body yet.

```
defmodule Simple do
  @type atom_list :: list(atom)
  @spec count_atoms(atom_list) :: non_neg_integer
  def count_atoms(list) do
    # ...
  end
end
```

Let's run dialyzer on our code. To make life simple, we'll use the dialixir library to add a dialyzer task to mix:

```
typespecs/simple/mix.exs
defp deps do
  [
    { :dialyxir, "~> 0.5", only: [:dev], runtime: false }
  ]
end
```

Fetch the library and build our project:

```
$ mix deps.get
$ mix compile
```

Now we're ready to analyze our code. However, the first time we do this, dialyzer needs to construct a massive data structure containing all the types and APIs in both Erlang and Elixir. This lets it check not just our code, but also that our code is interacting correctly with the rest of the world. Building this data structure is slow: expect it to take 10 to 20 minutes! But once done, it won't be repeated.

```
mix dialyzer
Compiling 2 files (.ex)
warning: variable "list" is unused
  lib/simple.ex:8

Generated simple app
Checking PLT...
[:compiler, :elixir, :kernel, :stdlib]
Finding suitable PLTs
Looking up modules in dialyxir_erlang-20.2.2_elixir-1.6.0-rc.0_deps-dev.plt
...
Checking 391 modules in dialyxir_erlang-20.2.2_elixir-1.6.0-rc.0_deps-dev.plt
Adding 48 modules to dialyxir_erlang-20.2.2_elixir-1.6.0-rc.0_deps-dev.plt

Starting Dialyzer
dialyzer args: [
  check_plt: false,
```

```

init_plt: '/Users/dave/Work/Bookshelf/titles/elixir16/Book/code/typespecs
/simple/_build/dev/dialyxir_erlang-20.2.2_elixir-1.6.0-rc.0_deps-dev.plt',
  files_rec: ['/Users/dave/Work/Bookshelf/titles/elixir16/Book/code/typespe
cs/simple/_build/dev/lib/simple/ebin'],
  warnings: [:unknown]
]
done in 0m1.18s
lib/simple.ex:7: Invalid type specification for function
  'Elixir.Simple':count_atoms/1. The success typing is (_) -> 'nil'
done (warnings were emitted)

```

Ouch! Let's run it again:

```

mix dialyzer --no-check
Starting Dialyzer
dialyzer args: [
  .
  .
]
done in 0m1.4s

lib/simple.ex:7: Invalid type specification for function
  'Elixir.Simple':count_atoms/1. The success typing is (_) -> 'nil'
done (warnings were emitted)

```

Those last three lines are the important ones. They're complaining that the typespec for `count_atoms` doesn't agree with the implementation. The *success typing* (think of this as the *actual type*)³ returns `nil`, but the spec says it is a nonnegative integer. Dialyzer has caught our stubbed-out body.

Let's fix that:

```

typespecs/simple/lib/simple.ex
defmodule Simple do
  @type atom_list :: list(atom)
  @spec count_atoms(atom_list) :: non_neg_integer
  def count_atoms(list) do
    length list
  end
end

```

and run dialyzer again:

```

$ mix dialyzer
Compiling 1 file (.ex)
Checking PLT...

done in 0m1.34s
done (passed successfully)

```

3. http://www.it.uu.se/research/group/hipe/papers/succ_types.pdf

Let's add a second module that calls our `count_atoms` function:

```
typespecs/simple/lib/simple/client.ex
defmodule Client do
  @spec other_function() :: non_neg_integer
  def other_function do
    Simple.count_atoms [1, 2, 3]
  end
end
```

Compile and dialyze:

```
19:18:53> mix dialyzer
Compiling 1 file (.ex)
Generated simple app
Checking PLT...
done in 0m1.37s

lib/simple/client.ex:3: Function other_function/0 has no local return
lib/simple/client.ex:4: The call 'Elixir.Simple':count_atoms([1 | 2
  | 3,...]) breaks the contract (atom_list()) -> non_neg_integer()
done (warnings were emitted)
```

That's pretty cool. Dialyzer noticed that we called `count_atoms` with a list of integers, but it is specified to receive a list of atoms. It also decided this would raise an error, so the function would never return (that's the *no local return* warning). Let's fix that:

```
defmodule Client do
  @spec other_function() :: non_neg_integer
  def other_function do
    Simple.count_atoms [:a, :b, :c]
  end
end

$ mix dialyzer
Compiling 1 file (.ex)
done in 0m1.03s
done (passed successfully)
```

And so it goes....

Dialyzer and Type Inference

In this appendix, we've shown dialyzer working with type specs that we added to our functions. But it also does a credible job with unannotated code. This is because dialyzer knows the types of the built-in functions (remember the long wait the first time we ran it) and can infer (some of) your function types from this. Here's a simple example:

```
defmodule NoSpecs do
  def length_plus_n(list, n) do
    length(list) + n
  end
  def call_it do
    length_plus_n(2, 1)
  end
end
```

Compile this, and run dialyzer:

```
$ mix dialyzer
...
done in 0m1.28s

lib/nospecs.ex:5: Function call_it/0 has no local return
lib/nospecs.ex:6: The call 'Elixir.NoSpecs':length_plus_n(1,2) will
  never return since it differs in the 1st argument from the success
  typing arguments: ([any()],number())
done (warnings were emitted)
```

Here it noticed that the `length_plus_n` function called `length` on its first parameter, and `length` requires a list as an argument. This means `length_plus_n` also needs a list argument, and so it complains.

What happens if we change the call to `length_plus_n([:a, :b], :c)`?

```
defmodule NoSpecs do
  def length_plus_n(list, n) do
    length(list) + n
  end
  def call_it do
    length_plus_n([1, 2], :c)
  end
end

$ mix dialyzer
done in 0m1.29s

lib/nospecs.ex:5: Function call_it/0 has no local return
lib/nospecs.ex:6: The call 'Elixir.NoSpecs':length_plus_n([1, 2],'c')
  will never return since it differs in the 2nd argument from the
  success typing arguments: ([any()],number())
done (warnings were emitted)
```

This is even cooler. It knows that `+` (which is implemented as a function) takes two numeric arguments. When we pass an atom as the second parameter, dialyzer recognizes that this makes no sense, and complains. But look at the error. It isn't complaining about the addition. Instead, it has assigned a default typespec to our function, based on its analysis of what we call inside that function.

This is *success typing*. Dialyzer attempts to infer the most permissive types that are compatible with the code—it assumes the code is correct until it finds a contradiction. This makes it a powerful tool, as it can make assumptions as it runs.

Does that mean you don't need @spec attributes? That's your call. Try it with and without. Often, adding a @spec will further constrain a function's type signature. We saw this with our count_of_atoms function, where the spec made it explicit that we expected a list of atoms as an argument.

Ultimately, dialyzer is a tool, not a test of your coding chops. Use it as such, but don't waste time adding specs to get a gold star.

Bibliography

- [Arm13] Joe Armstrong. *Programming Erlang (2nd edition)*. The Pragmatic Bookshelf, Raleigh, NC, 2nd, 2013.

Index

SYMBOLS

- ! (exclamation mark)
 - in names, 34
 - relaxed Boolean operator, 36
 - atoms, 26
 - raising exceptions, 139
 - != strict inequality operator, 35
 - != value inequality operator, 35
- "" (double quotes)
 - atoms, 26
 - enclosing strings, 73, 117, 124–130, 307
- """ (double quotes, triple), heredocs, 118, 120, 348
- # (hash sign), preceding comments, 35
- #{} (string interpolation), 44
- \$ (dollar sign), terminal prompt, 4
- & (ampersand)
 - function capture operator, 48–50, 64, 76
 - &(&1) identity function, 295
 - && relaxed Boolean operator, 36
- " (single quotes), enclosing strings, 73, 117, 121–123, 307
- ''' (single quotes, triple), heredocs, 118, 120
- () (parentheses)
 - enclosing function arguments, 41
- named functions, 54, 64
- nesting functions, 46
- pipelines, 64
- using with, 39
- * (asterisk), multiplication operator, 36
- <> (angle brackets), <>> enclosing binaries, 362
- () (parentheses), grouping type specifications, 363
- .. (range operator), 363
- [] (square brackets), as nil value, 362
- _ (underscore), special variable, 362
- | (vertical bar), join operator, 363
- + (plus sign)
 - addition operator, 36
 - ++ list concatenation, 30, 36, 81
- , (comma), formatter tool and, 192
- (minus sign)
 - subtraction operator, 36
 - list difference, 30
- > operator, in functions, 41
- . (dot notation)
 - anonymous function calls, 41
 - maps, 32
 - module names and function calls, 68
 - nested modules, 66
 - nesting dictionary structures, 90
 - structs, 88
- ... (ellipsis), iex prompt, 4
- / (forward slash), division operator, 36
- : (colon)
 - atoms, 26
 - preceding Erlang functions, 69
- < (left angle bracket), less-than operator, 35
- <- operator, pattern matching with with, 38
- <>>
 - comprehension generators, 113
 - enclosing binaries, 32
- <= less-than-or-equal-to operator, 35, 102
- <>
 - binaries concatenation, 36
 - enclosing binaries, 123
- = (equal sign)
 - match operator, 15, 20
 - == strict equality operator, 35
 - == value equality operator, 35
- > (right angle bracket), greater-than operator, 35
- >= greater-than-or-equal-to operator, 35
- ? (question mark)
 - atoms, 26
 - in names, 34
- @ (at sign)
 - atoms, 26
 - module attributes, 67

- [] (square brackets)
 - accessing maps, 31
 - c helper return value, 10
 - enclosing lists, 30
 - list creation, 16
 - \ (backslash), preceding escape sequences, 117
 - \| (backslash, double), default parameters, 60
 - ^ (caret), pin operator
 - in function parameters, 48
 - in pattern matching, 19, 87
 - _ (underscore)
 - in decimals, 26
 - list processing, 74–75
 - in names, 34
 - special variable, 18, 43, 74–75
 - { } (braces)
 - enclosing maps, 31, 88
 - string interpolation (#{}), 44, 117
 - tuples, 28
 - | (vertical bar)
 - > pipe operator, 63, 150
 - // relaxed Boolean operator, 36
 - join operator, 78
 - pipe character, lists, 71
 - pipelines, 2
 - |> pipe operator, 63, 150
 - ~ (tilde), preceding sigils, 118
-
- A**
- Access module, 93
 - accessors, 91
 - accumulator, processes, 203–205
 - actor model of concurrency, 197
 - actors, defined, 197
 - add_worker, 267
 - addition operator, 36
 - after clause, try function, 356–357
 - after pattern, process messages, 201
 - agents
 - about, 293
 - anagram example, 297–300
 - distributing, 299
-
- B**
- backslash (), preceding escape sequences, 117
 - backslash, double (\|), default parameters, 60
 - :bad_call error, 238
 - :bad_cast error, 238
 - Beam, *see* Erlang VM
 - .beam files, 281
 - @behaviour attribute, 320
 - behaviours, *see also* OTP behaviours
 - exercises, 326
 - starting dependencies as, 154
 - terminology, 155, 277
 - umbrella projects, 351–354
 - visualizations, 189

- linking modules with, 319–326
- names, 34
- source-code control methods example, 319–322
- big**, binary qualifier, 130
- binaries**
 - about, 32, 123, 362
 - binary functions, 125–130
 - binary integers, 26
 - bit extraction, 124
 - comprehensions, 112
 - concatenating, 36
 - converting to integers, 148
 - double-quoted strings as, 124
 - exercises, 130–131
 - modifiers for, 123
 - overriding binary operator, 315
 - pattern matching, 130
 - qualifiers for, 130
 - splitting, 130
 - storing in binaries, 124
 - using, 123–132
- binary type**, 130
- bind_quoted**: option, 313
- binding**
 - explicitly maintaining, 353
 - injecting values with, 312, 314
 - macros, 312, 314
 - variables in nested functions, 46
 - variables in pattern matching, 18, 87
- bit_size**, 123
- bits type**, 130
- bitstring type**, 130, 331, 362–364
- bitstrings**
 - about, 130
 - comprehensions, 112
 - protocols, 331
- Boolean operators**, 35, 59
- Boolean values**, 35, 364, 366
- boundary conditions**, 179
- braces {}**
 - enclosing maps, 31, 88
 - string interpolation (#{}), 44, 117
 - tuples, 28
- brackets**, *see* square brackets
- break!**, 172
- breakpoints**, 169–173
- byte_size**, 123
- bytes type**, 130
- C**
- c helper function**, 10, 53
- ?c notation**, 122
- ~c sigil**, 119
- ~c sigil**, 119
- caching**, Fibonacci server example, 215
- Caesar cypher**, 331
- Calendar library**, 34
- Calendar module**, 33
- call**, 233–234, 238
- @callback**, 319
- callbacks**
 - defining behaviours, 319
 - flagging with @impl attribute, 321
 - GenServer default, 238
 - tracing example, 322–326
- callers**, listing in mix, 186
- can_retry**, 358
- Capistrano**, 285
- capitalization**, *see* case
- capitalize**, 125
- capture_jo**, 162
- cards**, dealing, 102
- caret (^)**, pin operator
 - in function parameters, 48
 - in pattern matching, 19, 87
- case**
 - agents, 296
 - binary functions, 125, 129
 - conventions for, 24, 34
 - Erlang documentation, 69
 - filtering and pattern-matching lists, 79
 - pattern matching, 27
 - strings, 24, 125, 129
- case statement**, 137, 311
- cast**, 234, 238
- catch clause**, try function, 356
- catch function**, 357
- character lists**
 - exercises, 123
- sigils**, 119
- terminology**, 120–121
- using**, 121–123
- characters**, *see also* character lists
 - converting values into, 344
 - escaping, 117
 - integer code, 122
 - sigils, 119
- Chars protocol**, 344
- check_all**, 180
- child specifications**, 255, 295
- child_list**, 248
- child_spec**, 255, 295
- children**
 - child specifications, 255, 295
 - names, 267
 - supervision, 248, 295
- classes**, 1
- CLI module**, *see* GitHub project example
- closures**, 46
- code**
 - binding code fragments, 313
 - for this book, 9
 - converting strings, 315
 - Elixir source code, 316
 - evaluating code fragments, 314
 - formatting, 35, 190–193, 320
 - hot upgrades, 234, 282, 287
 - injection with macros, 304–310
 - recompiling source code, 234
 - releasing, 282–292
 - using representation as code in macros, 307–310
 - working with large code examples, 44
- code blocks**, do block for, 36
- code_change**, 239, 290
- codepoints**, 125
- Collectable API**, 111
- Collectable protocol**, 110, 338–340

- collections, 99–115, *see also* binaries; lists; maps; tuples
 about, 362
 cards, dealing, 102
 Collectable protocol, 110, 338–340
 collection types, 25, 28–33
 comparing, 100
 comprehensions, 111–114
 concatenating, 100
 converting into lists, 100
 creating, 100
 enumerating, 85
 exercises, 102
 filtering, 100
 folding elements, 101
 inserting elements into, 110
 iterating with for, 85
 joining, 101
 mapping, 100
 merging, 101
 predicate operations, 101
 processing with Enum, 99–103
 processing with Stream, 99, 103–110
 protocols, 331
 selecting specific elements, 100
 sorting, 100, 102
 splitting, 100
- colon (:)
 atoms, 26
 preceding Erlang functions, 69
- color
 color constants example, 349–350
 disabling colorization, 9
 iex options, 8
- columns, formatting table, 160
- comma (,), formatter tool and, 192
- command line
 executable for projects, 163–164
 parsing, 145–149
- command pipelines,
 see pipelines
- comments
 formatter tool and, 192
 regular expressions, 27
- syntax for, 35
 tests based on, 173–177
- comparison operators, 35, 59
- compile, 280
- compiling
 about, 9
 libraries, 153
 logging options, 164
 messages bug, 204
 modules, 53
 OTP applications, 280
 recompiling, 150, 234
- components
 applications as, 155, 242, 277
 sequence server example, 242–244
- composability, streams, 103
- comprehensions, 111–114
- concat, 100
- concatenating
 binaries, 36
 collections, 100
 lists, 30, 36, 81
- concurrency, 197–217, *see also* nodes; OTP; processes
 about, 11, 197
 actor model, 197
 exercises, 205, 209, 211, 215
 Fibonacci server example, 211–215
 parallel map function, 2, 210
 performance, 203–205, 215
 problems with multiple core systems, ix
- cond, 134–137
- conditional logic vs. guard clauses, 60
- config option, 158
- config.exs file, 157
- config/ directory, 144
- configuration
 files and directories, 144, 157
 overriding default name, 158
- control flow, 133–139
- cookie option, 222
- cookies, 222
- copying data, immutability and, 23
- count, 334, 337, 345
- countdown timer example
 using streams, 108–110
- coupling and focal points, 258
- coverex, 184
- create_processes, 203
- CSV, sigil exercise, 350
- Ctrl-\, exiting iex, 4
 Ctrl-C, exiting iex, 4
 Ctrl-G q, exiting iex, 4
- curly braces, *see* braces
- cycle function, streams, 106
- cypher, Caesar, 331
-
- D**
- ~D sigil, 119
- data
 copying and immutability, 23
 immutable data as known data, 22
 transformation, GitHub project example, 149–162
 transforming and immutability, 22, 24
 transforming in functional programming, 1–3, 22, 24, 64
 transforming in production line diagram, 142, 168
- Date type, 33
- dates
 sigils, 119
 types, 33
- DateTime type, 34
- dealing cards, 102
- debug, 164, 235
- debugging
 debug trace, 235
 with iex, 8, 169–173
 injecting breakpoints, 170
 logging message levels, 164
 resources on, 173
 setting breakpoints, 172
- decimal integers, 26, 122
- def, 53, 323
- defexception, 358
- defimpl, 330–332
- defmacro, 305
- defmodule, 53
- defp, 63

- defprotocol, 330
 defstruct, 88, 332
 delimiters
 regular expressions, 27
 sigils, 119
 dependencies
 adding, 152
 finding, 157
 loading, 157
 managing with mix, 152, 186
 reporting on, 186
 starting, 154
 visualizations, 186
 deploy directory, 285
 deployment
 automated, 285
 hot upgrades, 234, 282, 287
 with ssh, 285
 upgrading during, 287–291
 deps, 152
 describe tool, ExUnit, 177
 descriptor, for application, 277
 destructuring, maps, 85
 Deutsch, L. Peter, 114
 development vs. production version, 284
 Dialyzer tool, 362, 366–371
 Dict module, 365
 dictionary types, 83–95, *see also* keyword lists; maps; sets; structs
 cautions, 95
 choosing, 83
 defined, 83
 nested accessors, 91
 nesting, 89–95
 using, 83–89
 difference, lists, 30
 dir_walker, 265
 directives, module, 66
 directories
 application specification files, 280
 creating, 143
 names, 145
 OTP applications, 284
 projects, 143–145
 source-code formatting, 190
 tests, 145, 147, 264
 Distillery release manager, 169, 282–292
 distributed failover, 292
 distributing, agents and tasks, 299
 div operator, 36
 division operator, 36
 do block
 scope, 36–40
 shortcut, 40
 syntax, 54
 do: syntax, 54, 133, 324
 docs, 167
 doctest tool, ExUnit, 176
 documentation, 354, 361, *see also* resources for this book projects, 166
 sample iex sessions in, 175
 dollar sign (\$), terminal prompt, 4
 :done, 269
 dot command, 186
 dot notation ()
 anonymous function calls, 41
 maps, 32
 module names and function calls, 68
 nested modules, 66
 nesting dictionary structures, 90
 structs, 88
 double quotes, heredocs ("""), 118, 120, 348
 double quotes (")
 atoms, 26
 enclosing strings, 73, 117, 124–130, 307
 :DOWN message, 209
 downcase function, 125
 downgrading, 288
 do...end block, 54
 duck typing, 98
 Duper example, 257–275
 about, 257
 performance, 273
 questions for organizing, 257–262
 sequence diagram, 260
 setup, 262–272
 supervision structure diagram, 262
 tests, 264
 duplicate, 125
 duplicate file example,
 see Duper example
 duplication
 immutability and, 23
 strings, 125
 dynamic nested accessors, 91
 dynamic supervisors, 266
-
- E**
- earmark, 166
 elem, 93
 Elixir
 about, x–xi
 advantages, 1–3
 basics, 25–40
 code formatter, 35, 320
 community support, xi, 10
 compiling and running, 9
 Distillery release manager, 169, 282–292
 exiting, 4
 formatting conventions, 10, 35
 installing, 3
 interactively running, 4–10
 message bug, 204
 resources on, xii, 10, 151
 source code, 316
 versions, 3
 elixir command, 10
 Elixometer, 190
 ellipsis (...), iex prompt, 4
 else: keyword, 133
 empty lists, 183
 empty?, 101
 end keyword, 41, 76
 ends_with? function, 125
 enoent error, 29
 Enum module, 366
 about, 99
 keyword lists, 84
 performance, 103
 using, 99–103
 Enumerable protocol, 99, 334–338
 environment, organizing questions, 257–258, 275
 equal sign (=)
 match operator, 15, 20

=== strict equality operator, 35
 == value equality operator, 35
 -erl parameter, Elixir, 205
Erlang
 about, ix
 cookies and, 223
 ETS tables and server-monitoring tools, 188
 forcing strings into Erlang terms, 121
 functions, calling, 44, 69
 libraries, finding, 69, 151
 naming conventions, 69
 pattern matching, 20
 resources on, 69, 151
 running two versions at same time, 288
 type specifications, 320, 361
 using function capture operator (&), 49
Erlang VM
 hot upgrades, 234
 increasing default values, 205
 nodes and, 219
error function, 357
error level logging, 164
error messages
 ExUnit, 147
 tests, 178
errors, *see also exceptions*
 error messages, 147, 178
 logging message levels, 164
 organizing questions, 257, 260, 275
 raising, 357
escape, 313
escape sequences
 ANSI, 326
 strings, 117
escript utility, 163
ETS tables, server-monitoring tools, 188
eval_quoted, 314, 316
eval_string, 315, 353
evaluator example of umbrella project, 351–354
events, server-monitoring tools, 190
.ex files, 9, 44
except clause, overriding sigils, 348
except: parameter, import directive, 67
exceptions, 355–359
 catching, 357
 defining, 358
 designing with, 138
 exiting, 357
 pattern matching, 356, 358
 raising, 138, 355–357
 re-raising, 357
 throwing, 357
exclamation mark (!)
 in names, 34
 relaxed Boolean operator, 36
 atoms, 26
 raising exceptions, 139
!= strict inequality operator, 35
!= value inequality operator, 35
excoveralls tool, 184–186
ExDoc, 166
exercises
 about, 10
 adding dependencies, 153
 anonymous functions, 42, 45, 47
 behaviours, 326
 binaries, 130–131
 character lists, 123
 collections, 102
 comprehensions, 114
 concurrency, 205, 209, 211, 215
 control flow, 139
 functions, 42, 45, 47, 50, 55, 57, 62, 70
 GitHub project example, 148, 153, 160
 guard clauses, 326
 libraries, 70
 lists and recursion, 77, 81
 macros, 311, 316, 326
 modules, 55, 70
 multiple processes, 205, 209, 211, 215
 named functions, 62
 nodes, 221, 226–227
 OTP applications, 282, 292
 OTP servers, 233, 237, 244
OTP supervisors, 250, 254
pattern matching, 18–19
project organization, 148, 153
protocols, 331, 345
resources for, 10
sigils, 350
strings, 123, 130–131
tests, 148, 316
weather fetching project, 168
exit
 exceptions, 357
 iex, 4
 preventing in mix, 273
 trapping, 208
exit function, 357
:EXIT message, 208
.exs files, 9, 44
extended mode, regular expressions, 27
ExUnit
 about, 146
 code-quoting exercise, 316
 doctest tool, 176
 GitHub project tests, basic, 146–149
 GitHub project tests, comments, 173–177
 GitHub project tests, formatting, 161
 GitHub project tests, sorting, 159
 property-based tests, 179–183
 resources on, 179, 183
 setup feature, 178
 structuring tests, 177–179
F
f option, regular expressions, 27
factorial implementation example, 55, 60
failover, distributed, 292
@fallback_to_any syntax, 331
false Boolean value, 35
fetch project, *see GitHub project example*
fetchable?, 320
Fibonacci examples
 server example, 211–215
 unfolding streams, 107

- file extensions, 9, 44
 File module, Elixir, 44
 File module, Erlang, 44
 files, *see also* Duper example
 file extensions, 9, 44
 loading, 10
 naming conventions, 9
 project directory, 143–
 145
 tests, 145, 147
 traversal, 265
 filter, 100
 filters
 collections, 100
 comprehensions, 112
 StreamData, 183
 first function, 126
 fixtures, 179
 FizzBuzz example, 45, 134–
 137, 139
 flatten, 81
 flattening, lists, 81
 float type, 130, 331
 floating-point numbers
 about, 26, 364
 as binaries, 124
 jaro_distance function, 126
 pattern matching, 130
 protocols, 331
 sigil exercise, 350
 fn keyword, 41, 76
 focal points
 defined, 258
 organizing questions,
 257–258, 275
 folding
 collection elements into
 single value, 101
 lists, 81
 foldl, 81
 foldr, 81
 for, iterating with, 85
 format, 190, 320
 format_error, 44
 format_status, 237, 239
 .formatter.exs file, 144
 formatting
 code formatter, 35, 320
 conventions, 10
 output formatter, 166
 source-code formatting,
 190–193
 tables, 160–162
 forward slash (/), division operator, 36
 fragments, code
 binding, 313
 evaluating, 314
 Frequency module example, 296
 fully qualified names, nodes, 219
 function capture operator (&), 48–50, 64, 76
 function type, 331
 functional programming
 vs. object-oriented programming, 1, 11
 thinking about, 11
 as transforming data, 1–
 3, 22, 24, 64
 FunctionClauseError, 357
 functions, *see also* helper functions; named functions; protocols; tasks
 agents, 295
 anonymous functions,
 41–51
 binary functions, 125–
 130
 creating, 41
 as data transformers, 3
 dictionary-access, 90
 differentiating between
 Elixir and Erlang, 44
 dynamic nested accessors, 91
 Erlang, calling, 44, 69
 exercises, 42, 45, 47, 50,
 55, 57, 62, 70
 guard clauses, 137
 help files, 6
 listing unknown in mix,
 186
 looping functions and
 timeouts, 272
 nested module functions,
 accessing, 65
 nesting, 45–47
 parallel processing, 3
 parameterized, 46
 passing as a key, 92
 passing as arguments,
 47–51
 pattern matching, 42
 private, 63
 recursive lists, 72–82
 server-monitoring tools,
 190
 specs for, 365–366
 syntax, 4, 6, 41–51
 as a type, 25
 updating, 263
 :functions atom, import directive, 67
-
- G**
- g option, regular expressions, 27
 garbage collection, ix, 23
 generators
 comprehensions, 111
 registering PIDs, 223
 GenServer
 about, 230
 adding behaviour with
 use, 231, 238
 default callbacks, 238
 overriding defaults, 255
 sequence server example,
 231–237, 240–244
 when to use, 300
 get, 153–154, 295
 get_and_update_in, 91–94
 get_env, 281
 get_in, 91–94
 get_status, 236
 GitHub project example
 adding libraries to, 151–
 156
 API for fetching issues,
 141
 command line, parsing,
 145–149
 command-line executable
 for, 163–164
 comments, 173–177
 configuring URLs, 157
 converting JSON list, 156
 directory tree for, 143–
 145
 documentation for, 166
 exercises, 148, 153, 160
 extracting issues from
 list, 159
 fetching data, 149–162
 formatting table, 160–162
 libraries for, 151
 logging for, 164–166
 overview, 141–142
 project, creating, 142–
 145
 sorting data, 158
 tests for, 146–149, 159,
 161, 173–186
 version control for, 144
 .gitignore file, 144

- global names, agents, 296, 299
graphemes, binary functions, 125–128
graphemes function, 126
greater-than-or-equal-to operator (\geq), 35
greediness, 27, 103
group leader, for nodes, 221, 227
group_leader function, 227
guard clauses
 case, 137
 vs. conditional logic, 60
exercises, 326
multiple processes, 204
using, 58–60
-
- H**
- h helper function, 5–6
Hackney library, 157
handle_call, 232, 238–239
handle_cast, 234, 238
handle_info, 238–239, 271
hash sign (#), preceding comments, 35
head
 building lists, 74
 pattern-matching complex lists, 78–81
 processing lists, 72–74
 splitting from binaries, 130
 understanding lists, 29, 71
Hello, World example
 creating, 9
 understanding processes, 198–202
help, 143
help files
 functions, 6
 mix, 143
helper functions
 c helper, 10, 53
 function capture operator (`&`), 48–50, 64, 76
 iex, 5–8
heredocs, 118, 120, 348
hex, 151, 157
hexadecimal integers, 26
:hibernate, 239
horizontal space, removal by formatter, 192
hot upgrades, 234, 282, 287
HTTPoison library, 151–156, 277
hungry consumer model, 260
hygiene and macros, 313–314
-
- I**
- i helper function, 7
i option, regular expressions, 27
identity function, 295
IEEE 754 double precision, 26
iex
 calling source files, 44
 customizing, 8
 debugging with, 8, 169–173
 exiting, 4
 helper functions, 5–8
 lists display in, 73
 running, 4–10
 sample sessions in documentation, 175
 starting servers manually, 232
 uses, 8
 working with large code examples, 44
iex.exs file, 8
if
 control flow, 133
 macro example, 303, 310
immutability, 21–24, *see also* state
@impl attribute, 321
import, 66, 162, 314
import_config, 158
import_file, 10
importing
 configuration, 158
 files, 10
 modules, 66, 162
 printing and, 162
 scope and, 314
in operator, 30, 36, 59
indentation, 10, 35
infinite streams, 105
info level logging, 164
init
 GenServer default callbacks, 238
 OTP servers, 231
starting workers before initialization, 269
supervisors, 267
inspect, 329, 332, 340–345
Inspect protocol, 330, 340–344
installation, Elixir, 3
integer type, 130, 331
integers
 about, 26, 364
 as binaries, 124
 code, 122
 converting strings to, 148
 lists display in iex, 73
 pattern matching, 130
 protocols, 331
 range of, representing, 363
into function, 338–340
into: parameter, comprehensions, 113
I/O servers, 226–227
issues project example,
 see GitHub project example
iterate function, streams, 106
iteration
 collections, 85
 comprehensions, 112
 streams, 106
 timers, 214
-
- J**
- jaro_distance function, 126
join function, 101
join operator (`|>`), 78, 363
join operators
 guard clauses, 59
 strings, 36
joining
 collections, 101
 lists, 78
 strings, 36
JSON, reformatting responses, 156
-
- K**
- Kernel module
 macros with same name, 324
 removing operator definitions, 315
 resources on, 316
Kerr, Jessica, 22
key function, 94
key! function, 94

- key/value pairs, *see also* dictionary types
 adding new key to a map, 87
 keyword lists, 30
 maps, 31
- keydelete, 81
- KeyError, 32
- keyfind, 81
- keyreplace, 81
- keys
 adding new key to a map, 87
 passing functions as, 92
 updating, 263
- keyword lists
 about, 30, 84, 97
 choosing, 83
 comprehensions, 112
 vs. maps, 31
 nested accessors, 91
 parameters, 281
 primitive vs. functional, 97
- Keyword module, 84, 97
- killing, processes, 206–210,
see also termination
- _kwlist_, 84
-
- L**
- ~1 sigil, 347
- last function, 126
- laziness, streams, 103–110
- length, 126
- less-than-or-equal-to operator
 (\leq), 35, 102
- lexical scope, modules, 66
- lib/ directory, 144–145
- libraries
 adding, 151–156
 compiling, 153
 directories, 144–145
 exercises, 70
 finding, 69, 151
 resources on, 69, 151, 227
 starting, 154
 using, 151–153
- LineSigil module, 347, 352
- linking
 modules with behaviours, 319–326
 processes, 207, 209, 247
 subprojects, 353
- Lisp, 29
- List module, 81, 97, 121
- list type, 331
- List.Chars protocol, 344
- lists, *see also* character lists;
 keyword lists
 about, 25, 29, 97, 362, 364
 building, 74
 as child specifications, 255
 comprehension, 111–114
 concatenating, 30, 36, 81
 converting collections to, 100
 converting into functions, 49
 creating, 16
 difference of, 30
 displaying in hex, 73
 empty, 75, 183
 exercises, 77, 81
 filtering, 78–81
 flattening, 81
 folding, 81
 functions, 81
 inserting elements into, 111
 joining, 78
 length of, 72
 of lists, 78–81
 mapping values to a new list, 75
 membership, 30
 ordering, 30, 135
 pattern matching, 16–20, 72–73, 75, 78–81
 performance, 30, 135
 pipe character () in, 71
 primitive vs. functional, 97
 processing, 72–74, 78
 processing multiple values in, 78
 protocols, 331
 recursion, 71–82
 reducing, 76–78
 replacing items in, 81
 swapping values in, 78
 updating, 81
- little, binary qualifier, 130
- load charts, 188
- local hostname, OS X and, 220
- local names
 agents, 296
 OTP servers, 240
 processes, 240
- logging
 debug trace, 235
 GitHub project example, 164–166
 message levels, 164
 sequence server project, 290
- looping functions and time-outs, 272
-
- M**
- m option, regular expressions, 27
- macros, 303–316
 binding, 312, 314
 cautions, 303
 defining, 305, 324
 exercises, 311, 316, 326
 importing from modules, 66
 load order, 306
 overriding operators, 315
 requiring, 67
 resources on, 316
 scope, 313
 understanding, 304–310
 using representation as code, 307–310
- :macros atom, import directive, 67
- main_module:, 163
- make_ref, 28
- map
 collections, 100
 creating function for lists, 75
 FizzBuzz example, 136
 passing anonymous functions as arguments, 47
- Map API, 84
- maps, *see also* structs
 about, 25, 31, 84, 97
 accessing, 31
 choosing, 83
 collections, 100
 destructuring, 85
 vs. keyword lists, 31
 nested accessors, 91
 pattern matching, 85–87
 primitive vs. functional, 97
 protocols, 331
 searching, 85–87
 sets, 95
 updating, 87, 263
- MapSet module, 95

- match operator (=), 15, 20
 MatchError exception, 139
 matching, *see* pattern matching
 max, 100
 max_by, 100
 member?, 101, 334, 337
 membership, lists, 30
 memory
 garbage collection, 23
 hibernating servers and, 239
 organizing questions, 258–259
 past limits on, ix
 restarting server after crash, 250
 server-monitoring tools, 189
 merging, collections, 101
 message attribute, 138
 message field, exceptions, 355, 358
 messages
 bug, 204
 exceptions, 138, 355, 358
 ExUnit, 147
 handle_info and, 239
 sending between processes, 198–202, 204
 server-monitoring tools, 190
 system messages in debug trace, 236
 metadata, module attributes, 67
 metaprogramming, 11
 methods, tracing example, 322–326
 MIDI example of protocols, 333–345
 migrating, server state, 288–291
 min_length, 183
 minus sign (-)
 subtraction operator, 36
 -- list difference, 30
 mix
 about, 142
 adding libraries, 151–156
 application specification file, creating, 278
 configuring OTP applications, 278–280
 creating projects, 142–145
 loading dependencies, 157
 logger, adding, 164–166
 managing dependencies, 152, 186
 packaging code, 163
 preventing exit, 273
 recompiling application with, 150
 resources on, 143
 running tests, 147
 -S option, 155, 232
 source-code control methods, 319
 source-code formatting, 190
 tasks list, 143
 umbrella projects, 351–354
 mix xref functions, 186
 mix.exs file, 143, 145, 151
 Mnesia, 229
 mod: option, 279
 modules, 53–70, *see also* protocols
 about, 53, 65
 aliasing, 67
 application entry points, 279
 attributes, 67
 child_spec, 255
 compiling, 53
 creating, 53
 directives, 66
 exercises, 55, 70
 help files, 6
 importing functions and macros, 66
 linking with behaviours, 319–326
 location, 145
 macros load order, 306
 names, 34, 68, 145, 240
 nesting, 65
 requiring, 67
 resources on, 69
 running two versions at same time, 288
 scope, 36, 66
 syntax, 54
 wrapping agents and tasks in, 300
 wrapping structs, 88
 wrapping with module functions, 240
 monitoring, processes, 208, 247
 multi-app projects, *see* umbrella projects
 multiple processes, running, *see* concurrency
 multiplication operator, 36
 myers_difference function, 126
-
- N**
- ~n for newline in strings, 121
 ~N sigil, 119
 NaiveDateTime type, 34
 --name option, nodes, 219
 name: option, nodes, 240
 named functions, 53–70
 about, 53
 arity, 54
 creating, 53
 default parameters, 60–63
 exercises, 62
 function capture operator (&), 50
 guard clauses, 58–60
 importing from modules, 66
 module prefix for, 65
 name conventions, 53
 pattern matching, 55–58
 pipe operator with, 63
 private, 63
 syntax, 54
 names
 agents, 296, 299
 applications, registering, 280
 behaviours, 34
 case and, 24
 children, 267
 directories, 145
 Erlang functions, 69
 file extensions, 9, 44
 local, 240, 296
 local hostname and OS X, 220
 modules, 34, 68, 145, 240
 named functions, 53
 naming conventions, 34, 44, 53, 69
 nodes, 219
 OTP applications, 280
 OTP servers, 240
 overriding default config file name, 158
 processes, 223–226, 240

- processes in OTP, 240
 projects, 145
 protocols, 34
 records, 34
 resources on, 69
 source files, 9, 44
 tests, 147
 variables, 34
- National Oceanic and Atmospheric Administration, 168
- native, binary qualifier, 130
- neg_integer type, 363
- negation operators, 59
- Nerves project, xi
- nesting
 accessors, 91
 anonymous functions, 45–47
 comprehensions, 111
 dictionary types, 89–95
 modules, 65
 pattern matching, 137
 structs, 89–95
- `new_map`, 87
- newline characters
 regular expressions, 27
 single-quoted strings, 121
- `next_codepoint`, 127
- `next_grapheme`, 127
- nil value, 35, 362
- no-halt flag, 273
- nodes, 219–228, *see also* OTP
 about, 219
 agents, 295, 299
 connecting, 220, 299
 exercises, 221, 226–227
 group leader for, 221, 227
 hierarchy of, 221
 I/O server, 226–227
 names, 219
 naming processes, 223–226
 number in PID, 221
 running multiple, 220
 security, 222
- `non_neg_integer` type, 363–364
- `none` type, 362
- `nonempty`, 183
- `not` operator, 35
- notifications
 process deaths, 206–210
 server notification example, 223–226
- O**
- object-oriented programming,
 vs. functional programming, 1, 11
- `:observer` tool, 187–190
- octal integers, 26
- `on_exit`, 179
- `:one_for_all` supervision strategy, 251, 268
- `:one_for_one` supervision strategy, 251, 267
- online resources, *see* resources for this book
- only parameter, import directive, 67
- `@opaque` attribute, 365
- `open`, 44
- Open Telecom Platform,
see OTP
- `open!`, 139
- operators
 about, 35
 guard clauses, 59
 list of, 35
 overriding with macros, 315
- options parsing, 145
- `or` operator, 35
- order
 comparison operators, 35
 function capture operator (&), 49
 list reversal, 135
 lists, 30, 135
 macros, 306
 named functions and pattern matching, 57
- OS X, local hostname and, 220
- OTP, *see also* Duper example;
 OTP applications; OTP servers; OTP supervisors
 about, 208, 229, 292
 agents, 295–300
 tasks, 293–295, 297–300
- OTP applications, 277–292
 about, 229
 compiling, 280
 configuring, 278–280
 creating, 278
 directory structure for, 284
 entry points, 279
 exercises, 282, 292
 hot upgrades, 282, 287
- names, 280
- parameters, 281
- processes in, 229
- registering, 280
- releasing code, 282–292
- sequence server example, 278–292
- specification files for, 278, 280
- terminology, 155, 277
- upgrading while deploying, 287–291
- versioning, 283–284, 287, 290
- OTP behaviours
 about, 229
 declaring, 320
 defined, 319
 defining, 319
- server (GenServer), 230–231, 238
- supervisor, 230, 247
- OTP servers, 229–245
 about, 230
 exercises, 233, 237, 244
 GenServer behavior for, 230–231, 238
- GenServer callback functions, 238
- hibernating, 239
- making into component, 242–244
- naming, 240
- one-way calls, 234
- organizing by characteristics, 259
- passing and returning tuples, 233
- processes in, naming, 240
- recompiling, 234
- replacing without stopping, 239
- sequence server example, 230–237, 240–244, 248–255, 278–292
- start order, 261
- starting manually, 232
- state managing, 230, 238
- statistics, 235
- status display, 239
- supervision strategy, 251, 266–267
- tasks as, 294
- terminating, 239, 244
- tracing execution, 235–237
- understanding, 230–237

- when to use, 300
- writing basic, 230–237
- OTP supervisors, 247–256
 - about, 230, 255
 - creating, 248
 - defined, 247
 - dynamic supervisors, 266
 - exercises, 250, 254
 - managing process state
 - across restarts, 250–254
 - reliability, 255, 281
 - restarting after termination, 249
 - sequence server example, 248–255, 278–292
 - starting servers in order, 261
 - subsupervisor, 261
 - supervision strategy, 251, 266–267
 - tasks, 294
 - worker restart options, 254
- output formatter, 166
- P**
- +P parameter, Erlang VM, 205
- packaging projects, 163–164
- pad_leading, 127
- pad_trailing, 127
- padding strings, 127
- parallel map function, 2, 210
- parallel processing, 2–3, *see also* concurrency
- parameterized functions, 46
- parameters
 - default, 60–63
 - keyword lists, 281
 - pinned values and, 48
- parentheses (())
- enclosing function arguments, 41
- grouping type specifications, 363
- named functions, 54, 64
- nesting functions, 46
- pipelines, 64
- using with, 39
- Parser module, 122
- parsing character lists, 122
- pattern matching, 15–20
 - anonymous functions, 42
 - assignment and, 15, 20, 42
 - binaries, 130
- binding variables, 18
- case and, 27, 79
- case construct, 137
- character lists, 121
- comprehensions, 111
- cond macro, 136
- default parameters, 60–63
- defined, 17
- exceptions, 356, 358
- exercises, 18–19
- forcing variable values, 19
- functions, 42
- guard clauses, 58–60
- ignoring values, 18
- inside patterns, 80
- lists, 16–20, 72–73, 75, 78–81
- maps, 85–87
- named functions, 55–58
- nesting, 137
- OTP servers, 233
- pinned values, 48
- process messages, 199
- regular expressions options, 27
- replacing strings, 128
- strings, 128, 130
- structs, 88
- tuples, 29
- understanding, 15–20
- wildcards, 18, 43, 358
- with with, 37–40
- PCRE, 27
- performance
 - double-quoted strings, 124
 - Duper, 273
 - Enum module, 103
 - garbage collection, 23
 - immutability and, 23
 - list reversal, 135
 - lists, 30, 135
 - processes, 203–205, 215
 - streams, 105
- Perl 5-compatible regular expressions, 27
- :permanent restart option, 254
- persistence, 215
- Phoenix, xi
- PID
 - calling, 28
 - defined, 28, 198
 - I/O server, 226–227
 - implementing protocols, 330
- naming processes, 223–226
- node number, 221
- parallel map function, 210
- registering, 223, 226
- tuples, 29
- PID type, 331
- pin operator (^)
 - in function parameters, 48
 - in pattern matching, 19, 87
- pinned values, 48
- Pinterest, 190
- pipe operator (|>), 63, 150
- pipelines
 - about, 2
 - lists, 71
 - named functions, 63
 - pipe operator (|>), 63, 150
- plus sign (+)
 - addition operator, 36
 - ++ list concatenation, 30, 36, 81
- pmap, 2, 210
- poison library, 156
- polymorphism
 - protocols and, 345
 - structs and, 89
- pop, 94
- port type, 331
- ports
 - defined, 28
 - port type, 331
- pos_integer type, 363
- precision, floats, 26
- primitive data types, 97
- printable?, 127
- private functions, 63
- process identifier, *see* PID
- Process_monitor function, 209
- processes, *see also* concurrency; OTP supervisors
 - about, 3, 11
 - creating, 203
 - exercises, 205, 209, 211, 215
 - Fibonacci server example, 211–215
 - garbage collection, 23
 - increasing default values, 205
 - killing, 206–210

- linking, 207, 209, 247
- managing process state across restarts, 250–254
- monitoring, 208, 247
- naming, 223–226, 240
- naming in OTP, 240
- parallel map function, 210
- performance, 203–205, 215
- PID for, 198
- running on different nodes, 221
- sending messages between, 198–202, 204
- starting, 198
- suspending, 209
- termination messages, 239
- trapping exit of, 208
- understanding, 198–202
- production vs. development version, 284
- Programming Erlang*, 256
- project function, 166
- projects, *see also* GitHub project example
 - adding libraries to, 151–156
 - command-line executable for, 163–164
 - configuring, 157
 - creating, 142–145
 - creating documentation for, 166
 - directory tree for, 143–145
 - logger, adding, 164–166
 - names, 145
 - organizing, 141–168
 - packaging, 163–164
 - source-code formatting, 190
 - subprojects, 351, 353
 - umbrella projects, 351–354
- property-based testing, 179–183
- protocols, 329–345
 - about, 114
 - built-in, 333–345
 - defined, 329
 - defining, 329
 - exercises, 331, 345
 - implementing, 330–332
 - names, 34
 - as polymorphism, 345
- structs and, 332
- types available to implement, 331
- pry, 170
- pull model, 260
- push model, 259
- put_in, 90
- put_new, 87
- puts
 - help files, 6
 - PIDs and nodes, 227
- Q**
- question mark (?)
 - atoms, 26
 - in names, 34
- questions for organizing applications, 257–262, 275
- Quixir, 180
- quotation marks, *see* double quotes; single quotes
- quote, 305–306, 311–312, 314
- R**
- r command, 234
- ~R sigil, 119
- ~r sigil, 27, 119
- race conditions, monitoring processes, 209
- raise, 138, 355–357
- raising, exceptions, 138, 355–357
- range operator (..), 363
- ranges, 26–27, 33
- Rashkovskii, Yurii, 361
- README.md file, 144
- real, 180
- receive, 199–202
- records
 - names, 34
 - protocols and, 331
- recursion
 - exercises, 77, 81
 - guard clauses for, 58
 - lists, 71–82
 - with named functions, 56
 - sending messages between processes, 202
 - tail recursion, 202
- reduce
 - accumulator example, 204
 - collection elements, 101
- exercises, 345
- protocols example, 334–337
- reducing
 - accumulator example, 204
 - collection elements into single value, 101
 - exercise, 345
 - lists, 76–78
 - protocols example, 334–337
- reductions value, debug trace, 236
- references, 28, 331
- Regex module, 28
- registered: option, 280
- registering
 - applications, 280
 - PIDs, 223, 226
- regular expressions
 - about, 26
 - options, 27
 - sigils, 119
 - splitting strings, 128
 - writing, 27
- reject, 100
- releases
 - code, 282–292
 - defined, 282
 - Distillery release manager, 169, 282–292
 - downgrading versions, 288
 - generating, 283–285
 - production version, 284
- reliability, OTP supervisors, 255, 281
- rem operator, 36
- remainder operator, 36
- repeatedly function, streams, 106
- replace, 128
- replace_at, 81
- reply, 232
- representation, using as code, 307–310
- require Logger, 166
- require directive, 67, 166, 306
- reraise, 357
- rescue clause, try function, 356
- resource, 105, 107–110
- resources
 - streaming, 105, 107–110

success typing, 368
 system messages, 236
 type specifications, 366
 resources for this book
 algebra documents, 344
 APIs, 99
 code files, 9
 debugging, 173
 Elixir, xii, 10, 151, 316
 Elixir source code, 316
 Erlang, 69, 151
 exercises, 10
 ExUnit, 179, 183
 guard clauses, 59
 Kernel module, 316
 libraries, 69, 151, 227
 macros, 316
 mix, 143
 modules, 69
 naming conventions, 69
 operator macros, 316
 quote, 314
 sys module, 236
 tasks, 295
 type-check functions, 59
 :rest_for_one supervision strategy, 252
 :restart option, 254
 restarts
 managing process state
 across, 250–254
 server after crash, 250
 worker restart options,
 254
 reverse, 128, 135
 run
 about, 145
 data-transformation example, 149
 multiple processes, 204
 scheduler example, 213
 running, organizing questions and, 257, 261, 275
 runtime
 logging options, 165
 nested accessors, 91
 organizing questions, 257, 259, 275
 RuntimeError, 138, 355, 357
 Russell, Cody, 215

S

s helper function, 362
 -S option, mix, 155, 232
 s option, regular expressions, 27
 ~S sigil, 119
 ~s sigil, 119
 say, 109
 scaling, automated, 292
 scheduler, Fibonacci server example, 211–215
 scope
 comprehensions, 113
 local scope, 37
 macros, 313
 modules, 36, 66
 quoted fragments, 314
 sigils, 348–349
 understanding, 36–40
 variables, 36–40, 46
 SCOWL, 331
 security, nodes, 222
 self, 28, 199
 send, 199–202, 239
 send sender..., 199
 send_after, 270
 separation of logic and implementation, sequence server example, 242–244
 sequence server example
 basic, 230–237
 interface cleanup, 240–244
 OTP application, 278–292
 supervisors, 248–255
 servers, *see also* OTP servers
 migrating server state, 288–291
 organizing by characteristics, 259
 server notification example, 223–226
 server-monitoring tools, 187–190
 start order, 261
 starting manually, 232
 sets, 95
 setup function, 178
 setup_all function, 179
 setups, testing, 178
 sigil_C, 348
 sigil_R, 348
 sigil_S, 348
 sigil_W, 348
 sigil_c, 348
 sigil_r, 348
 sigil_s, 347–348
 sigil_w, 348
 sigils, 118–120, 347–354
 signed, binary qualifier, 130
 :simple_one_for_one supervision strategy, 266
 single quotes, heredocs (""), 118, 120
 single quotes ('), enclosing strings, 73, 117, 121–123, 307
 size, binary qualifier, 130
 skipping generated values, 183
 slice, 128, 334, 337
 --sname option, nodes, 219
 sort, 100, 102, 158
 sort_into_descending_order, 159
 sorting
 collections, 100, 102
 data in GitHub project example, 158
 source files
 calling, 44
 names, 9, 44
 source-code control methods
 example, 319–322
 source-code formatting, 190–193
 spawn, 198, 204, 221
 spawn_link, 207, 209
 spawn_monitor, 209
 @spec attribute, 362, 365–366, 371
 special variable (), 18, 43, 74–75
 specification files, application, 278, 280
 split, 100, 128
 split_while, 100
 splitting
 binaries, 130
 collections, 100
 strings, 127–128, 130
 square brackets ([])
 accessing maps, 31
 c helper return value, 10
 enclosing lists, 30
 list creation, 16
 as nil value, 362
 ssh, 285
 stack trace, 357
 start, 248, 281
 start_link
 about, 232

- agents example, 296
 - defaults, 255
 - init and, 238
 - linking tasks, 294
 - sequence server example, 240
 - supervision, 249, 267
 - tracing server execution, 235
 - starting
 - applications, 261
 - organizing questions and, 257, 261, 275
 - processes, 198
 - restarting server after crash, 250
 - servers manually, 232
 - workers, 268–269
 - starts_with?, 129
 - state
 - agents, 295–297
 - hibernating servers and, 239
 - immutability in Elixir, 21–24
 - managing process state across restarts, 250–254
 - migrating server state, 288–291
 - naming processes, 226
 - in object-oriented programming, 1
 - OTP servers, 230, 238
 - unfolding streams, 107
 - versioning and, 283
 - statistics, debug trace, 235
 - :statistics option, 235
 - Stream
 - about, 99
 - creating streams, 105–110
 - using, 103–110
 - StreamData, 180–183
 - streams
 - about, 99
 - collectable, 111
 - as composable enumerators, 103
 - countdown timer example, 108–110
 - creating, 103, 105–110
 - infinite, 105
 - performance, 105
 - processing collections, 103–110
 - shortcut for, 105
 - streaming external resources, 105, 107–110
 - unfolding, 105, 107
 - when to use, 110
 - string interpolation, 117, 344, 348
 - String module, 125–130
 - String.Chars protocol, 344
 - string_to_quoted, 315
 - strings
 - about, 25
 - case, 24, 125, 129
 - codepoints in, 125
 - comprehensions, 112
 - converting to another string, 126
 - converting to code, 315
 - converting to integers, 148
 - converting values into, 344
 - displaying in iex, 73
 - double-quoted, 73, 117, 124–130, 307
 - duplicating, 125
 - escape sequences, 117
 - evaluating directly, 315
 - exercises, 123, 130–131
 - extracting elements from, 125–126
 - function capture operator (&), 49
 - heredocs, 118
 - interpolation, 44, 117, 344, 348
 - length of, 126
 - padding, 127
 - pattern matching, 128, 130
 - prefixes of, 129
 - printability of, 127
 - replacing elements in, 128
 - reversing, 128
 - sigils, 119
 - similarity of two strings, 126
 - single-quoted, 73, 117, 121–123, 307
 - slicing, 128
 - splitting, 127–128, 130
 - substrings of, 128
 - suffix of, testing, 125
 - terminology, 120
 - trimming, 129
 - understanding, 117–123
 - validity of, 130
 - structs
 - about, 25, 88
 - cautions, 95
 - defining, 363
 - nesting, 89–95
 - pattern matching, 88
 - protocols and, 332
 - updating, 88
 - subprojects
 - creating, 351
 - linking, 353
 - subsupervisor, 261
 - subtraction operator, 36
 - success typing, 368, 371
 - sum, 76
 - sup flag, 248
 - supervision strategy, 251, 266–267
 - supervisors, *see* OTP supervisors
 - suspending processes, 209
 - swap, 78
 - sys module, 236
 - system information, monitoring, 187–190
 - system messages, debug trace, 236
 - system types, 25, 28
-
- T**
- t helper function, 362
 - t option for ssh, 286
 - T sigil, 119
 - Taarnskov, Lau, 34
 - tables
 - formatting, 160–162
 - server-monitoring tools for Erlang ETS tables, 188
 - tail
 - about, 29, 71
 - building lists, 74
 - processing lists, 72–74
 - splitting from binaries, 130
 - tail-call optimization, 202
 - take, 100, 159, 337
 - take_every, 100
 - take_while, 100
 - tasks
 - anagram example, 297–300
 - defined, 293
 - distributing, 299

- mix, 143
 - resources on, 295
 - using, 293–295
 - when to use, 300
 - wrapping in modules, 300
 - Tate, Bruce, 3
 - tc library, 204, 214
 - templates, test files, 147
 - :temporary restart option, 254
 - terminal sessions, 4
 - terminate, 239, 252
 - termination
 - messages, 239
 - OTP servers, 239, 244
 - processes, 206–210, 294
 - supervision strategy, 251
 - tasks, 294
 - test coverage, 184–186
 - test task, 147
 - test/ directory, 145, 147
 - tests
 - Duper example, 264
 - exercises, 148, 316
 - file extensions for, 9
 - files and directories, 145, 147, 264
 - GitHub project, basic, 146–149
 - GitHub project, comments, 173–177
 - GitHub project, formatting, 161
 - GitHub project, sorting, 159
 - names, 147
 - property-based, 179–183
 - setups, 178
 - structuring, 177–179, 264
 - test coverage, 184–186
 - Theseus, 291
 - threads, 3, 197
 - throw, 357
 - tilde (~), preceding sigils, 118
 - Time type, 34
 - timeout parameter, 239
 - timeouts
 - GenServer, 238–239
 - looping functions and, 272
 - messages between processes, 201
 - timers, example of, 108–110, 204, 214
 - times
 - sigils, 119
 - types, 33
 - timestamps, debug trace, 236
 - to_charlist, 344
 - to_integer, 148
 - to_list, 100
 - to_process, 214
 - to_string, 315, 345
 - Torres, Devin, 210
 - tracing
 - methods call example, 322–326
 - OTP server execution, 235–237
 - stack trace, 357
 - transforming
 - > operator, 64
 - data and immutability, 22, 24
 - data in functional programming, 1–3, 22, 24, 64
 - data production line diagram, 142, 168
 - data, GitHub project example, 149–162
 - pipelines for, 2
 - :transient restart option, 255, 272
 - trapping process exit, 208
 - trim, 129
 - trim_leading, 129
 - trim_trailing, 129
 - true Boolean value, 35
 - truthy values, 35, 364, 366
 - try, 356
 - tuple type, 363
 - tuples
 - about, 28, 363
 - accessing in lists, 81
 - converting into functions, 49
 - pattern matching, 29
 - protocols, 331
 - @type attribute, 364–365
 - type checking, 362, 366–371
 - type inference, 369–371
 - type specifications, 320, 361–366
 - type-check functions, guard clauses, 59
 - typed maps, *see* structs
 - #Typeprefix, 340
 - @typep attribute, 365
 - types, *see also* dictionary types
 - aliases for, 364
 - built-in, 25–34, 364
 - collection, 25, 28–33
 - collections, 362
 - combining, 363
 - dates and times, 33
 - defined, 362
 - list, 362
 - local to modules, 365
 - new, defining, 364–365
 - primitive vs. functional, 97
 - protocols that can be implemented for, 331
 - system types, 25, 28
 - type specifications, 320
 - type-check functions, 59
 - understanding, 97
 - value, 25–28
-
- ## U
- U option, regular expressions, 27
 - u option, regular expressions, 27
 - umbrella flag, 351
 - umbrella projects, 351–354
 - unary operator, overriding, 315
 - underscore ()_
 - in decimals, 26
 - list processing, 74–75
 - in names, 34
 - special variable, 18, 43, 74–75, 362
 - unfold, 105, 107
 - ungreedy option, 27
 - Unicode, regular expressions, 27
 - union operator (), 363
 - unless, 133
 - unquote, 305, 309, 311
 - unquote_splicing, 310
 - unsigned, binary qualifier, 130
 - upcase, 129
 - update, 263, 295
 - update_in, 90
 - updating
 - agents, 295
 - functions, 263
 - hot upgrades, 282, 287

- lists, 81
 maps, 87, 263
 nested dictionary structures, 90
 structs, 88
- upgrades
 downgrading, 288
 during deployment, 287–291
 hot upgrades, 234, 282, 287
 version numbers, 287, 290
`_upto`, 135
- URLs, configuring, 157
- use
 about, 322
`GenServer`, 231, 238
 linking modules with behaviours, 319–326
 overriding defaults, 255
 tasks, 295
`_using_`, 322, 325, 349
- UTF strings, binaries and, 33
- UTF-8 encoding, 124, 130
- `utf8` type, 130
- `utf16` type, 130
- `utf32` type, 130
-
- V**
- `valid?`, 130
- Valim, José, 180, 215, 361
- values
 atom names, 27
 converting into characters, 344
 converting into strings, 344
 explicitly maintaining bindings, 353
 folding elements into single, 101
 forcing in pattern matching, 19
 help files, 7
 injecting with bindings, 312, 314
 injecting with unquote, 309
 inspecting protocol, 340–344
-
- W**
- `~w`
 formatting strings as Erlang term, 121
`sigil`, 119
- `~W` sigil, 119
- warn level logging, 164
- warnings, listing in mix, 186
- weather fetching project, 168
- website resources, *see* resources for this book
- when keyword, guard clauses, 58–60
- whitespace
 regular expressions, 27
`sigils`, 119
 trimming from strings, 129
- wildcards, pattern matching, 18, 43, 358
- with expression, using, 37–40
- with_index, 101
- workers
 about, 270
 adding, 267
 hungry consumer model, 260
 OTP supervisors and, 247
 push model, 259
 restart options, 254
 specifying, 255
 starting, 254, 268–269
- wrapping
 agents and tasks in modules, 300
 function capture operator (`&`), 50
 modules, 240
 structs, 88
-
- X**
- `x` option, regular expressions, 27
- `xref` functions, 186
-
- Z**
- `zip`, 101

Thank you!

How did you enjoy this book? Please let us know. Take a moment and email us at support@pragprog.com with your feedback. Tell us your story and you could win free ebooks. Please use the subject line "Book Feedback."

Ready for your next great Pragmatic Bookshelf book? Come on over to <https://pragprog.com> and use the coupon code BUYANOTHER2018 to save 30% on your next ebook.

Void where prohibited, restricted, or otherwise unwelcome. Do not use ebooks near water. If rash persists, see a doctor. Doesn't apply to *The Pragmatic Programmer* ebook because it's older than the Pragmatic Bookshelf itself. Side effects may include increased knowledge and skill, increased marketability, and deep satisfaction. Increase dosage regularly.

And thank you for your continued support,

Andy Hunt, Publisher



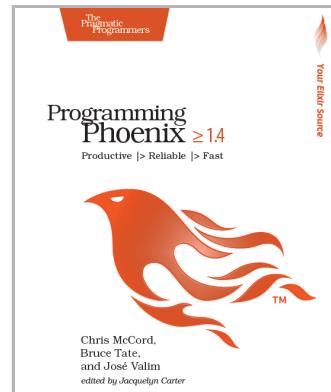
A Better Web with Phoenix and Elm

Elixir and Phoenix on the server side with Elm on the front end gets you the best of both worlds in both worlds!

Programming Phoenix ≥ 1.4

Don't accept the compromise between fast and beautiful: you can have it all. Phoenix creator Chris McCord, Elixir creator José Valim, and award-winning author Bruce Tate walk you through building an application that's fast and reliable. At every step, you'll learn from the Phoenix creators not just what to do, but why. Packed with insider insights and completely updated for Phoenix 1.4, this definitive guide will be your constant companion in your journey from Phoenix novice to expert, as you build the next generation of web applications.

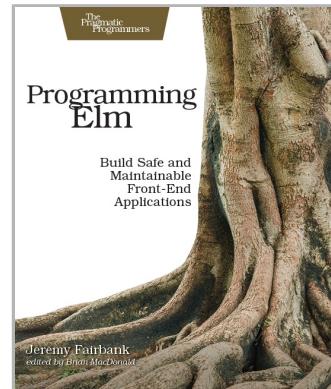
Chris McCord, Bruce Tate and José Valim
(325 pages) ISBN: 9781680502268. \$45.95
<https://pragprog.com/book/phoenix14>



Programming Elm

Elm brings the safety and stability of functional programming to front-end development, making it one of the most popular new languages. Elm's functional nature and static typing means that run-time errors are nearly impossible, and it compiles to JavaScript for easy web deployment. This book helps you take advantage of this new language in your web site development. Learn how the Elm Architecture will help you create fast applications. Discover how to integrate Elm with JavaScript so you can update legacy applications. See how Elm tooling makes deployment quicker and easier.

Jeremy Fairbank
(250 pages) ISBN: 9781680502855. \$40.95
<https://pragprog.com/book/jfelm>



Dive Deep in to OTP and Absinthe

Put it all together with Elixir, OTP, and Phoenix. Dive into GraphQL for better APIs in Elixir. It's all here.

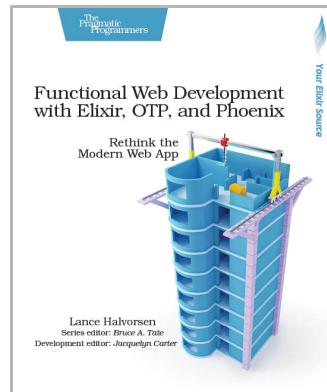
Functional Web Development with Elixir, OTP, and Phoenix

Elixir and Phoenix are generating tremendous excitement as an unbeatable platform for building modern web applications. For decades OTP has helped developers create incredibly robust, scalable applications with unparalleled uptime. Make the most of them as you build a stateful web app with Elixir, OTP, and Phoenix. Model domain entities without an ORM or a database. Manage server state and keep your code clean with OTP Behaviours. Layer on a Phoenix web interface without coupling it to the business logic. Open doors to powerful new techniques that will get you thinking about web development in fundamentally new ways.

Lance Halvorsen

(218 pages) ISBN: 9781680502435. \$45.95

<https://pragprog.com/book/lhelp>



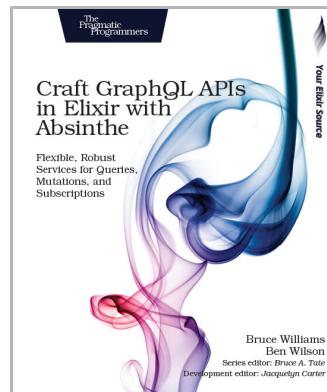
Craft GraphQL APIs in Elixir with Absinthe

Your domain is rich and interconnected, and your API should be too. Upgrade your web API to GraphQL, leveraging its flexible queries to empower your users, and its declarative structure to simplify your code. Absinthe is the GraphQL toolkit for Elixir, a functional programming language designed to enable massive concurrency atop robust application architectures. Written by the creators of Absinthe, this book will help you take full advantage of these two groundbreaking technologies. Build your own flexible, high-performance APIs using step-by-step guidance and expert advice you won't find anywhere else.

Bruce Williams and Ben Wilson

(302 pages) ISBN: 9781680502558. \$47.95

<https://pragprog.com/book/wwgraphql>



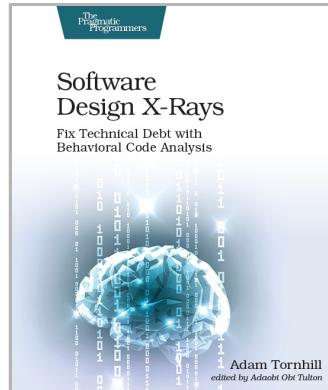
Fix Your Hidden Problems

From technical debt to deployment in the very real, very messy world, we've got the tools you need to fix the hidden problems before they become disasters.

Software Design X-Rays

Are you working on a codebase where cost overruns, death marches, and heroic fights with legacy code monsters are the norm? Battle these adversaries with novel ways to identify and prioritize technical debt, based on behavioral data from how developers work with code. And that's just for starters. Because good code involves social design, as well as technical design, you can find surprising dependencies between people and code to resolve coordination bottlenecks among teams. Best of all, the techniques build on behavioral data that you already have: your version-control system. Join the fight for better code!

Adam Tornhill
(274 pages) ISBN: 9781680502725. \$45.95
<https://pragprog.com/book/atevol>



Release It! Second Edition

A single dramatic software failure can cost a company millions of dollars—but can be avoided with simple changes to design and architecture. This new edition of the best-selling industry standard shows you how to create systems that run longer, with fewer failures, and recover better when bad things happen. New coverage includes DevOps, microservices, and cloud-native architecture. Stability antipatterns have grown to include systemic problems in large-scale systems. This is a must-have pragmatic guide to engineering for production systems.

Michael Nygard
(376 pages) ISBN: 9781680502398. \$47.95
<https://pragprog.com/book/mnne2>



Pragmatic Programming

We'll show you how to be more pragmatic and effective, from design to daily practice.

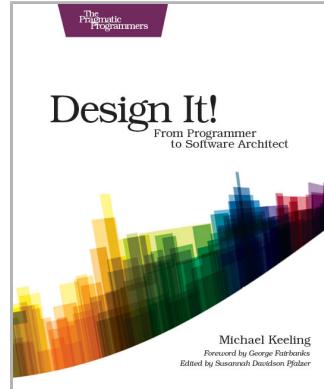
Design It!

Don't engineer by coincidence—design it like you mean it! Grounded by fundamentals and filled with practical design methods, this is the perfect introduction to software architecture for programmers who are ready to grow their design skills. Ask the right stakeholders the right questions, explore design options, share your design decisions, and facilitate collaborative workshops that are fast, effective, and fun. Become a better programmer, leader, and designer. Use your new skills to lead your team in implementing software with the right capabilities—and develop awesome software!

Michael Keeling

(358 pages) ISBN: 9781680502091. \$41.95

<https://pragprog.com/book/mkdsa>



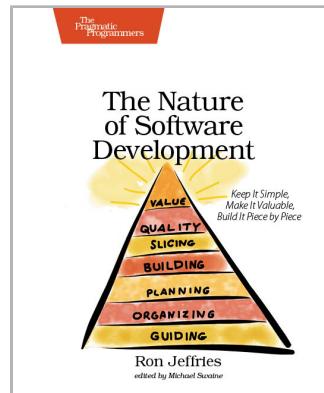
The Nature of Software Development

You need to get value from your software project. You need it "free, now, and perfect." We can't get you there, but we can help you get to "cheaper, sooner, and better." This book leads you from the desire for value down to the specific activities that help good Agile projects deliver better software sooner, and at a lower cost. Using simple sketches and a few words, the author invites you to follow his path of learning and understanding from a half century of software development and from his engagement with Agile methods from their very beginning.

Ron Jeffries

(176 pages) ISBN: 9781941222379. \$24

<https://pragprog.com/book/rjnsd>



The Joy of Mazes and Math

Rediscover the joy and fascinating weirdness of mazes and pure mathematics.

Mazes for Programmers

A book on mazes? Seriously?

Yes!

Not because you spend your day creating mazes, or because you particularly like solving mazes.

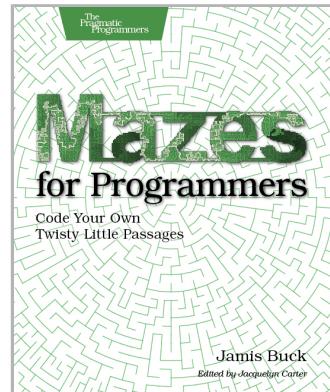
But because it's fun. Remember when programming used to be fun? This book takes you back to those days when you were starting to program, and you wanted to make your code do things, draw things, and solve puzzles. It's fun because it lets you explore and grow your code, and reminds you how it feels to just think.

Sometimes it feels like you live your life in a maze of twisty little passages, all alike. Now you can code your way out.

Jamis Buck

(286 pages) ISBN: 9781680500554. \$38

<https://pragprog.com/book/jbmaze>



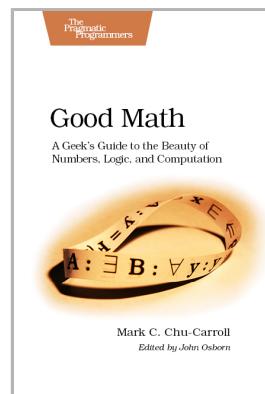
Good Math

Mathematics is beautiful—and it can be fun and exciting as well as practical. *Good Math* is your guide to some of the most intriguing topics from two thousand years of mathematics: from Egyptian fractions to Turing machines; from the real meaning of numbers to proof trees, group symmetry, and mechanical computation. If you've ever wondered what lay beyond the proofs you struggled to complete in high school geometry, or what limits the capabilities of the computer on your desk, this is the book for you.

Mark C. Chu-Carroll

(282 pages) ISBN: 9781937785338. \$34

<https://pragprog.com/book/mcmath>



The Pragmatic Bookshelf

The Pragmatic Bookshelf features books written by developers for developers. The titles continue the well-known Pragmatic Programmer style and continue to garner awards and rave reviews. As development gets more and more difficult, the Pragmatic Programmers will be there with more titles and products to help you stay on top of your game.

Visit Us Online

This Book's Home Page

<https://pragprog.com/book/elixir16>

Source code from this book, errata, and other resources. Come give us feedback, too!

Keep Up to Date

<https://pragprog.com>

Join our announcement mailing list (low volume) or follow us on twitter @pragprog for new titles, sales, coupons, hot tips, and more.

New and Noteworthy

<https://pragprog.com/news>

Check out the latest pragmatic developments, new titles and other offerings.

Buy the Book

If you liked this eBook, perhaps you'd like to have a paper copy of the book. It's available for purchase at our store: <https://pragprog.com/book/elixir16>

Contact Us

Online Orders: <https://pragprog.com/catalog>

Customer Service: support@pragprog.com

International Rights: translations@pragprog.com

Academic Use: academic@pragprog.com

Write for Us: <http://write-for-us.pragprog.com>

Or Call: +1 800-699-7764