

You Don't Know JS Yet: Get Started - 2nd Edition

Foreword

The first time I saw a tweet advertising the crowdfunding campaign for the original book series *You Don't Know JS*, I thought to myself that whoever the hell this Kyle Simpson is can get bent. Of course I know JavaScript! At the time I had been working in JavaScript for many years with some of tech's biggest names, so I felt justified in looking down my nose at this sweeping claim.

Once the funding for the book completed, I observed that a lot of people were really excited to read this book series and I decided to go ahead and give it a shot—you know, just to show everyone that I *do know* JavaScript. Once I dug in and poured over the texts, I experienced a profound mix of learning, surprise, and even a little anger. Kyle has a knack for saying something that challenges my world view and making me think about it until I realize what he said was actually true (I'll never admit this to him though!).

So it turned out I didn't know JavaScript. I didn't know why I had adopted certain patterns; I didn't know why JavaScript behaves the way it does in certain situations, and I didn't know much of the nuance of the languages that I assumed I did. Many were things I didn't know I didn't know, and I was worse off as a developer for it.

This is what this book series is great for. It's not just for someone picking up the language for the first time (though it's for them, too); it's for all software craftspeople who want to master their tools, who want to understand the ins and outs of their trade, and who want to select the proper methods for solving problems.

The thing I appreciate about Kyle and his work is that he is delightfully uninfluenced by the zeitgeist of the programming world around him. That isn't to say he isn't aware of what's happening in the community but rather that it doesn't sway his pursuit of the best answers to the correct questions. This often has put him at odds with the latest "best practices," but in reality this is exactly what you need: a perspective removed from temporal influence and simply speaking to the underlying truth. That's why this series is so good. The first edition of *You Don't Know JS* is still accurate, years later! Not many things stand this test of time given the shifting sands of the JavaScript landscape.

Let's talk a moment about this first book itself, *Get Started*. You may be tempted to skip it as you may think that you already have "gotten started," but it's worth reading this one! You'd be surprised how much depth, nuance, and oddity there is in the underlying building blocks of JavaScript, and it's quite essential for you to grapple with these underpinnings before venturing forward into the constructs of the language. It's exactly the sort of foundation you need,

to really know JavaScript.

So, do your future self a favor and dig into this book and unlock the knowledge within. These solid foundations will serve you better than any framework ever will; those come and go but we'll still be writing JavaScript itself for decades to come. Keep an open mind and challenge your preconceived notions.

Because, as I found out myself, you probably don't know JavaScript (yet).

Brian Holt Senior Program Manager Visual Studio Code and Node.js on Azure
Microsoft

You Don't Know JS Yet - 2nd Edition

Preface

Welcome to the 2nd edition of the widely acclaimed *You Don't Know JS* (**YDKJS**) book series: *You Don't Know JS Yet* (**YDKJSY**).

If you've read any of the 1st edition books, you can expect a refreshed approach in these new ones, with plenty of updated coverage of what's changed in JS over the last five years. But what I hope and believe you'll still *get* is the same commitment to respecting JS and digging into what really makes it tick.

If this is your first time reading these books, I'm glad you're here. Prepare for a deep and extensive journey into all the corners of JavaScript.

If you are new to programming or JS, be aware that these books are not intended as a gentle "intro to JavaScript." This material is, at times, complex and challenging, and goes much deeper than is typical for a first-time learner. You're welcome here no matter what your background is, but these books are written assuming you're already comfortable with JS and have at least 6–9 months experience with it.

The Parts

These books approach JavaScript intentionally opposite of how *The Good Parts* treats the language. No, that doesn't mean we're looking at *the bad parts*, but rather, exploring **all the parts**.

You may have been told, or felt yourself, that JS is a deeply flawed language that was poorly designed and inconsistently implemented. Many have asserted that it's the worst most popular language in the world; that nobody writes JS because they want to, only because they have to given its place at the center of the web. That's a ridiculous, unhealthy, and wholly condescending claim.

Millions of developers write JavaScript every day, and many of them appreciate and respect the language.

Like any great language, it has its brilliant parts as well as its scars. Even the creator of JavaScript himself, Brendan Eich, laments some of those parts as mistakes. But he's wrong: they weren't mistakes at all. JS is what it is today—the world's most ubiquitous and thus most influential programming language—precisely because of *all those parts*.

Don't buy the lie that you should only learn and use a small collection of *good parts* while avoiding all the bad stuff. Don't buy the "X is the new Y" snake oil, that some new feature of the language instantly relegates all usage of a previous feature as obsolete and ignorant. Don't listen when someone says your code isn't "modern" because it isn't yet using a stage-0 feature that was only proposed a few weeks ago!

Every part of JS is useful. Some parts are more useful than others. Some parts require you to be more careful and intentional.

I find it absurd to try to be a truly effective JavaScript developer while only using a small sliver of what the language has to offer. Can you imagine a construction worker with a toolbox full of tools, who only uses their hammer and scoffs at the screwdriver or tape measure as inferior? That's just silly.

My unreserved claim is that you should go about learning all parts of JavaScript, and where appropriate, use them! And if I may be so bold as to suggest: it's time to discard any JS books that tell you otherwise.

The Title?

So what's the title of the series all about?

I'm not trying to insult you with criticism about your current lack of knowledge or understanding of JavaScript. I'm not suggesting you can't or won't be able to learn JavaScript. I'm not boasting about secret advanced insider wisdom that I and only a select few possess.

Seriously, all those were real reactions to the original series title before folks even read the books. And they're baseless.

The primary point of the title "You Don't Know JS Yet" is to point out that most JS developers don't take the time to really understand how the code that they write works. They know *that* it works—that it produces a desired outcome. But they either don't understand exactly *how*, or worse, they have an inaccurate mental model for the *how* that falters on closer scrutiny.

I'm presenting a gentle but earnest challenge to you the reader, to set aside the assumptions you have about JS, and approach it with fresh eyes and an invigorated curiosity that leads you to ask *why* for every line of code you write. Why does it do what it does? Why is one way better or more appropriate than the other half-dozen ways you could have accomplished it? Why do all the "popular kids" say to do X with your code, but it turns out that Y might be a better choice?

I added "Yet" to the title, not only because it's the second edition, but because ultimately I want these books to challenge you in a hopeful rather than discouraging way.

But let me be clear: I don't think it's possible to ever fully *know* JS. That's not an achievement to be obtained, but a goal to strive after. You don't finish knowing everything about JS, you just keep learning more and more as you spend more time with the language. And the deeper you go, the more you revisit what you *knew* before, and you re-learn it from that more experienced perspective.

I encourage you to adopt a mindset around JavaScript, and indeed all of software

development, that you will never fully have mastered it, but that you can and should keep working to get closer to that end, a journey that will stretch for the entirety of your software development career, and beyond.

You can always know JS better than you currently do. That's what I hope these YDKJSY books represent.

The Mission

The case doesn't really need to be made for why developers should take JS seriously—I think it's already more than proven worthy of first-class status among the world's programming languages.

But a different, more important case still needs to be made, and these books rise to that challenge.

I've taught more than 5,000 developers from teams and companies all over the world, in more than 25 countries on six continents. And what I've seen is that far too often, what *counts* is generally just the result of the program, not how the program is written or how/why it works.

My experience not only as a developer but in teaching many other developers tells me: you will always be more effective in your development work if you more completely understand how your code works than you are solely *just* getting it to produce a desired outcome.

In other words, *good enough to work* is not, and should not be, *good enough*.

All developers regularly struggle with some piece of code not working correctly, and they can't figure out why. But far too often, JS developers will blame this on the language rather than admitting it's their own understanding that is falling short. These books serve as both the question and answer: why did it do *this*, and here's how to get it to do *that* instead.

My mission with YDKJSY is to empower every single JS developer to fully own the code they write, to understand it and to write with intention and clarity.

The Path

Some of you have started reading this book with the goal of completing all six books, back to back.

I would like to caution you to consider changing that plan.

It is not my intention that YDKJSY be read straight through. The material in these books is dense, because JavaScript is powerful, sophisticated, and in parts rather complex. Nobody can really hope to *download* all this information to their brains in a single pass and retain any significant amount of it. That's unreasonable, and it's foolish to try.

My suggestion is you take your time going through YDKJSY. Take one chapter, read it completely through start to finish, and then go back and re-read it section by section. Stop in between each section, and practice the code or ideas from that section. For larger concepts, it probably is a good idea to expect to spend several days digesting, re-reading, practicing, then digesting some more.

You could spend a week or two on each chapter, and a month or two on each book, and a year or more on the whole series, and you would still not be squeezing every ounce of YDKJSY out.

Don't binge these books; be patient and spread out your reading. Interleave reading with lots of practice on real code in your job or on projects you participate in. Wrestle with the opinions I've presented along the way, debate with others, and most of all, disagree with me! Run a study group or book club. Teach mini-workshops at your office. Write blog posts on what you've learned. Speak about these topics at local JS meetups.

It's never my goal to convince you to agree with my opinion, but to encourage you to own and be able to defend your opinions. You can't get *there* with an expedient read-through of these books. That's something that takes a long while to emerge, little by little, as you study and ponder and re-visit.

These books are meant to be a field-guide on your wanderings through JavaScript, from wherever you currently are with the language, to a place of deeper understanding. And the deeper you understand JS, the more questions you will ask and the more you will have to explore! That's what I find so exciting!

I'm so glad you're embarking on this journey, and I am so honored you would consider and consult these books along the way. It's time to start *getting to know JS*.

You Don't Know JS Yet: Get Started - 2nd Edition

Chapter 1: What *Is* JavaScript?

You don't know JS, yet. Neither do I, not fully anyway. None of us do. But we can all start getting to know JS better.

In this first chapter of the first book of the *You Don't Know JS Yet* (YDKJSY) series, we will take some time to build a foundation to move forward on. We need to start by covering a variety of important background housekeeping details, clearing up some myths and misconceptions about what the language really is (and isn't!).

This is valuable insight into the identity and process of how JS is organized and maintained; all JS developers should understand it. If you want to get to know JS, this is how to *get started* taking the first steps in that journey.

About This Book

I emphasize the word *journey* because *knowing JS* is not a destination, it's a direction. No matter how much time you spend with the language, you will always be able to find something else to learn and understand a little better. So don't look at this book as something to rush through for a quick achievement. Instead, patience and persistence are best as you take these first few steps.

Following this background chapter, the rest of the book lays out a high-level map of what you will find as you dig into and study JS with the YDKJSY books.

In particular, Chapter 4 identifies three main pillars around which the JS language is organized: scope/closures, prototypes/objects, and types/coercion. JS is a broad and sophisticated language, with many features and capabilities. But all of JS is founded on these three foundational pillars.

Keep in mind that even though this book is titled "Get Started," it's **not intended as a beginner/intro book**. This book's main job is to get you ready for studying JS deeply throughout the rest of the series; it's written assuming you already have familiarity with JS over at least several months experience before moving on in YDKJSY. So to get the most out of *Get Started*, make sure you spend plenty of time writing JS code to build up your experience.

Even if you've already written a lot of JS before, this book should not be skimmed over or skipped; take your time to fully process the material here. **A good start always depends on a solid first step.**

What's With That Name?

The name JavaScript is probably the most mistaken and misunderstood programming language name.

Is this language related to Java? Is it only the script form for Java? Is it only for writing scripts and not real programs?

The truth is, the name JavaScript is an artifact of marketing shenanigans. When Brendan Eich first conceived of the language, he code-named it Mocha. Internally at Netscape, the brand LiveScript was used. But when it came time to publicly name the language, “JavaScript” won the vote.

Why? Because this language was originally designed to appeal to an audience of mostly Java programmers, and because the word “script” was popular at the time to refer to lightweight programs. These lightweight “scripts” would be the first ones to embed inside of pages on this new thing called the web!

In other words, JavaScript was a marketing ploy to try to position this language as a palatable alternative to writing the heavier and more well-known Java of the day. It could just as easily have been called “WebJava,” for that matter.

There are some superficial resemblances between JavaScript’s code and Java code. Those similarities don’t particularly come from shared development, but from both languages targeting developers with assumed syntax expectations from C (and to an extent, C++).

For example, we use the { to begin a block of code and the } to end that block of code, just like C/C++ and Java. We also use the ; to punctuate the end of a statement.

In some ways, legal relationships run even deeper than the syntax. Oracle (via Sun), the company that still owns and runs Java, also owns the official trademark for the name “JavaScript” (via Netscape). This trademark is almost never enforced, and likely couldn’t be at this point.

For these reasons, some have suggested we use JS instead of JavaScript. That is a very common shorthand, if not a good candidate for an official language branding itself. Indeed, these books use JS almost exclusively to refer to the language.

Further distancing the language from the Oracle-owned trademark, the official name of the language specified by TC39 and formalized by the ECMA standards body is **ECMAScript**. And indeed, since 2016, the official language name has also been suffixed by the revision year; as of this writing, that’s ECMAScript 2019, or otherwise abbreviated ES2019.

In other words, the JavaScript/JS that runs in your browser or in Node.js, is *an* implementation of the ES2019 standard.

NOTE:

Don't use terms like "JS6" or "ES8" to refer to the language. Some do, but those terms only serve to perpetuate confusion. "ES20xx" or just "JS" are what you should stick to.

Whether you call it JavaScript, JS, ECMAScript, or ES2019, it's most definitely not a variant of the Java language!

"Java is to JavaScript as ham is to hamster." —Jeremy Keith, 2009

Language Specification

I mentioned TC39, the technical steering committee that manages JS. Their primary task is managing the official specification for the language. They meet regularly to vote on any agreed changes, which they then submit to ECMA, the standards organization.

JS's syntax and behavior are defined in the ES specification.

ES2019 happens to be the 10th major numbered specification/revision since JS's inception in 1995, so in the specification's official URL as hosted by ECMA, you'll find "10.0":

<https://www.ecma-international.org/ecma-262/10.0/>

The TC39 committee is comprised of between 50 and about 100 different people from a broad section of web-invested companies, such as browser makers (Mozilla, Google, Apple) and device makers (Samsung, etc). All members of the committee are volunteers, though many of them are employees of these companies and so may receive compensation in part for their duties on the committee.

TC39 meets generally about every other month, usually for about three days, to review work done by members since the last meeting, discuss issues, and vote on proposals. Meeting locations rotate among member companies willing to host.

All TC39 proposals progress through a five-stage process—of course, since we're programmers, it's 0-based!—Stage 0 through Stage 4. You can read more about the Stage process here: <https://tc39.es/process-document/>

Stage 0 means roughly, someone on TC39 thinks it's a worthy idea and plans to champion and work on it. That means lots of ideas that non-TC39 members "propose," through informal means such as social media or blog posts, are really "pre-stage 0." You have to get a TC39 member to champion a proposal for it to be considered "Stage 0" officially.

Once a proposal reaches "Stage 4" status, it is eligible to be included in the next yearly revision of the language. It can take anywhere from several months to a few years for a proposal to work its way through these stages.

All proposals are managed in the open, on TC39's Github repository: <https://github.com/tc39/proposals>

Anyone, whether on TC39 or not, is welcome to participate in these public discussions and the processes for working on the proposals. However, only TC39 members can attend meetings and vote on the proposals and changes. So in effect, the voice of a TC39 member carries a lot of weight in where JS will go.

Contrary to some established and frustratingly perpetuated myth, there are *not* multiple versions of JavaScript in the wild. There's just **one JS**, the official standard as maintained by TC39 and ECMA.

Back in the early 2000s, when Microsoft maintained a forked and reverse-engineered (and not entirely compatible) version of JS called "JScript," there were legitimately "multiple versions" of JS. But those days are long gone. It's outdated and inaccurate to make such claims about JS today.

All major browsers and device makers have committed to keeping their JS implementations compliant with this one central specification. Of course, engines implement features at different times. But it should never be the case that the v8 engine (Chrome's JS engine) implements a specified feature differently or incompatibly as compared to the SpiderMonkey engine (Mozilla's JS engine).

That means you can learn **one JS**, and rely on that same JS everywhere.

The Web Rules Everything About (JS)

While the array of environments that run JS is constantly expanding (from browsers, to servers (Node.js), to robots, to lightbulbs, to...), the one environment that rules JS is the web. In other words, how JS is implemented for web browsers is, in all practicality, the only reality that matters.

For the most part, the JS defined in the specification and the JS that runs in browser-based JS engines is the same. But there are some differences that must be considered.

Sometimes the JS specification will dictate some new or refined behavior, and yet that won't exactly match with how it works in browser-based JS engines. Such a mismatch is historical: JS engines have had 20+ years of observable behaviors around corner cases of features that have come to be relied on by web content. As such, sometimes the JS engines will refuse to conform to a specification-dictated change because it would break that web content.

In these cases, often TC39 will backtrack and simply choose to conform the specification to the reality of the web. For example, TC39 planned to add a `contains(..)` method for Arrays, but it was found that this name conflicted with old JS frameworks still in use on some sites, so they changed the name to a non-conflicting `includes(..)`. The same happened with a comedic/tragic JS *community crisis* dubbed "smooshgate," where the planned `flatten(..)` method was eventually renamed `flat(..)`.

But occasionally, TC39 will decide the specification should stick firm on some point even though it is unlikely that browser-based JS engines will ever conform.

The solution? Appendix B, “Additional ECMAScript Features for Web Browsers”.¹ The JS specification includes this appendix to detail out any known mismatches between the official JS specification and the reality of JS on the web. In other words, these are exceptions that are allowed *only* for web JS; other JS environments must stick to the letter of the law.

Section B.1 and B.2 cover *additions* to JS (syntax and APIs) that web JS includes, again for historical reasons, but which TC39 does not plan to formally specify in the core of JS. Examples include 0-prefixed octal literals, the global `escape(..) / unescape(..)` utilities, String “helpers” like `anchor(..)` and `blink()`, and the RegExp `compile(..)` method.

Section B.3 includes some conflicts where code may run in both web and non-web JS engines, but where the behavior *could* be observably different, resulting in different outcomes. Most of the listed changes involve situations that are labeled as early errors when code is running in strict mode.

Appendix B *gotchas* aren’t encountered very often, but it’s still a good idea to avoid these constructs to be future safe. Wherever possible, adhere to the JS specification and don’t rely on behavior that’s only applicable in certain JS engine environments.

Not All (Web) JS...

Is this code a JS program?

```
alert("Hello, JS!");
```

Depends on how you look at things. The `alert(..)` function shown here is not included in the JS specification, but it *is* in all web JS environments. Yet, you won’t find it in Appendix B, so what gives?

Various JS environments (like browser JS engines, Node.js, etc.) add APIs into the global scope of your JS programs that give you environment-specific capabilities, like being able to pop an alert-style box in the user’s browser.

In fact, a wide range of JS-looking APIs, like `fetch(..)`, `getCurrentLocation(..)`, and `getUserMedia(..)`, are all web APIs that look like JS. In Node.js, we can access hundreds of API methods from various built-in modules, like `fs.write(..)`.

Another common example is `console.log(..)` (and all the other `console.*` methods!). These are not specified in JS, but because of their universal utility

¹ECMAScript 2019 Language Specification, Appendix B: Additional ECMAScript Features for Web Browsers, <https://www.ecma-international.org/ecma-262/10.0/#sec-additional-ecmascript-features-for-web-browsers> (latest as of time of this writing in January 2020)

are defined by pretty much every JS environment, according to a roughly agreed consensus.

So `alert(..)` and `console.log(..)` are not defined by JS. But they *look* like JS. They are functions and object methods and they obey JS syntax rules. The behaviors behind them are controlled by the environment running the JS engine, but on the surface they definitely have to abide by JS to be able to play in the JS playground.

Most of the cross-browser differences people complain about with “JS is so inconsistent!” claims are actually due to differences in how those environment behaviors work, not in how the JS itself works.

So an `alert(..)` call *is* JS, but `alert` itself is really just a guest, not part of the official JS specification.

It’s Not Always JS

Using the console/REPL (Read-Evaluate-Print-Loop) in your browser’s Developer Tools (or Node) feels like a pretty straightforward JS environment at first glance. But it’s not, really.

Developer Tools are... tools for developers. Their primary purpose is to make life easier for developers. They prioritize DX (Developer Experience). It is *not* a goal of such tools to accurately and purely reflect all nuances of strict-spec JS behavior. As such, there’s many quirks that may act as “gotchas” if you’re treating the console as a *pure* JS environment.

This convenience is a good thing, by the way! I’m glad Developer Tools make developers’ lives easier! I’m glad we have nice UX charms like auto-complete of variables/properties, etc. I’m just pointing out that we can’t and shouldn’t expect such tools to *always* adhere strictly to the way JS programs are handled, because that’s not the purpose of these tools.

Since such tools vary in behavior from browser to browser, and since they change (sometimes rather frequently), I’m not going to “hardcode” any of the specific details into this text, thereby ensuring this book text is outdated quickly.

But I’ll just hint at some examples of quirks that have been true at various points in different JS console environments, to reinforce my point about not assuming native JS behavior while using them:

- Whether a `var` or `function` declaration in the top-level “global scope” of the console actually creates a real global variable (and mirrored `window` property, and vice versa!).
- What happens with multiple `let` and `const` declarations in the top-level “global scope.”
- Whether `"use strict";` on one line-entry (pressing `<enter>` after) enables strict mode for the rest of that console session, the way it would on

the first line of a .js file, as well as whether you can use "use strict"; beyond the "first line" and still get strict mode turned on for that session.

- How non-strict mode `this` default-binding works for function calls, and whether the "global object" used will contain expected global variables.
- How hoisting (see Book 2, *Scope & Closures*) works across multiple line entries.
- ...several others

The developer console is not trying to pretend to be a JS compiler that handles your entered code exactly the same way the JS engine handles a .js file. It's trying to make it easy for you to quickly enter a few lines of code and see the results immediately. These are entirely different use cases, and as such, it's unreasonable to expect one tool to handle both equally.

Don't trust what behavior you see in a developer console as representing *exact* to-the-letter JS semantics; for that, read the specification. Instead, think of the console as a "JS-friendly" environment. That's useful in its own right.

Many Faces

The term "paradigm" in programming language context refers to a broad (almost universal) mindset and approach to structuring code. Within a paradigm, there are myriad variations of style and form that distinguish programs, including countless different libraries and frameworks that leave their unique signature on any given code.

But no matter what a program's individual style may be, the big picture divisions around paradigms are almost always evident at first glance of any program.

Typical paradigm-level code categories include procedural, object-oriented (OO/classes), and functional (FP):

- Procedural style organizes code in a top-down, linear progression through a pre-determined set of operations, usually collected together in related units called procedures.
- OO style organizes code by collecting logic and data together into units called classes.
- FP style organizes code into functions (pure computations as opposed to procedures), and the adaptations of those functions as values.

Paradigms are neither right nor wrong. They're orientations that guide and mold how programmers approach problems and solutions, how they structure and maintain their code.

Some languages are heavily slanted toward one paradigm—C is procedural, Java/C++ are almost entirely class oriented, and Haskell is FP through and through.

But many languages also support code patterns that can come from, and even mix and match from, different paradigms. So called “multi-paradigm languages” offer ultimate flexibility. In some cases, a single program can even have two or more expressions of these paradigms sitting side by side.

JavaScript is most definitely a multi-paradigm language. You can write procedural, class-oriented, or FP-style code, and you can make those decisions on a line-by-line basis instead of being forced into an all-or-nothing choice.

Backwards & Forwards

One of the most foundational principles that guides JavaScript is preservation of *backwards compatibility*. Many are confused by the implications of this term, and often confuse it with a related but different term: *forwards compatibility*.

Let’s set the record straight.

Backwards compatibility means that once something is accepted as valid JS, there will not be a future change to the language that causes that code to become invalid JS. Code written in 1995—however primitive or limited it may have been!—should still work today. As TC39 members often proclaim, “we don’t break the web!”

The idea is that JS developers can write code with confidence that their code won’t stop working unpredictably because a browser update is released. This makes the decision to choose JS for a program a more wise and safe investment, for years into the future.

That “guarantee” is no small thing. Maintaining backwards compatibility, stretched out across almost 25 years of the language’s history, creates an enormous burden and a whole slew of unique challenges. You’d be hard pressed to find many other examples in computing of such a commitment to backwards compatibility.

The costs of sticking to this principle should not be casually dismissed. It necessarily creates a very high bar to including changing or extending the language; any decision becomes effectively permanent, mistakes and all. Once it’s in JS, it can’t be taken out because it might break programs, even if we’d really, really like to remove it!

There are some small exceptions to this rule. JS has had some backwards-incompatible changes, but TC39 is extremely cautious in doing so. They study existing code on the web (via browser data gathering) to estimate the impact of such breakage, and browsers ultimately decide and vote on whether they’re willing to take the heat from users for a very small-scale breakage weighed against the benefits of fixing or improving some aspect of the language for many more sites (and users).

These kinds of changes are rare, and are almost always in corner cases of usage that are unlikely to be observably breaking in many sites.

Compare *backwards compatibility* to its counterpart, *forwards compatibility*. Being forwards-compatible means that including a new addition to the language in a program would not cause that program to break if it were run in an older JS engine. **JS is not forwards-compatible**, despite many wishing such, and even incorrectly believing the myth that it is.

HTML and CSS, by contrast, are forwards-compatible but not backwards-compatible. If you dug up some HTML or CSS written back in 1995, it's entirely possible it would not work (or work the same) today. But, if you use a new feature from 2019 in a browser from 2010, the page isn't "broken" – the unrecognized CSS/HTML is skipped over, while the rest of the CSS/HTML would be processed accordingly.

It may seem desirable for forwards-compatibility to be included in programming language design, but it's generally impractical to do so. Markup (HTML) or styling (CSS) are declarative in nature, so it's much easier to "skip over" unrecognized declarations with minimal impact to other recognized declarations.

But chaos and non-determinism would ensue if a programming language engine selectively skipped statements (or even expressions!) that it didn't understand, as it's impossible to ensure that a subsequent part of the program wasn't expecting the skipped-over part to have been processed.

Though JS isn't, and can't be, forwards-compatible, it's critical to recognize JS's backwards compatibility, including the enduring benefits to the web and the constraints and difficulties it places on JS as a result.

Jumping the Gaps

Since JS is not forwards-compatible, it means that there is always the potential for a gap between code that you can write that's valid JS, and the oldest engine that your site or application needs to support. If you run a program that uses an ES2019 feature in an engine from 2016, you're very likely to see the program break and crash.

If the feature is a new syntax, the program will in general completely fail to compile and run, usually throwing a syntax error. If the feature is an API (such as ES6's `Object.is(...)`), the program may run up to a point but then throw a runtime exception and stop once it encounters the reference to the unknown API.

Does this mean JS developers should always lag behind the pace of progress, using only code that is on the trailing edge of the oldest JS engine environments they need to support? No!

But it does mean that JS developers need to take special care to address this gap.

For new and incompatible syntax, the solution is transpiling. Transpiling is a contrived and community-invented term to describe using a tool to convert the

source code of a program from one form to another (but still as textual source code). Typically, forwards-compatibility problems related to syntax are solved by using a transpiler (the most common one being Babel (<https://babeljs.io>)) to convert from that newer JS syntax version to an equivalent older syntax.

For example, a developer may write a snippet of code like:

```
if (something) {  
    let x = 3;  
    console.log(x);  
}  
else {  
    let x = 4;  
    console.log(x);  
}
```

This is how the code would look in the source code tree for that application. But when producing the file(s) to deploy to the public website, the Babel transpiler might convert that code to look like this:

```
var x$0, x$1;  
if (something) {  
    x$0 = 3;  
    console.log(x$0);  
}  
else {  
    x$1 = 4;  
    console.log(x$1);  
}
```

The original snippet relied on `let` to create block-scoped `x` variables in both the `if` and `else` clauses which did not interfere with each other. An equivalent program (with minimal re-working) that Babel can produce just chooses to name two different variables with unique names, producing the same non-interference outcome.

NOTE:

The `let` keyword was added in ES6 (in 2015). The preceding example of transpiling would only need to apply if an application needed to run in a pre-ES6 supporting JS environment. The example here is just for simplicity of illustration. When ES6 was new, the need for such a transpilation was quite prevalent, but in 2020 it's much less common to need to support pre-ES6 environments. The "target" used for transpilation is thus a sliding window that shifts upward only as decisions are made for a site/application to stop supporting some old browser/engine.

You may wonder: why go to the trouble of using a tool to convert from a

newer syntax version to an older one? Couldn't we just write the two variables and skip using the `let` keyword? The reason is, it's strongly recommended that developers use the latest version of JS so that their code is clean and communicates its ideas most effectively.

Developers should focus on writing the clean, new syntax forms, and let the tools take care of producing a forwards-compatible version of that code that is suitable to deploy and run on the oldest-supported JS engine environments.

Filling the Gaps

If the forwards-compatibility issue is not related to new syntax, but rather to a missing API method that was only recently added, the most common solution is to provide a definition for that missing API method that stands in and acts as if the older environment had already had it natively defined. This pattern is called a polyfill (aka "shim").

Consider this code:

```
// getSomeRecords() returns us a promise for some
// data it will fetch
var pr = getSomeRecords();

// show the UI spinner while we get the data
startSpinner();

pr
  .then(renderRecords)    // render if successful
  .catch(showError)       // show an error if not
  .finally(hideSpinner)   // always hide the spinner
```

This code uses an ES2019 feature, the `finally(..)` method on the promise prototype. If this code were used in a pre-ES2019 environment, the `finally(..)` method would not exist, and an error would occur.

A polyfill for `finally(..)` in pre-ES2019 environments could look like this:

```
if (!Promise.prototype.finally) {
  Promise.prototype.finally = function f(fn){
    return this.then(
      function t(v){
        return Promise.resolve( fn() )
          .then(function t(){
            return v;
          });
      },
      function c(e){
        return Promise.resolve( fn() )
          .then(function t(){
```

```

        throw e;
    });
}
);
}
}

```

WARNING:

This is only a simple illustration of a basic (not entirely spec-compliant) polyfill for `finally(...)`. Don't use this polyfill in your code; always use a robust, official polyfill wherever possible, such as the collection of polyfills/shims in ES-Shim.

The `if` statement protects the polyfill definition by preventing it from running in any environment where the JS engine has already defined that method. In older environments, the polyfill is defined, but in newer environments the `if` statement is quietly skipped.

Transpilers like Babel typically detect which polyfills your code needs and provide them automatically for you. But occasionally you may need to include/define them explicitly, which works similar to the snippet we just looked at.

Always write code using the most appropriate features to communicate its ideas and intent effectively. In general, this means using the most recent stable JS version. Avoid negatively impacting the code's readability by trying to manually adjust for the syntax/API gaps. That's what tools are for!

Transpilation and polyfilling are two highly effective techniques for addressing that gap between code that uses the latest stable features in the language and the old environments a site or application needs to still support. Since JS isn't going to stop improving, the gap will never go away. Both techniques should be embraced as a standard part of every JS project's production chain going forward.

What's in an Interpretation?

A long-debated question for code written in JS: is it an interpreted script or a compiled program? The majority opinion seems to be that JS is an interpreted (scripting) language. But the truth is more complicated than that.

For much of the history of programming languages, “interpreted” languages and “scripting” languages have been looked down on as inferior compared to their compiled counterparts. The reasons for this acrimony are numerous, including the perception that there is a lack of performance optimization, as well as dislike of certain language characteristics, such as scripting languages generally using dynamic typing instead of the “more mature” statically typed languages.

Languages regarded as “compiled” usually produce a portable (binary) representation of the program that is distributed for execution later. Since we don’t really observe that kind of model with JS (we distribute the source code, not the binary form), many claim that disqualifies JS from the category. In reality, the distribution model for a program’s “executable” form has become drastically more varied and also less relevant over the last few decades; to the question at hand, it doesn’t really matter so much anymore what form a program gets passed around.

These misinformed claims and criticisms should be set aside. The real reason it matters to have a clear picture on whether JS is interpreted or compiled relates to the nature of how errors are handled.

Historically, scripted or interpreted languages were executed in generally a top-down and line-by-line fashion; there’s typically not an initial pass through the program to process it before execution begins (see Figure 1).

Fig. 1: Interpreted/Scripted Execution

In scripted or interpreted languages, an error on line 5 of a program won’t be discovered until lines 1 through 4 have already executed. Notably, the error on line 5 might be due to a runtime condition, such as some variable or value having an unsuitable value for an operation, or it may be due to a malformed statement/command on that line. Depending on context, deferring error handling to the line the error occurs on may be a desirable or undesirable effect.

Compare that to languages which do go through a processing step (typically, called parsing) before any execution occurs, as illustrated in Figure 2:

Fig. 2: Parsing + Compilation + Execution

In this processing model, an invalid command (such as broken syntax) on line 5 would be caught during the parsing phase, before any execution has begun, and none of the program would run. For catching syntax (or otherwise “static”) errors, generally it’s preferred to know about them ahead of any doomed partial execution.

So what do “parsed” languages have in common with “compiled” languages? First, all compiled languages are parsed. So a parsed language is quite a ways down the road toward being compiled already. In classic compilation theory, the last remaining step after parsing is code generation: producing an executable form.

Once any source program has been fully parsed, it’s very common that its subsequent execution will, in some form or fashion, include a translation from the parsed form of the program—usually called an Abstract Syntax Tree (AST)—to that executable form.

In other words, parsed languages usually also perform code generation before execution, so it’s not that much of a stretch to say that, in spirit, they’re compiled languages.

JS source code is parsed before it is executed. The specification requires as much, because it calls for “early errors”—statically determined errors in code, such as a duplicate parameter name—to be reported before the code starts executing. Those errors cannot be recognized without the code having been parsed.

So **JS is a parsed language**, but is it *compiled*?

The answer is closer to yes than no. The parsed JS is converted to an optimized (binary) form, and that “code” is subsequently executed (Figure 2); the engine does not commonly switch back into line-by-line execution (like Figure 1) mode after it has finished all the hard work of parsing—most languages/engines wouldn’t, because that would be highly inefficient.

To be specific, this “compilation” produces a binary byte code (of sorts), which is then handed to the “JS virtual machine” to execute. Some like to say this VM is “interpreting” the byte code. But then that means Java, and a dozen other JVM-driven languages, for that matter, are interpreted rather than compiled. Of course, that contradicts the typical assertion that Java/etc are compiled languages.

Interestingly, while Java and JavaScript are very different languages, the question of interpreted/compiled is pretty closely related between them!

Another wrinkle is that JS engines can employ multiple passes of JIT (Just-In-Time) processing/optimization on the generated code (post parsing), which again could reasonably be labeled either “compilation” or “interpretation” depending on perspective. It’s actually a fantastically complex situation under the hood of a JS engine.

So what do these nitty-gritty details boil down to? Step back and consider the entire flow of a JS source program:

1. After a program leaves a developer’s editor, it gets transpiled by Babel, then packed by Webpack (and perhaps half a dozen other build processes), then it gets delivered in that very different form to a JS engine.
2. The JS engine parses the code to an AST.
3. Then the engine converts that AST to a kind-of byte code, a binary intermediate representation (IR), which is then refined/converted even further by the optimizing JIT compiler.
4. Finally, the JS VM executes the program.

To visualize those steps, again:

Fig. 3: Parsing, Compiling, and Executing JS

Is JS handled more like an interpreted, line-by-line script, as in Figure 1, or is it handled more like a compiled language that’s processed in one-to-several passes first, before execution (as in Figures 2 and 3)?

I think it’s clear that in spirit, if not in practice, **JS is a compiled language**.

And again, the reason that matters is, since JS is compiled, we are informed of static errors (such as malformed syntax) before our code is executed. That is a substantively different interaction model than we get with traditional “scripting” programs, and arguably more helpful!

Web Assembly (WASM)

One dominating concern that has driven a significant amount of JS’s evolution is performance, both how quickly JS can be parsed/compiled and how quickly that compiled code can be executed.

In 2013, engineers from Mozilla Firefox demonstrated a port of the Unreal 3 game engine from C to JS. The ability for this code to run in a browser JS engine at full 60fps performance was predicated on a set of optimizations that the JS engine could perform specifically because the JS version of the Unreal engine’s code used a style of code that favored a subset of the JS language, named “ASM.js”.

This subset is valid JS written in ways that are somewhat uncommon in normal coding, but which signal certain important typing information to the engine that allow it to make key optimizations. ASM.js was introduced as one way of addressing the pressures on the runtime performance of JS.

But it’s important to note that ASM.js was never intended to be code that was authored by developers, but rather a representation of a program having been transpiled from another language (such as C), where these typing “annotations” were inserted automatically by the tooling.

Several years after ASM.js demonstrated the validity of tooling-created versions of programs that can be processed more efficiently by the JS engine, another group of engineers (also, initially, from Mozilla) released Web Assembly (WASM).

WASM is similar to ASM.js in that its original intent was to provide a path for non-JS programs (C, etc.) to be converted to a form that could run in the JS engine. Unlike ASM.js, WASM chose to additionally get around some of the inherent delays in JS parsing/compilation before a program can execute, by representing the program in a form that is entirely unlike JS.

WASM is a representation format more akin to Assembly (hence, its name) that can be processed by a JS engine by skipping the parsing/compilation that the JS engine normally does. The parsing/compilation of a WASM-targeted program happen ahead of time (AOT); what’s distributed is a binary-packed program ready for the JS engine to execute with very minimal processing.

An initial motivation for WASM was clearly the potential performance improvements. While that continues to be a focus, WASM is additionally motivated by the desire to bring more parity for non-JS languages to the web platform. For example, if a language like Go supports threaded programming, but JS (the

language) does not, WASM offers the potential for such a Go program to be converted to a form the JS engine can understand, without needing a threads feature in the JS language itself.

In other words, WASM relieves the pressure to add features to JS that are mostly/exclusively intended to be used by transpiled programs from other languages. That means JS feature development can be judged (by TC39) without being skewed by interests/demands in other language ecosystems, while still letting those languages have a viable path onto the web.

Another perspective on WASM that's emerging is, interestingly, not even directly related to the web (W). WASM is evolving to become a cross-platform virtual machine (VM) of sorts, where programs can be compiled once and run in a variety of different system environments.

So, WASM isn't only for the web, and WASM also isn't JS. Ironically, even though WASM runs in the JS engine, the JS language is one of the least suitable languages to source WASM programs with, because WASM relies heavily on static typing information. Even TypeScript (TS)—ostensibly, JS + static types—is not quite suitable (as it stands) to transpile to WASM, though language variants like AssemblyScript are attempting to bridge the gap between JS/TS and WASM.

This book isn't about WASM, so I won't spend much more time discussing it, except to make one final point. *Some* folks have suggested WASM points to a future where JS is excised from, or minimized in, the web. These folks often harbor ill feelings about JS, and want some other language—any other language!—to replace it. Since WASM lets other languages run in the JS engine, on its face this isn't an entirely fanciful fairytale.

But let me just state simply: WASM will not replace JS. WASM significantly augments what the web (including JS) can accomplish. That's a great thing, entirely orthogonal to whether some people will use it as an escape hatch from having to write JS.

Strictly Speaking

Back in 2009 with the release of ES5, JS added *strict mode* as an opt-in mechanism for encouraging better JS programs.

The benefits of strict mode far outweigh the costs, but old habits die hard and the inertia of existing (aka “legacy”) code bases is really hard to shift. So sadly, more than 10 years later, strict mode's *optionality* means that it's still not necessarily the default for JS programmers.

Why strict mode? Strict mode shouldn't be thought of as a restriction on what you can't do, but rather as a guide to the best way to do things so that the JS engine has the best chance of optimizing and efficiently running the code. Most JS code is worked on by teams of developers, so the *strict*-ness of strict mode

(along with tooling like linters!) often helps collaboration on code by avoiding some of the more problematic mistakes that slip by in non-strict mode.

Most strict mode controls are in the form of *early errors*, meaning errors that aren't strictly syntax errors but are still thrown at compile time (before the code is run). For example, strict mode disallows naming two function parameters the same, and results in an early error. Some other strict mode controls are only observable at runtime, such as how `this` defaults to `undefined` instead of the global object.

Rather than fighting and arguing with strict mode, like a kid who just wants to defy whatever their parents tell them not to do, the best mindset is that strict mode is like a linter reminding you how JS *should* be written to have the highest quality and best chance at performance. If you find yourself feeling handcuffed, trying to work around strict mode, that should be a blaring red warning flag that you need to back up and rethink the whole approach.

Strict mode is switched on per file with a special pragma (nothing allowed before it except comments/whitespace):

```
// only whitespace and comments are allowed
// before the use-strict pragma
"use strict";
// the rest of the file runs in strict mode
```

WARNING:

Something to be aware of is that even a stray `;` all by itself appearing before the strict mode pragma will render the pragma useless; no errors are thrown because it's valid JS to have a string literal expression in a statement position, but it also will silently *not* turn on strict mode!

Strict mode can alternatively be turned on per-function scope, with exactly the same rules about its surroundings:

```
function someOperations() {
  // whitespace and comments are fine here
  "use strict";

  // all this code will run in strict mode
}
```

Interestingly, if a file has strict mode turned on, the function-level strict mode pragmas are disallowed. So you have to pick one or the other.

The **only** valid reason to use a per-function approach to strict mode is when you are converting an existing non-strict mode program file and need to make the changes little by little over time. Otherwise, it's vastly better to simply turn strict mode on for the entire file/program.

Many have wondered if there would ever be a time when JS made strict mode the default? The answer is, almost certainly not. As we discussed earlier around backwards compatibility, if a JS engine update started assuming code was strict mode even if it's not marked as such, it's possible that this code would break as a result of strict mode's controls.

However, there are a few factors that reduce the future impact of this non-default "obscurity" of strict mode.

For one, virtually all transpiled code ends up in strict mode even if the original source code isn't written as such. Most JS code in production has been transpiled, so that means most JS is already adhering to strict mode. It's possible to undo that assumption, but you really have to go out of your way to do so, so it's highly unlikely.

Moreover, a wide shift is happening toward more/most new JS code being written using the ES6 module format. ES6 modules assume strict mode, so all code in such files is automatically defaulted to strict mode.

Taken together, strict mode is largely the de facto default even though technically it's not actually the default.

Defined

JS is an implementation of the ECMAScript standard (version ES2019 as of this writing), which is guided by the TC39 committee and hosted by ECMA. It runs in browsers and other JS environments such as Node.js.

JS is a multi-paradigm language, meaning the syntax and capabilities allow a developer to mix and match (and bend and reshape!) concepts from various major paradigms, such as procedural, object-oriented (OO/classes), and functional (FP).

JS is a compiled language, meaning the tools (including the JS engine) process and verify a program (reporting any errors!) before it executes.

With our language now *defined*, let's start getting to know its ins and outs.

You Don't Know JS Yet: Get Started - 2nd Edition

Chapter 2: Surveying JS

The best way to learn JS is to start writing JS.

To do that, you need to know how the language works, and that's what we'll focus on here. Even if you've programmed in other languages before, take your time getting comfortable with JS, and make sure to practice each piece.

This chapter is not an exhaustive reference on every bit of syntax of the JS language. It's also not intended to be a complete "intro to JS" primer.

Instead, we're just going to survey some of the major topic areas of the language. Our goal is to get a better *feel* for it, so that we can move forward writing our own programs with more confidence. We'll revisit many of these topics in successively more detail as you go through the rest of this book, and the rest of the series.

Please don't expect this chapter to be a quick read. It's long and there's plenty of detail to chew on. Take your time.

TIP:

If you're still getting familiar with JS, I suggest you reserve plenty of extra time to work through this chapter. Take each section and ponder and explore the topic for awhile. Look through existing JS programs and compare what you see in them to the code and explanations (and opinions!) presented here. You will get a lot more out of the rest of the book and series with a solid foundation of JS's *nature*.

Each File is a Program

Almost every website (web application) you use is comprised of many different JS files (typically with the .js file extension). It's tempting to think of the whole thing (the application) as one program. But JS sees it differently.

In JS, each standalone file is its own separate program.

The reason this matters is primarily around error handling. Since JS treats files as programs, one file may fail (during parse/compile or execution) and that will not necessarily prevent the next file from being processed. Obviously, if your application depends on five .js files, and one of them fails, the overall application will probably only partially operate, at best. It's important to ensure that each file works properly, and that to whatever extent possible, they handle failure in other files as gracefully as possible.

It may surprise you to consider separate .js files as separate JS programs. From the perspective of your usage of an application, it sure seems like one big program. That's because the execution of the application allows these individual *programs* to cooperate and act as one program.

NOTE:

Many projects use build process tools that end up combining separate files from the project into a single file to be delivered to a web page. When this happens, JS treats this single combined file as the entire program.

The only way multiple standalone .js files act as a single program is by sharing their state (and access to their public functionality) via the “global scope.” They mix together in this global scope namespace, so at runtime they act as a whole.

Since ES6, JS has also supported a module format in addition to the typical standalone JS program format. Modules are also file-based. If a file is loaded via module-loading mechanism such as an `import` statement or a `<script type="module">` tag, all its code is treated as a single module.

Though you wouldn’t typically think about a module—a collection of state and publicly exposed methods to operate on that state—as a standalone program, JS does in fact still treat each module separately. Similar to how “global scope” allows standalone files to mix together at runtime, importing a module into another allows runtime interoperation between them.

Regardless of which code organization pattern (and loading mechanism) is used for a file (standalone or module), you should still think of each file as its own (mini) program, which may then cooperate with other (mini) programs to perform the functions of your overall application.

Values

The most fundamental unit of information in a program is a value. Values are data. They’re how the program maintains state. Values come in two forms in JS: **primitive** and **object**.

Values are embedded in programs using *literals*:

```
greeting("My name is Kyle.");
```

In this program, the value "My name is Kyle." is a primitive string literal; strings are ordered collections of characters, usually used to represent words and sentences.

I used the double-quote " character to *delimit* (surround, separate, define) the string value. But I could have used the single-quote ' character as well. The choice of which quote character is entirely stylistic. The important thing, for

the sake of code readability and maintainability, is to pick one and to use it consistently throughout the program.

Another option to delimit a string literal is to use the back-tick ` character. However, this choice is not merely stylistic; there's a behavioral difference as well. Consider:

```
console.log("My name is ${ firstName }.");
// My name is ${ firstName }.

console.log('My name is ${ firstName }.');
// My name is ${ firstName }.

console.log(`My name is ${ firstName }.`);
// My name is Kyle.
```

Assuming this program has already defined a variable `firstName` with the string value "Kyle", the ` -delimited string then resolves the variable expression (indicated with `${ ... }`) to its current value. This is called **interpolation**.

The back-tick ` -delimited string can be used without including interpolated expressions, but that defeats the whole purpose of that alternate string literal syntax:

```
console.log(
  `Am I confusing you by omitting interpolation?`
);
// Am I confusing you by omitting interpolation?
```

The better approach is to use " or ' (again, pick one and stick to it!) for strings *unless you need* interpolation; reserve ` only for strings that will include interpolated expressions.

Other than strings, JS programs often contain other primitive literal values such as booleans and numbers:

```
while (false) {
  console.log(3.141592);
}
```

`while` represents a loop type, a way to repeat operations *while* its condition is true.

In this case, the loop will never run (and nothing will be printed), because we used the `false` boolean value as the loop conditional. `true` would have resulted in a loop that keeps going forever, so be careful!

The number `3.141592` is, as you may know, an approximation of mathematical PI to the first six digits. Rather than embed such a value, however, you would typically use the predefined `Math.PI` value for that purpose. Another variation

on numbers is the `bigint` (big-integer) primitive type, which is used for storing arbitrarily large numbers.

Numbers are most often used in programs for counting steps, such as loop iterations, and accessing information in numeric positions (i.e., an array index). We'll cover arrays/objects in a little bit, but as an example, if there was an array called `names`, we could access the element in its second position like this:

```
console.log(`My name is ${ names[1] }.`);
// My name is Kyle.
```

We used `1` for the element in the second position, instead of `2`, because like in most programming languages, JS array indices are 0-based (`0` is the first position).

In addition to strings, numbers, and booleans, two other *primitive* values in JS programs are `null` and `undefined`. While there are differences between them (some historic and some contemporary), for the most part both values serve the purpose of indicating *emptiness* (or absence) of a value.

Many developers prefer to treat them both consistently in this fashion, which is to say that the values are assumed to be indistinguishable. If care is taken, this is often possible. However, it's safest and best to use only `undefined` as the single empty value, even though `null` seems attractive in that it's shorter to type!

```
while (value != undefined) {
  console.log("Still got something!");
}
```

The final primitive value to be aware of is a symbol, which is a special-purpose value that behaves as a hidden unguessable value. Symbols are almost exclusively used as special keys on objects:

```
hitchhikersGuide[ Symbol("meaning of life") ];
// 42
```

You won't encounter direct usage of symbols very often in typical JS programs. They're mostly used in low-level code such as in libraries and frameworks.

Arrays And Objects

Besides primitives, the other value type in JS is an object value.

As mentioned earlier, arrays are a special type of object that's comprised of an ordered and numerically indexed list of data:

```
var names = [ "Frank", "Kyle", "Peter", "Susan" ];

names.length;
// 4
```

```
names[0];  
// Frank
```

```
names[1];  
// Kyle
```

JS arrays can hold any value type, either primitive or object (including other arrays). As we'll see toward the end of Chapter 3, even functions are values that can be held in arrays or objects.

NOTE:

Functions, like arrays, are a special kind (aka, sub-type) of object. We'll cover functions in more detail in a bit.

Objects are more general: an unordered, keyed collection of any various values. In other words, you access the element by a string location name (aka "key" or "property") rather than by its numeric position (as with arrays). For example:

```
var me = {  
  first: "Kyle",  
  last: "Simpson",  
  age: 39,  
  specialties: [ "JS", "Table Tennis" ]  
};  
  
console.log(`My name is ${ me.first }.`);
```

Here, `me` represents an object, and `first` represents the name of a location of information in that object (value collection). Another syntax option that accesses information in an object by its property/key uses the square-brackets `[]`, such as `me["first"]`.

Value Type Determination

For distinguishing values, the `typeof` operator tells you its built-in type, if primitive, or "object" otherwise:

```
typeof 42;          // "number"  
typeof "abc";       // "string"  
typeof true;        // "boolean"  
typeof undefined;   // "undefined"  
typeof null;        // "object" -- oops, bug!  
typeof { "a": 1 };  // "object"  
typeof [1,2,3];    // "object"  
typeof function hello(){}; // "function"
```

WARNING:

`typeof null` unfortunately returns "object" instead of the expected "null". Also, `typeof` returns the specific "function" for functions, but not the expected "array" for arrays.

Converting from one value type to another, such as from string to number, is referred to in JS as "coercion." We'll cover this in more detail later in this chapter.

Primitive values and object values behave differently when they're assigned or passed around. We'll cover these details in Appendix A, "Values vs References."

Declaring and Using Variables

To be explicit about something that may not have been obvious in the previous section: in JS programs, values can either appear as literal values (as many of the preceding examples illustrate), or they can be held in variables; think of variables as just containers for values.

Variables have to be declared (created) to be used. There are various syntax forms that declare variables (aka, "identifiers"), and each form has different implied behaviors.

For example, consider the `var` statement:

```
var myName = "Kyle";
var age;
```

The `var` keyword declares a variable to be used in that part of the program, and optionally allows an initial assignment of a value.

Another similar keyword is `let`:

```
let myName = "Kyle";
let age;
```

The `let` keyword has some differences to `var`, with the most obvious being that `let` allows a more limited access to the variable than `var`. This is called "block scoping" as opposed to regular or function scoping.

Consider:

```
var adult = true;

if (adult) {
  var myName = "Kyle";
  let age = 39;
  console.log("Shhh, this is a secret!");
}
```

```
console.log(myName);
// Kyle

console.log(age);
// Error!
```

The attempt to access `age` outside of the `if` statement results in an error, because `age` was block-scoped to the `if`, whereas `myName` was not.

Block-scoping is very useful for limiting how widespread variable declarations are in our programs, which helps prevent accidental overlap of their names.

But `var` is still useful in that it communicates “this variable will be seen by a wider scope (of the whole function)”. Both declaration forms can be appropriate in any given part of a program, depending on the circumstances.

NOTE:

It’s very common to suggest that `var` should be avoided in favor of `let` (or `const!`), generally because of perceived confusion over how the scoping behavior of `var` has worked since the beginning of JS. I believe this to be overly restrictive advice and ultimately unhelpful. It’s assuming you are unable to learn and use a feature properly in combination with other features. I believe you *can* and *should* learn any features available, and use them where appropriate!

A third declaration form is `const`. It’s like `let` but has an additional limitation that it must be given a value at the moment it’s declared, and cannot be re-assigned a different value later.

Consider:

```
const myBirthday = true;
let age = 39;

if (myBirthday) {
  age = age + 1;    // OK!
  myBirthday = false; // Error!
}
```

The `myBirthday` constant is not allowed to be re-assigned.

`const` declared variables are not “unchangeable”, they just cannot be re-assigned. It’s ill-advised to use `const` with object values, because those values can still be changed even though the variable can’t be re-assigned. This leads to potential confusion down the line, so I think it’s wise to avoid situations like:

```

const actors = [
  "Morgan Freeman", "Jennifer Aniston"
];

actors[2] = "Tom Cruise"; // OK :(
actors = []; // Error!

```

The best semantic use of a **const** is when you have a simple primitive value that you want to give a useful name to, such as using `myBirthday` instead of `true`. This makes programs easier to read.

TIP:

If you stick to using **const** only with primitive values, you avoid any confusion of re-assignment (not allowed) vs. mutation (allowed)! That's the safest and best way to use **const**.

Besides `var` / `let` / `const`, there are other syntactic forms that declare identifiers (variables) in various scopes. For example:

```

function hello(myName) {
  console.log(`Hello, ${ myName }.`);
}

hello("Kyle");
// Hello, Kyle.

```

The identifier `hello` is created in the outer scope, and it's also automatically associated so that it references the function. But the named parameter `myName` is created only inside the function, and thus is only accessible inside that function's scope. `hello` and `myName` generally behave as `var`-declared.

Another syntax that declares a variable is a `catch` clause:

```

try {
  someError();
}
catch (err) {
  console.log(err);
}

```

The `err` is a block-scoped variable that exists only inside the `catch` clause, as if it had been declared with `let`.

Functions

The word “function” has a variety of meanings in programming. For example, in the world of Functional Programming, “function” has a precise mathematical definition and implies a strict set of rules to abide by.

In JS, we should consider “function” to take the broader meaning of another related term: “procedure.” A procedure is a collection of statements that can be invoked one or more times, may be provided some inputs, and may give back one or more outputs.

From the early days of JS, function definition looked like:

```
function awesomeFunction(coolThings) {  
  // ...  
  return amazingStuff;  
}
```

This is called a function declaration because it appears as a statement by itself, not as an expression in another statement. The association between the identifier `awesomeFunction` and the function value happens during the compile phase of the code, before that code is executed.

In contrast to a function declaration statement, a function expression can be defined and assigned like this:

```
// let awesomeFunction = ...  
// const awesomeFunction = ...  
var awesomeFunction = function(coolThings) {  
  // ...  
  return amazingStuff;  
};
```

This function is an expression that is assigned to the variable `awesomeFunction`. Different from the function declaration form, a function expression is not associated with its identifier until that statement during runtime.

It’s extremely important to note that in JS, functions are values that can be assigned (as shown in this snippet) and passed around. In fact, JS functions are a special type of the object value type. Not all languages treat functions as values, but it’s essential for a language to support the functional programming pattern, as JS does.

JS functions can receive parameter input:

```
function greeting(myName) {  
  console.log(`Hello, ${myName}!`);  
}  
  
greeting("Kyle"); // Hello, Kyle!
```

In this snippet, `myName` is called a parameter, which acts as a local variable inside the function. Functions can be defined to receive any number of parameters, from none upward, as you see fit. Each parameter is assigned the argument value that you pass in that position ("Kyle", here) of the call.

Functions also can return values using the `return` keyword:

```

function greeting(myName) {
  return `Hello, ${ myName }!`;
}

var msg = greeting("Kyle");

```

```
console.log(msg); // Hello, Kyle!
```

You can only `return` a single value, but if you have more values to return, you can wrap them up into a single object/array.

Since functions are values, they can be assigned as properties on objects:

```

var whatToSay = {
  greeting() {
    console.log("Hello!");
  },
  question() {
    console.log("What's your name?");
  },
  answer() {
    console.log("My name is Kyle.");
  }
};

whatToSay.greeting();
// Hello!

```

In this snippet, references to three functions (`greeting()`, `question()`, and `answer()`) are included in the object held by `whatToSay`. Each function can be called by accessing the property to retrieve the function reference value. Compare this straightforward style of defining functions on an object to the more sophisticated `class` syntax discussed later in this chapter.

There are many varied forms that `functions` take in JS. We dig into these variations in Appendix A, “So Many Function Forms.”

Comparisons

Making decisions in programs requires comparing values to determine their identity and relationship to each other. JS has several mechanisms to enable value comparison, so let’s take a closer look at them.

Equal...ish

The most common comparison in JS programs asks the question, “Is this X value *the same as* that Y value?” What exactly does “the same as” really mean to JS, though?

For ergonomic and historical reasons, the meaning is more complicated than the obvious *exact identity* sort of matching. Sometimes an equality comparison intends *exact* matching, but other times the desired comparison is a bit broader, allowing *closely similar* or *interchangeable* matching. In other words, we must be aware of the nuanced differences between an **equality** comparison and an **equivalence** comparison.

If you've spent any time working with and reading about JS, you've certainly seen the so-called "triple-equals" `==` operator, also described as the "strict equality" operator. That seems rather straightforward, right? Surely, "strict" means strict, as in narrow and *exact*.

Not *exactly*.

Yes, most values participating in an `==` equality comparison will fit with that *exact same* intuition. Consider some examples:

```
3 === 3.0;          // true
"yes" === "yes";    // true
null === null;      // true
false === false;    // true

42 === "42";        // false
"hello" === "Hello"; // false
true === 1;          // false
0 === null;          // false
"" === null;         // false
null === undefined; // false
```

NOTE:

Another way `==`'s equality comparison is often described is, "checking both the value and the type". In several of the examples we've looked at so far, like `42 === "42"`, the *type* of both values (number, string, etc.) does seem to be the distinguishing factor. There's more to it than that, though. **All** value comparisons in JS consider the type of the values being compared, not *just* the `==` operator. Specifically, `==` disallows any sort of type conversion (aka, "coercion") in its comparison, where other JS comparisons *do* allow coercion.

But the `==` operator does have some nuance to it, a fact many JS developers gloss over, to their detriment. The `==` operator is designed to *lie* in two cases of special values: `NaN` and `-0`. Consider:

```
NaN === NaN;        // false
0 === -0;            // true
```

In the case of `NaN`, the `==` operator *lies* and says that an occurrence of `NaN` is not equal to another `NaN`. In the case of `-0` (yes, this is a real, distinct value

you can use intentionally in your programs!), the `==` operator *lies* and says it's equal to the regular `0` value.

Since the *lying* about such comparisons can be bothersome, it's best to avoid using `==` for them. For `NaN` comparisons, use the `Number.isNaN(..)` utility, which does not *lie*. For `-0` comparison, use the `Object.is(..)` utility, which also does not *lie*. `Object.is(..)` can also be used for non-*lying* `NaN` checks, if you prefer. Humorously, you could think of `Object.is(..)` as the “quadruple-equals” `=====`, the really-really-strict comparison!

There are deeper historical and technical reasons for these *lies*, but that doesn't change the fact that `==` is not actually *strictly exactly equal* comparison, in the *strictest* sense.

The story gets even more complicated when we consider comparisons of object values (non-primitives). Consider:

```
[ 1, 2, 3 ] === [ 1, 2, 3 ];      // false
{ a: 42 } === { a: 42 }           // false
(x => x * 2) === (x => x * 2)   // false
```

What's going on here?

It may seem reasonable to assume that an equality check considers the *nature* or *contents* of the value; after all, `42 == 42` considers the actual `42` value and compares it. But when it comes to objects, a content-aware comparison is generally referred to as “structural equality.”

JS does not define `==` as *structural equality* for object values. Instead, `==` uses *identity equality* for object values.

In JS, all object values are held by reference (see “Values vs References” in Appendix A), are assigned and passed by reference-copy, **and** to our current discussion, are compared by reference (identity) equality. Consider:

```
var x = [ 1, 2, 3 ];

// assignment is by reference-copy, so
// y references the *same* array as x,
// not another copy of it.
var y = x;

y === x;                  // true
y === [ 1, 2, 3 ];        // false
x === [ 1, 2, 3 ];        // false
```

In this snippet, `y == x` is true because both variables hold a reference to the same initial array. But the `== [1,2,3]` comparisons both fail because `y` and `x`, respectively, are being compared to new *different* arrays `[1,2,3]`. The array structure and contents don't matter in this comparison, only the **reference identity**.

JS does not provide a mechanism for structural equality comparison of object values, only reference identity comparison. To do structural equality comparison, you'll need to implement the checks yourself.

But beware, it's more complicated than you'll assume. For example, how might you determine if two function references are "structurally equivalent"? Even stringifying to compare their source code text wouldn't take into account things like closure. JS doesn't provide structural equality comparison because it's almost intractable to handle all the corner cases!

Coercive Comparisons

Coercion means a value of one type being converted to its respective representation in another type (to whatever extent possible). As we'll discuss in Chapter 4, coercion is a core pillar of the JS language, not some optional feature that can reasonably be avoided.

But where coercion meets comparison operators (like equality), confusion and frustration unfortunately crop up more often than not.

Few JS features draw more ire in the broader JS community than the `==` operator, generally referred to as the "loose equality" operator. The majority of all writing and public discourse on JS condemns this operator as poorly designed and dangerous/bug-ridden when used in JS programs. Even the creator of the language himself, Brendan Eich, has lamented how it was designed as a big mistake.

From what I can tell, most of this frustration comes from a pretty short list of confusing corner cases, but a deeper problem is the extremely widespread misconception that it performs its comparisons without considering the types of its compared values.

The `==` operator performs an equality comparison similarly to how the `===` performs it. In fact, both operators consider the type of the values being compared. And if the comparison is between the same value type, both `==` and `===` **do exactly the same thing, no difference whatsoever.**

If the value types being compared are different, the `==` differs from `===` in that it allows coercion before the comparison. In other words, they both want to compare values of like types, but `==` allows type conversions *first*, and once the types have been converted to be the same on both sides, then `==` does the same thing as `===`. Instead of "loose equality," the `==` operator should be described as "coercive equality."

Consider:

```
42 == "42";           // true
1 == true;            // true
```

In both comparisons, the value types are different, so the `==` causes the non-number values ("42" and `true`) to be converted to numbers (42 and 1, respectively) before the comparisons are made.

Just being aware of this nature of `==`—that it prefers primitive numeric comparisons—helps you avoid most of the troublesome corner cases, such as staying away from a gotchas like `"" == 0` or `0 == false`.

You may be thinking, “Oh, well, I will just always avoid any coercive equality comparison (using `==` instead) to avoid those corner cases”! Eh, sorry, that’s not quite as likely as you would hope.

There’s a pretty good chance that you’ll use relational comparison operators like `<`, `>` (and even `<=` and `>=`).

Just like `==`, these operators will perform as if they’re “strict” if the types being relationally compared already match, but they’ll allow coercion first (generally, to numbers) if the types differ.

Consider:

```
var arr = [ "1", "10", "100", "1000" ];
for (let i = 0; i < arr.length && arr[i] < 500; i++) {
  // will run 3 times
}
```

The `i < arr.length` comparison is “safe” from coercion because `i` and `arr.length` are always numbers. The `arr[i] < 500` invokes coercion, though, because the `arr[i]` values are all strings. Those comparisons thus become `1 < 500`, `10 < 500`, `100 < 500`, and `1000 < 500`. Since that fourth one is false, the loop stops after its third iteration.

These relational operators typically use numeric comparisons, except in the case where **both** values being compared are already strings; in this case, they use alphabetical (dictionary-like) comparison of the strings:

```
var x = "10";
var y = "9";

x < y;      // true, watch out!
```

There’s no way to get these relational operators to avoid coercion, other than to just never use mismatched types in the comparisons. That’s perhaps admirable as a goal, but it’s still pretty likely you’re going to run into a case where the types *may* differ.

The wiser approach is not to avoid coercive comparisons, but to embrace and learn their ins and outs.

Coercive comparisons crop up in other places in JS, such as conditionals (`if`, etc.), which we’ll revisit in Appendix A, “Coercive Conditional Comparison.”

How We Organize in JS

Two major patterns for organizing code (data and behavior) are used broadly across the JS ecosystem: classes and modules. These patterns are not mutually exclusive; many programs can and do use both. Other programs will stick with just one pattern, or even neither!

In some respects, these patterns are very different. But interestingly, in other ways, they're just different sides of the same coin. Being proficient in JS requires understanding both patterns and where they are appropriate (and not!).

Classes

The terms “object-oriented,” “class-oriented,” and “classes” are all very loaded full of detail and nuance; they're not universal in definition.

We will use a common and somewhat traditional definition here, the one most likely familiar to those with backgrounds in “object-oriented” languages like C++ and Java.

A class in a program is a definition of a “type” of custom data structure that includes both data and behaviors that operate on that data. Classes define how such a data structure works, but classes are not themselves concrete values. To get a concrete value that you can use in the program, a class must be *instantiated* (with the `new` keyword) one or more times.

Consider:

```
class Page {
  constructor(text) {
    this.text = text;
  }

  print() {
    console.log(this.text);
  }
}

class Notebook {
  constructor() {
    this.pages = [];
  }

  addPage(text) {
    var page = new Page(text);
    this.pages.push(page);
  }

  print() {
```

```

        for (let page of this.pages) {
            page.print();
        }
    }
}

var mathNotes = new Notebook();
mathNotes.addPage("Arithmetic: + - * / ...");
mathNotes.addPage("Trigonometry: sin cos tan ...");

mathNotes.print();
// ...

```

In the `Page` class, the data is a string of text stored in a `this.text` member property. The behavior is `print()`, a method that dumps the text to the console.

For the `Notebook` class, the data is an array of `Page` instances. The behavior is `addPage(..)`, a method that instantiates new `Page` pages and adds them to the list, as well as `print()` (which prints out all the pages in the notebook).

The statement `mathNotes = new Notebook()` creates an instance of the `Notebook` class, and `page = new Page(text)` is where instances of the `Page` class are created.

Behavior (methods) can only be called on instances (not the classes themselves), such as `mathNotes.addPage(..)` and `page.print()`.

The `class` mechanism allows packaging data (`text` and `pages`) to be organized together with their behaviors (e.g., `addPage(..)` and `print()`). The same program could have been built without any `class` definitions, but it would likely have been much less organized, harder to read and reason about, and more susceptible to bugs and subpar maintenance.

Class Inheritance Another aspect inherent to traditional “class-oriented” design, though a bit less commonly used in JS, is “inheritance” (and “polymorphism”). Consider:

```

class Publication {
    constructor(title,author,pubDate) {
        this.title = title;
        this.author = author;
        this.pubDate = pubDate;
    }

    print() {
        console.log(`Title: ${this.title}
        By: ${this.author}`)
    }
}

```

```

        ${ this.pubDate }
    );
}
}

```

This `Publication` class defines a set of common behavior that any publication might need.

Now let's consider more specific types of publication, like `Book` and `BlogPost`:

```

class Book extends Publication {
    constructor(bookDetails) {
        super(
            bookDetails.title,
            bookDetails.author,
            bookDetails.publishedOn
        );
        this.publisher = bookDetails.publisher;
        this.ISBN = bookDetails.ISBN;
    }

    print() {
        super.print();
        console.log(`
            Publisher: ${ this.publisher }
            ISBN: ${ this.ISBN }
        `);
    }
}

class BlogPost extends Publication {
    constructor(title,author,pubDate,URL) {
        super(title,author,pubDate);
        this.URL = URL;
    }

    print() {
        super.print();
        console.log(this.URL);
    }
}

```

Both `Book` and `BlogPost` use the `extends` clause to *extend* the general definition of `Publication` to include additional behavior. The `super(..)` call in each constructor delegates to the parent `Publication` class's constructor for its initialization work, and then they do more specific things according to their respective publication type (aka, “sub-class” or “child class”).

Now consider using these child classes:

```
var YDKJS = new Book({
  title: "You Don't Know JS",
  author: "Kyle Simpson",
  publishedOn: "June 2014",
  publisher: "O'Reilly",
  ISBN: "123456-789"
});

YDKJS.print();
// Title: You Don't Know JS
// By: Kyle Simpson
// June 2014
// Publisher: O'Reilly
// ISBN: 123456-789

var forAgainstLet = new BlogPost(
  "For and against let",
  "Kyle Simpson",
  "October 27, 2014",
  "https://davidwalsh.name/for-and-against-let"
);

forAgainstLet.print();
// Title: For and against let
// By: Kyle Simpson
// October 27, 2014
// https://davidwalsh.name/for-and-against-let
```

Notice that both child class instances have a `print()` method, which was an override of the *inherited* `print()` method from the parent `Publication` class. Each of those overridden child class `print()` methods call `super.print()` to invoke the inherited version of the `print()` method.

The fact that both the inherited and overridden methods can have the same name and co-exist is called *polymorphism*.

Inheritance is a powerful tool for organizing data/behavior in separate logical units (classes), but allowing the child class to cooperate with the parent by accessing/using its behavior and data.

Modules

The module pattern has essentially the same goal as the class pattern, which is to group data and behavior together into logical units. Also like classes, modules can “include” or “access” the data and behaviors of other modules, for cooperation’s sake.

But modules have some important differences from classes. Most notably, the syntax is entirely different.

Classic Modules ES6 added a module syntax form to native JS syntax, which we'll look at in a moment. But from the early days of JS, modules was an important and common pattern that was leveraged in countless JS programs, even without a dedicated syntax.

The key hallmarks of a *classic module* are an outer function (that runs at least once), which returns an “instance” of the module with one or more functions exposed that can operate on the module instance’s internal (hidden) data.

Because a module of this form is *just a function*, and calling it produces an “instance” of the module, another description for these functions is “module factories”.

Consider the classic module form of the earlier `Publication`, `Book`, and `BlogPost` classes:

```
function Publication(title,author,pubDate) {
  var publicAPI = {
    print() {
      console.log(`Title: ${ title }
      By: ${ author }
      ${ pubDate }
    `);
    }
  };

  return publicAPI;
}

function Book(bookDetails) {
  var pub = Publication(
    bookDetails.title,
    bookDetails.author,
    bookDetails.publishedOn
  );

  var publicAPI = {
    print() {
      pub.print();
      console.log(`Publisher: ${ bookDetails.publisher }
      ISBN: ${ bookDetails.ISBN }
    `);
    }
  };

  return publicAPI;
}
```

```

        }
    };

    return publicAPI;
}

function BlogPost(title,author,pubDate,URL) {
    var pub = Publication(title,author,pubDate);

    var publicAPI = {
        print() {
            pub.print();
            console.log(URL);
        }
    };

    return publicAPI;
}

```

Comparing these forms to the `class` forms, there are more similarities than differences.

The `class` form stores methods and data on an object instance, which must be accessed with the `this.` prefix. With modules, the methods and data are accessed as identifier variables in scope, without any `this.` prefix.

With `class`, the “API” of an instance is implicit in the class definition—also, all data and methods are public. With the module factory function, you explicitly create and return an object with any publicly exposed methods, and any data or other unreferenced methods remain private inside the factory function.

There are other variations to this factory function form that are quite common across JS, even in 2020; you may run across these forms in different JS programs: AMD (Asynchronous Module Definition), UMD (Universal Module Definition), and CommonJS (classic Node.js-style modules). The variations are minor (not quite compatible). However, all of these forms rely on the same basic principles.

Consider also the usage (aka, “instantiation”) of these module factory functions:

```

var YDKJS = Book({
    title: "You Don't Know JS",
    author: "Kyle Simpson",
    publishedOn: "June 2014",
    publisher: "O'Reilly",
    ISBN: "123456-789"
});

YDKJS.print();
// Title: You Don't Know JS

```

```

// By: Kyle Simpson
// June 2014
// Publisher: O'Reilly
// ISBN: 123456-789

var forAgainstLet = BlogPost(
  "For and against let",
  "Kyle Simpson",
  "October 27, 2014",
  "https://davidwalsh.name/for-and-against-let"
);

forAgainstLet.print();
// Title: For and against let
// By: Kyle Simpson
// October 27, 2014
// https://davidwalsh.name/for-and-against-let

```

The only observable difference here is the lack of using `new`, calling the module factories as normal functions.

ES Modules ES modules (ESM), introduced to the JS language in ES6, are meant to serve much the same spirit and purpose as the existing *classic modules* just described, especially taking into account important variations and use cases from AMD, UMD, and CommonJS.

The implementation approach does, however, differ significantly.

First, there's no wrapping function to *define* a module. The wrapping context is a file. ESMs are always file-based; one file, one module.

Second, you don't interact with a module's "API" explicitly, but rather use the `export` keyword to add a variable or method to its public API definition. If something is defined in a module but not `exported`, then it stays hidden (just as with *classic modules*).

Third, and maybe most noticeably different from previously discussed patterns, you don't "instantiate" an ES module, you just `import` it to use its single instance. ESMs are, in effect, "singletons," in that there's only one instance ever created, at first `import` in your program, and all other `imports` just receive a reference to that same single instance. If your module needs to support multiple instantiations, you have to provide a *classic module-style* factory function on your ESM definition for that purpose.

In our running example, we do assume multiple-instantiation, so these following snippets will mix both ESM and *classic modules*.

Consider the file `publication.js`:

```

function printDetails(title,author,pubDate) {
  console.log(`
    Title: ${ title }
    By: ${ author }
    ${ pubDate }
  `);
}

export function create(title,author,pubDate) {
  var publicAPI = {
    print() {
      printDetails(title,author,pubDate);
    }
  };

  return publicAPI;
}

```

To import and use this module, from another ES module like `blogpost.js`:

```

import { create as createPub } from "publication.js";

function printDetails(pub,URL) {
  pub.print();
  console.log(URL);
}

export function create(title,author,pubDate,URL) {
  var pub = createPub(title,author,pubDate);

  var publicAPI = {
    print() {
      printDetails(pub,URL);
    }
  };

  return publicAPI;
}

```

And finally, to use this module, we import into another ES module like `main.js`:

```

import { create as newBlogPost } from "blogpost.js";

var forAgainstLet = newBlogPost(
  "For and against let",
  "Kyle Simpson",
  "October 27, 2014",
  "https://davidwalsh.name/for-and-against-let"

```

```
);

forAgainstLet.print();
// Title: For and against let
// By: Kyle Simpson
// October 27, 2014
// https://davidwalsh.name/for-and-against-let
```

NOTE:

The `as newBlogPost` clause in the `import` statement is optional; if omitted, a top-level function just named `create(..)` would be imported. In this case, I'm renaming it for readability's sake; its more generic factory name of `create(..)` becomes more semantically descriptive of its purpose as `newBlogPost(..)`.

As shown, ES modules can use *classic modules* internally if they need to support multiple-instantiation. Alternatively, we could have exposed a `class` from our module instead of a `create(..)` factory function, with generally the same outcome. However, since you're already using ESM at that point, I'd recommend sticking with *classic modules* instead of `class`.

If your module only needs a single instance, you can skip the extra layers of complexity: `export` its public methods directly.

The Rabbit Hole Deepens

As promised at the top of this chapter, we just glanced over a wide surface area of the main parts of the JS language. Your head may still be spinning, but that's entirely natural after such a firehose of information!

Even with just this “brief” survey of JS, we covered or hinted at a ton of details you should carefully consider and ensure you are comfortable with. I’m serious when I suggest: re-read this chapter, maybe several times.

In the next chapter, we’re going to dig much deeper into some important aspects of how JS works at its core. But before you follow that rabbit hole deeper, make sure you’ve taken adequate time to fully digest what we’ve just covered here.

You Don't Know JS Yet: Get Started - 2nd Edition

Chapter 3: Digging to the Roots of JS

If you've read Chapters 1 and 2, and taken the time to digest and percolate, you're hopefully starting to *get JS* a little more. If you skipped/skimmed them (especially Chapter 2), I recommend going back to spend some more time with that material.

In Chapter 2, we surveyed syntax, patterns, and behaviors at a high level. In this chapter, our attention shifts to some of the lower-level root characteristics of JS that underpin virtually every line of code we write.

Be aware: this chapter digs much deeper than you're likely used to thinking about a programming language. My goal is to help you appreciate the core of how JS works, what makes it tick. This chapter should begin to answer some of the "Why?" questions that may be cropping up as you explore JS. However, this material is still not an exhaustive exposition of the language; that's what the rest of the book series is for! Our goal here is still just to *get started*, and become more comfortable with, the *feel* of JS, how it ebbs and flows.

Don't run so quickly through this material that you get lost in the weeds. As I've said a dozen times already, **take your time**. Even still, you'll probably finish this chapter with remaining questions. That's OK, because there's a whole book series ahead of you to keep exploring!

Iteration

Since programs are essentially built to process data (and make decisions on that data), the patterns used to step through the data have a big impact on the program's readability.

The iterator pattern has been around for decades, and suggests a "standardized" approach to consuming data from a source one *chunk* at a time. The idea is that it's more common and helpful to iterate the data source—to progressively handle the collection of data by processing the first part, then the next, and so on, rather than handling the entire set all at once.

Imagine a data structure that represents a relational database `SELECT` query, which typically organizes the results as rows. If this query had only one or a couple of rows, you could handle the entire result set at once, and assign each row to a local variable, and perform whatever operations on that data that were appropriate.

But if the query has 100 or 1,000 (or more!) rows, you'll need iterative processing to deal with this data (typically, a loop).

The iterator pattern defines a data structure called an “iterator” that has a reference to an underlying data source (like the query result rows), which exposes a method like `next()`. Calling `next()` returns the next piece of data (i.e., a “record” or “row” from a database query).

You don’t always know how many pieces of data that you will need to iterate through, so the pattern typically indicates completion by some special value or exception once you iterate through the entire set and *go past the end*.

The importance of the iterator pattern is in adhering to a *standard* way of processing data iteratively, which creates cleaner and easier to understand code, as opposed to having every data structure/source define its own custom way of handling its data.

After many years of various JS community efforts around mutually agreed-upon iteration techniques, ES6 standardized a specific protocol for the iterator pattern directly in the language. The protocol defines a `next()` method whose return is an object called an *iterator result*; the object has `value` and `done` properties, where `done` is a boolean that is `false` until the iteration over the underlying data source is complete.

Consuming Iterators

With the ES6 iteration protocol in place, it’s workable to consume a data source one value at a time, checking after each `next()` call for `done` to be `true` to stop the iteration. But this approach is rather manual, so ES6 also included several mechanisms (syntax and APIs) for standardized consumption of these iterators.

One such mechanism is the `for..of` loop:

```
// given an iterator of some data source:  
var it = /* .. */;  
  
// loop over its results one at a time  
for (let val of it) {  
    console.log(`Iterator value: ${ val }`);  
}  
// Iterator value: ..  
// Iterator value: ..  
// ..
```

NOTE:

We’ll omit the manual loop equivalent here, but it’s definitely less readable than the `for..of` loop!

Another mechanism that’s often used for consuming iterators is the `...` operator. This operator actually has two symmetrical forms: *spread* and *rest* (or *gather*, as I prefer). The *spread* form is an iterator-consumer.

To *spread* an iterator, you have to have *something* to spread it into. There are two possibilities in JS: an array or an argument list for a function call.

An array spread:

```
// spread an iterator into an array,
// with each iterated value occupying
// an array element position.
var vals = [ ...it ];
```

A function call spread:

```
// spread an iterator into a function,
// call with each iterated value
// occupying an argument position.
doSomethingUseful( ...it );
```

In both cases, the iterator-spread form of `...` follows the iterator-consumption protocol (the same as the `for..of` loop) to retrieve all available values from an iterator and place (aka, spread) them into the receiving context (array, argument list).

Iterables

The iterator-consumption protocol is technically defined for consuming *iterables*; an iterable is a value that can be iterated over.

The protocol automatically creates an iterator instance from an iterable, and consumes *just that iterator instance* to its completion. This means a single iterable could be consumed more than once; each time, a new iterator instance would be created and used.

So where do we find iterables?

ES6 defined the basic data structure/collection types in JS as iterables. This includes strings, arrays, maps, sets, and others.

Consider:

```
// an array is an iterable
var arr = [ 10, 20, 30 ];

for (let val of arr) {
    console.log(`Array value: ${val}`);
}

// Array value: 10
// Array value: 20
// Array value: 30
```

Since arrays are iterables, we can shallow-copy an array using iterator consumption via the `...` spread operator:

```
var arrCopy = [ ...arr ];
```

We can also iterate the characters in a string one at a time:

```
var greeting = "Hello world!";
var chars = [ ...greeting ];
```

```
chars;
// [ "H", "e", "l", "l", "o", " ", 
//   "w", "o", "r", "l", "d", "!" ]
```

A Map data structure uses objects as keys, associating a value (of any type) with that object. Maps have a different default iteration than seen here, in that the iteration is not just over the map's values but instead its *entries*. An *entry* is a tuple (2-element array) including both a key and a value.

Consider:

```
// given two DOM elements, `btn1` and `btn2`

var buttonNames = new Map();
buttonNames.set(btn1, "Button 1");
buttonNames.set(btn2, "Button 2");

for (let [btn,btnName] of buttonNames) {
  btn.addEventListener("click",function onClick(){
    console.log(`Clicked ${ btnName }`);
  });
}
```

In the `for..of` loop over the default map iteration, we use the `[btn,btnName]` syntax (called “array destructuring”) to break down each consumed tuple into the respective key/value pairs (`btn1 / "Button 1"` and `btn2 / "Button 2"`).

Each of the built-in iterables in JS expose a default iteration, one which likely matches your intuition. But you can also choose a more specific iteration if necessary. For example, if we want to consume only the values of the above `buttonNames` map, we can call `values()` to get a values-only iterator:

```
for (let btnName of buttonNames.values()) {
  console.log(btnName);
}
// Button 1
// Button 2
```

Or if we want the index *and* value in an array iteration, we can make an *entries* iterator with the `entries()` method:

```
var arr = [ 10, 20, 30 ];

for (let [idx,val] of arr.entries()) {
```

```
    console.log(`[${idx}]: ${val}`);
}
// [0]: 10
// [1]: 20
// [2]: 30
```

For the most part, all built-in iterables in JS have three iterator forms available: keys-only (`keys()`), values-only (`values()`), and entries (`entries()`).

Beyond just using built-in iterables, you can also ensure your own data structures adhere to the iteration protocol; doing so means you opt into the ability to consume your data with `for..of` loops and the `...` operator. “Standardizing” on this protocol means code that is overall more readily recognizable and readable.

NOTE:

You may have noticed a nuanced shift that occurred in this discussion. We started by talking about consuming **iterators**, but then switched to talking about iterating over **iterables**. The iteration-consumption protocol expects an *iterable*, but the reason we can provide a direct *iterator* is that an iterator is just an iterable of itself! When creating an iterator instance from an existing iterator, the iterator itself is returned.

Closure

Perhaps without realizing it, almost every JS developer has made use of closure. In fact, closure is one of the most pervasive programming functionalities across a majority of languages. It might even be as important to understand as variables or loops; that's how fundamental it is.

Yet it feels kind of hidden, almost magical. And it's often talked about in either very abstract or very informal terms, which does little to help us nail down exactly what it is.

We need to be able to recognize where closure is used in programs, as the presence or lack of closure is sometimes the cause of bugs (or even the cause of performance issues).

So let's define closure in a pragmatic and concrete way:

Closure is when a function remembers and continues to access variables from outside its scope, even when the function is executed in a different scope.

We see two definitional characteristics here. First, closure is part of the nature of a function. Objects don't get closures, functions do. Second, to observe a closure, you must execute a function in a different scope than where that function was originally defined.

Consider:

```
function greeting(msg) {
  return function who(name) {
    console.log(`#${ msg }, ${ name }!`);
  };
}

var hello = greeting("Hello");
var howdy = greeting("Howdy");

hello("Kyle");
// Hello, Kyle!

hello("Sarah");
// Hello, Sarah!

howdy("Grant");
// Howdy, Grant!
```

First, the `greeting(..)` outer function is executed, creating an instance of the inner function `who(..)`; that function closes over the variable `msg`, which is the parameter from the outer scope of `greeting(..)`. When that inner function is returned, its reference is assigned to the `hello` variable in the outer scope. Then we call `greeting(..)` a second time, creating a new inner function instance, with a new closure over a new `msg`, and return that reference to be assigned to `howdy`.

When the `greeting(..)` function finishes running, normally we would expect all of its variables to be garbage collected (removed from memory). We'd expect each `msg` to go away, but they don't. The reason is closure. Since the inner function instances are still alive (assigned to `hello` and `howdy`, respectively), their closures are still preserving the `msg` variables.

These closures are not a snapshot of the `msg` variable's value; they are a direct link and preservation of the variable itself. That means closure can actually observe (or make!) updates to these variables over time.

```
function counter(step = 1) {
  var count = 0;
  return function increaseCount(){
    count = count + step;
    return count;
  };
}

var incBy1 = counter(1);
var incBy3 = counter(3);
```

```

incBy1();      // 1
incBy1();      // 2

incBy3();      // 3
incBy3();      // 6
incBy3();      // 9

```

Each instance of the inner `increaseCount()` function is closed over both the `count` and `step` variables from its outer `counter(..)` function's scope. `step` remains the same over time, but `count` is updated on each invocation of that inner function. Since closure is over the variables and not just snapshots of the values, these updates are preserved.

Closure is most common when working with asynchronous code, such as with callbacks. Consider:

```

function getSomeData(url) {
  ajax(url, function onResponse(resp){
    console.log(
      `Response (from ${ url }): ${ resp }`
    );
  });
}

getSomeData("https://some.url/wherever");
// Response (from https://some.url/wherever): ...

```

The inner function `onResponse(..)` is closed over `url`, and thus preserves and remembers it until the Ajax call returns and executes `onResponse(..)`. Even though `getSomeData(..)` finishes right away, the `url` parameter variable is kept alive in the closure for as long as needed.

It's not necessary that the outer scope be a function—it usually is, but not always—just that there be at least one variable in an outer scope accessed from an inner function:

```

for (let [idx,btn] of buttons.entries()) {
  btn.addEventListener("click",function onClick(){
    console.log(`Clicked on button ${ idx }!`);
  });
}

```

Because this loop is using `let` declarations, each iteration gets new block-scoped (aka, local) `idx` and `btn` variables; the loop also creates a new inner `onClick(..)` function each time. That inner function closes over `idx`, preserving it for as long as the click handler is set on the `btn`. So when each button is clicked, its handler can print its associated index value, because the handler remembers its respective `idx` variable.

Remember: this closure is not over the value (like 1 or 3), but over the variable `idx` itself.

Closure is one of the most prevalent and important programming patterns in any language. But that's especially true of JS; it's hard to imagine doing anything useful without leveraging closure in one way or another.

If you're still feeling unclear or shaky about closure, the majority of Book 2, *Scope & Closures* is focused on the topic.

this Keyword

One of JS's most powerful mechanisms is also one of its most misunderstood: the `this` keyword. One common misconception is that a function's `this` refers to the function itself. Because of how `this` works in other languages, another misconception is that `this` points the instance that a method belongs to. Both are incorrect.

As discussed previously, when a function is defined, it is *attached* to its enclosing scope via closure. Scope is the set of rules that controls how references to variables are resolved.

But functions also have another characteristic besides their scope that influences what they can access. This characteristic is best described as an *execution context*, and it's exposed to the function via its `this` keyword.

Scope is static and contains a fixed set of variables available at the moment and location you define a function, but a function's execution *context* is dynamic, entirely dependent on **how it is called** (regardless of where it is defined or even called from).

`this` is not a fixed characteristic of a function based on the function's definition, but rather a dynamic characteristic that's determined each time the function is called.

One way to think about the *execution context* is that it's a tangible object whose properties are made available to a function while it executes. Compare that to scope, which can also be thought of as an *object*; except, the *scope object* is hidden inside the JS engine, it's always the same for that function, and its *properties* take the form of identifier variables available inside the function.

```
function classroom(teacher) {
  return function study() {
    console.log(
      `#${ teacher } says to study ${ this.topic }`
    );
  };
}
var assignment = classroom("Kyle");
```

The outer `classroom(..)` function makes no reference to a `this` keyword, so it's just like any other function we've seen so far. But the inner `study()` function does reference `this`, which makes it a `this`-aware function. In other words, it's a function that is dependent on its *execution context*.

NOTE:

`study()` is also closed over the `teacher` variable from its outer scope.

The inner `study()` function returned by `classroom("Kyle")` is assigned to a variable called `assignment`. So how can `assignment()` (aka `study()`) be called?

```
assignment();
// Kyle says to study undefined -- Oops :(
```

In this snippet, we call `assignment()` as a plain, normal function, without providing it any *execution context*.

Since this program is not in strict mode (see Chapter 1, “Strictly Speaking”), context-aware functions that are called **without any context specified** default the context to the global object (`window` in the browser). As there is no global variable named `topic` (and thus no such property on the global object), `this.topic` resolves to `undefined`.

Now consider:

```
var homework = {
  topic: "JS",
  assignment: assignment
};

homework.assignment();
// Kyle says to study JS
```

A copy of the `assignment` function reference is set as a property on the `homework` object, and then it's called as `homework.assignment()`. That means the `this` for that function call will be the `homework` object. Hence, `this.topic` resolves to "JS".

Lastly:

```
var otherHomework = {
  topic: "Math"
};

assignment.call(otherHomework);
// Kyle says to study Math
```

A third way to invoke a function is with the `call(..)` method, which takes an object (`otherHomework` here) to use for setting the `this` reference for the function call. The property reference `this.topic` resolves to "Math".

The same context-aware function invoked three different ways, gives different answers each time for what object `this` will reference.

The benefit of `this`-aware functions—and their dynamic context—is the ability to more flexibly re-use a single function with data from different objects. A function that closes over a scope can never reference a different scope or set of variables. But a function that has dynamic `this` context awareness can be quite helpful for certain tasks.

Prototypes

Where `this` is a characteristic of function execution, a prototype is a characteristic of an object, and specifically resolution of a property access.

Think about a prototype as a linkage between two objects; the linkage is hidden behind the scenes, though there are ways to expose and observe it. This prototype linkage occurs when an object is created; it's linked to another object that already exists.

A series of objects linked together via prototypes is called the "prototype chain."

The purpose of this prototype linkage (i.e., from an object B to another object A) is so that accesses against B for properties/methods that B does not have, are *delegated* to A to handle. Delegation of property/method access allows two (or more!) objects to cooperate with each other to perform a task.

Consider defining an object as a normal literal:

```
var homework = {  
    topic: "JS"  
};
```

The `homework` object only has a single property on it: `topic`. However, its default prototype linkage connects to the `Object.prototype` object, which has common built-in methods on it like `toString()` and `valueOf()`, among others.

We can observe this prototype linkage *delegation* from `homework` to `Object.prototype`:

```
homework.toString(); // [object Object]
```

`homework.toString()` works even though `homework` doesn't have a `toString()` method defined; the delegation invokes `Object.prototype.toString()` instead.

Object Linkage

To define an object prototype linkage, you can create the object using the `Object.create(..)` utility:

```
var homework = {  
    topic: "JS"  
};  
  
var otherHomework = Object.create(homework);  
  
otherHomework.topic; // "JS"
```

The first argument to `Object.create(..)` specifies an object to link the newly created object to, and then returns the newly created (and linked!) object.

Figure 4 shows how the three objects (`otherHomework`, `homework`, and `Object.prototype`) are linked in a prototype chain:

Fig. 4: Objects in a prototype chain

Delegation through the prototype chain only applies for accesses to lookup the value in a property. If you assign to a property of an object, that will apply directly to the object regardless of where that object is prototype linked to.

TIP:

`Object.create(null)` creates an object that is not prototype linked anywhere, so it's purely just a standalone object; in some circumstances, that may be preferable.

Consider:

```
homework.topic;  
// "JS"  
  
otherHomework.topic;  
// "JS"  
  
otherHomework.topic = "Math";  
otherHomework.topic;  
// "Math"  
  
homework.topic;  
// "JS" -- not "Math"
```

The assignment to `topic` creates a property of that name directly on `otherHomework`; there's no effect on the `topic` property on `homework`.

The next statement then accesses `otherHomework.topic`, and we see the non-delegated answer from that new property: "Math".

Figure 5 shows the objects/properties after the assignment that creates the `otherHomework.topic` property:

Fig. 5: Shadowed property 'topic'

The `topic` on `otherHomework` is "shadowing" the property of the same name on the `homework` object in the chain.

NOTE:

Another frankly more convoluted but perhaps still more common way of creating an object with a prototype linkage is using the "prototypal class" pattern, from before `class` (see Chapter 2, "Classes") was added in ES6. We'll cover this topic in more detail in Appendix A, "Prototypal 'Classes'".

this Revisited

We covered the `this` keyword earlier, but its true importance shines when considering how it powers prototype-delegated function calls. Indeed, one of the main reasons `this` supports dynamic context based on how the function is called is so that method calls on objects which delegate through the prototype chain still maintain the expected `this`.

Consider:

```
var homework = {
  study() {
    console.log(`Please study ${ this.topic }`);
  }
};

var jsHomework = Object.create(homework);
jsHomework.topic = "JS";
jsHomework.study();
// Please study JS

var mathHomework = Object.create(homework);
mathHomework.topic = "Math";
mathHomework.study();
// Please study Math
```

The two objects `jsHomework` and `mathHomework` each prototype link to the single `homework` object, which has the `study()` function. `jsHomework` and `mathHomework` are each given their own `topic` property (see Figure 6).

Fig. 6: Two objects linked to a common parent

`jsHomework.study()` delegates to `homework.study()`, but its `this` (`this.topic`) for that execution resolves to `jsHomework` because of how the function is called, so `this.topic` is "JS". Similarly for `mathHomework.study()` delegating to `homework.study()` but still resolving `this` to `mathHomework`, and thus `this.topic` as "Math".

The preceding code snippet would be far less useful if `this` was resolved to `homework`. Yet, in many other languages, it would seem `this` would be `homework` because the `study()` method is indeed defined on `homework`.

Unlike many other languages, JS's `this` being dynamic is a critical component of allowing prototype delegation, and indeed `class`, to work as expected!

Asking "Why?"

The intended take-away from this chapter is that there's a lot more to JS under the hood than is obvious from glancing at the surface.

As you are *getting started* learning and knowing JS more closely, one of the most important skills you can practice and bolster is curiosity, and the art of asking "Why?" when you encounter something in the language.

Even though this chapter has gone quite deep on some of the topics, many details have still been entirely skimmed over. There's much more to learn here, and the path to that starts with you asking the *right* questions of your code. Asking the right questions is a critical skill of becoming a better developer.

In the final chapter of this book, we're going to briefly look at how JS is divided, as covered across the rest of the *You Don't Know JS Yet* book series. Also, don't skip Appendix B of this book, which has some practice code to review some of the main topics covered in this book.

You Don't Know JS Yet: Get Started - 2nd Edition

Chapter 4: The Bigger Picture

This book surveys what you need to be aware of as you *get started* with JS. The goal is to fill in gaps that readers newer to JS might have tripped over in their early encounters with the language. I also hope that we've hinted at enough deeper detail throughout to pique your curiosity to want to dig more into the language.

The rest of the books in this series are where we will unpack all of the rest of the language, in far greater detail than we could have done in a few brief chapters here.

Remember to take your time, though. Rather than rushing onto the next book in an attempt to churn through all the books expediently, spend some time going back over the material in this book. Spend some more time looking through code in your current projects, and comparing what you see to what's been discussed so far.

When you're ready, this final chapter divides the organization of the JS language into three main pillars, then offers a brief roadmap of what to expect from the rest of the book series, and how I suggest you proceed. Also, don't skip the appendices, especially Appendix B, "Practice, Practice, Practice!".

Pillar 1: Scope and Closure

The organization of variables into units of scope (functions, blocks) is one of the most foundational characteristics of any language; perhaps no other characteristic has a greater impact on how programs behave.

Scopes are like buckets, and variables are like marbles you put into those buckets. The scope model of a language is like the rules that help you determine which color marbles go in which matching-color buckets.

Scopes nest inside each other, and for any given expression or statement, only variables at that level of scope nesting, or in higher/outer scopes, are accessible; variables from lower/inner scopes are hidden and inaccessible.

This is how scopes behave in most languages, which is called lexical scope. The scope unit boundaries, and how variables are organized in them, is determined at the time the program is parsed (compiled). In other words, it's an author-time decision: where you locate a function/scope in the program determines what the scope structure of that part of the program will be.

JS is lexically scoped, though many claim it isn't, because of two particular characteristics of its model that are not present in other lexically scoped languages.

The first is commonly called *hoisting*: when all variables declared anywhere in a scope are treated as if they're declared at the beginning of the scope. The other is that `var`-declared variables are function scoped, even if they appear inside a block.

Neither hoisting nor function-scoped `var` are sufficient to back the claim that JS is not lexically scoped. `let`/`const` declarations have a peculiar error behavior called the “Temporal Dead Zone” (TDZ) which results in observable but unusable variables. Though TDZ can be strange to encounter, it's *also* not an invalidation of lexical scoping. All of these are just unique parts of the language that should be learned and understood by all JS developers.

Closure is a natural result of lexical scope when the language has functions as first-class values, as JS does. When a function makes reference to variables from an outer scope, and that function is passed around as a value and executed in other scopes, it maintains access to its original scope variables; this is closure.

Across all of programming, but especially in JS, closure drives many of the most important programming patterns, including modules. As I see it, modules are as *with the grain* as you can get, when it comes to code organization in JS.

To dig further into scope, closures, and how modules work, read Book 2, *Scope & Closures*.

Pillar 2: Prototypes

The second pillar of the language is the prototypes system. We covered this topic in-depth in Chapter 3 (“Prototypes”), but I just want to make a few more comments about its importance.

JS is one of very few languages where you have the option to create objects directly and explicitly, without first defining their structure in a class.

For many years, people implemented the class design pattern on top of prototypes—so-called “prototypal inheritance” (see Appendix A, “Prototypal ‘Classes’”)—and then with the advent of ES6’s `class` keyword, the language doubled-down on its inclination toward OO/class-style programming.

But I think that focus has obscured the beauty and power of the prototype system: the ability for two objects to simply connect with each other and cooperate dynamically (during function/method execution) through sharing a `this` context.

Classes are just one pattern you can build on top of such power. But another approach, in a very different direction, is to simply embrace objects as objects, forget classes altogether, and let objects cooperate through the prototype chain. This is called *behavior delegation*. I think delegation is more powerful than class inheritance, as a means for organizing behavior and data in our programs.

But class inheritance gets almost all the attention. And the rest goes to functional programming (FP), as the sort of “anti-class” way of designing programs. This saddens me, because it snuffs out any chance for exploration of delegation as a viable alternative.

I encourage you to spend plenty of time deep in Book 3, *Objects & Classes*, to see how object delegation holds far more potential than we’ve perhaps realized. This isn’t an anti-class message, but it is intentionally a “classes aren’t the only way to use objects” message that I want more JS developers to consider.

Object delegation is, I would argue, far more *with the grain* of JS, than classes (more on *grains* in a bit).

Pillar 3: Types and Coercion

The third pillar of JS is by far the most overlooked part of JS’s nature.

The vast majority of developers have strong misconceptions about how *types* work in programming languages, and especially how they work in JS. A tidal wave of interest in the broader JS community has begun to shift to “static typing” approaches, using type-aware tooling like TypeScript or Flow.

I agree that JS developers should learn more about types, and should learn more about how JS manages type conversions. I also agree that type-aware tooling can help developers, assuming they have gained and used this knowledge in the first place!

But I don’t agree at all that the inevitable conclusion of this is to decide JS’s type mechanism is bad and that we need to cover up JS’s types with solutions outside the language. We don’t have to follow the “static typing” way to be smart and solid with types in our programs. There are other options, if you’re just willing to go *against the grain* of the crowd, and *with the grain* of JS (again, more on that to come).

Arguably, this pillar is more important than the other two, in the sense that no JS program will do anything useful if it doesn’t properly leverage JS’s value types, as well as the conversion (coercion) of values between types.

Even if you love TypeScript/Flow, you are not going to get the most out of those tools or coding approaches if you aren’t deeply familiar with how the language itself manages value types.

To learn more about JS types and coercion, check out Book 4, *Types & Grammar*. But please don’t skip over this topic just because you’ve always heard that we should use `==` and forget about the rest.

Without learning this pillar, your foundation in JS is shaky and incomplete at best.

With the Grain

I have some advice to share on continuing your learning journey with JS, and your path through the rest of this book series: be aware of the *grain* (recall various references to *grain* earlier in this chapter).

First, consider the *grain* (as in, wood) of how most people approach and use JS. You've probably already noticed that these books cut against that *grain* in many respects. In YDKJSY, I respect you the reader enough to explain all the parts of JS, not only some select popular parts. I believe you're both capable and deserving of that knowledge.

But that is not what you'll find from a lot of other material out there. It also means that the more you follow and adhere to the guidance from these books—that you think carefully and analyze for yourself what's best in your code—the more you will stand out. That can be a good and bad thing. If you ever want to break out from the crowd, you're going to have to break from how the crowd does it!

But I've also had many people tell me that they quoted some topic/explanation from these books during a job interview, and the interviewer told the candidate they were wrong; indeed, people have reportedly lost out on job offers as a result.

As much as possible, I endeavor in these books to provide completely accurate information about JS, informed generally from the specification itself. But I also dose out quite a bit of my opinions on how you can interpret and use JS to the best benefit in your programs. I don't present opinion as fact, or vice versa. You'll always know which is which in these books.

Facts about JS are not really up for debate. Either the specification says something, or it doesn't. If you don't like what the specification says, or my relaying of it, take that up with TC39! If you're in an interview and they claim you're wrong on the facts, ask them right then and there if you can look it up in the specification. If the interviewer won't re-consider, then you shouldn't want to work there anyway.

But if you choose to align with my opinions, you have to be prepared to back up those choices with *why* you feel that way. Don't just parrot what I say. Own your opinions. Defend them. And if someone you were hoping to work with disagrees, walk away with your head still held high. It's a big JS, and there's plenty of room for lots of different ways.

In other words, don't be afraid to go against the *grain*, as I have done with these books and all my teachings. Nobody can tell you how you will best make use of JS; that's for you to decide. I'm merely trying to empower you in coming to your own conclusions, no matter what they are.

On the other hand, there's a *grain* you really should pay attention to and follow: the *grain* of how JS works, at the language level. There are things that work

well and naturally in JS, given the right practice and approach, and there are things you really shouldn't try to do in the language.

Can you make your JS program look like a Java, C#, or Perl program? What about Python or Ruby, or even PHP? To varying degrees, sure you can. But should you?

No, I don't think you should. I think you should learn and embrace the JS way, and make your JS programs as JS'y as is practical. Some will think that means sloppy and informal programming, but I don't mean that at all. I just mean that JS has a lot of patterns and idioms that are recognizably "JS," and going with that *grain* is the general path to the best success.

Finally, maybe the most important *grain* to recognize is how the existing program(s) you're working on, and developers you're working with, do stuff. Don't read these books and then try to change *all that grain* in your existing projects over night. That approach will always fail.

You'll have to shift these things little by little, over time. Work on building consensus with your fellow developers on why it's important to re-visit and reconsider an approach. But do so with just one small topic at a time, and let before-and-after code comparisons do most of the talking. Bring everyone on the team together to discuss, and push for decisions that are based on analysis and evidence from the code rather than the inertia of "our senior devs have always done it this way."

That's the most important advice I can impart to help you learn JS. Always keep looking for better ways to use what JS gives us to author more readable code. Everyone who works on your code, including your future self, will thank you!

In Order

So now you've got a broader perspective on what's left to explore in JS, and the right attitude to approach the rest of your journey.

But one of the most common practical questions I get at this point is, "What order should I read the books?" There is a straightforward answer... but it also depends.

My suggestion for most readers is to proceed through this series in this order:

1. Get started with a solid foundation of JS from *Get Started* (Book 1) – good news, you've already almost finished this book!
2. In *Scope & Closures* (Book 2), dig into the first pillar of JS: lexical scope, how that supports closure, and how the module pattern organizes code.
3. In *Objects & Classes* (Book 3), focus on the second pillar of JS: how JS's `this` works, how object prototypes support delegation, and how prototypes enable the `class` mechanism for OO-style code organization.

4. In *Types & Grammar* (Book 4), tackle the third and final pillar of JS: types and type coercion, as well as how JS's syntax and grammar define how we write our code.
5. With the **three pillars** solidly in place, *Sync & Async* (Book 5) then explores how we use flow control to model state change in our programs, both synchronously (right away) and asynchronously (over time).
6. The series concludes with *ES.Next & Beyond* (Book 6), a forward look at the near- and mid-term future of JS, including a variety of features likely coming to your JS programs before too long.

That's the intended order to read this book series.

However, Books 2, 3, and 4 can generally be read in any order, depending on which topic you feel most curious about and comfortable exploring first. But I don't recommend you skip any of these three books—not even *Types & Grammar*, as some of you will be tempted to do!—even if you think you already have that topic down.

Book 5 (*Sync & Async*) is crucial for deeply understanding JS, but if you start digging in and find it's too intimidating, this book can be deferred until you're more experienced with the language. The more JS you've written (and struggled with!), the more you'll come to appreciate this book. So don't be afraid to come back to it at a later time.

The final book in the series, *ES.Next & Beyond*, in some respects stands alone. It can be read at the end, as I suggest, or right after *Getting Started* if you're looking for a shortcut to broaden your radar of what JS is all about. This book will also be more likely to receive updates in the future, so you'll probably want to re-visit it occasionally.

However you choose to proceed with YDKJSY, check out the appendices of this book first, especially practicing the snippets in Appendix B, “Practice, Practice, Practice!” Did I mention you should go practice!? There's no better way to learn code than to write it.

You Don't Know JS Yet: Get Started - 2nd Edition

Appendix A: Exploring Further

In this appendix, we're going to explore some topics from the main chapter text in a bit more detail. Think of this content as an optional preview of some of the more nuanced details covered throughout the rest of the book series.

Values vs. References

In Chapter 2, we introduced the two main types of values: primitives and objects. But we didn't discuss yet one key difference between the two: how these values are assigned and passed around.

In many languages, the developer can choose between assigning/passing a value as the value itself, or as a reference to the value. In JS, however, this decision is entirely determined by the kind of value. That surprises a lot of developers from other languages when they start using JS.

If you assign/pass a value itself, the value is copied. For example:

```
var myName = "Kyle";  
  
var yourName = myName;
```

Here, the `yourName` variable has a separate copy of the "Kyle" string from the value that's stored in `myName`. That's because the value is a primitive, and primitive values are always assigned/passed as **value copies**.

Here's how you can prove there's two separate values involved:

```
var myName = "Kyle";  
  
var yourName = myName;  
  
myName = "Frank";  
  
console.log(myName);  
// Frank  
  
console.log(yourName);  
// Kyle
```

See how `yourName` wasn't affected by the re-assignment of `myName` to "Frank"? That's because each variable holds its own copy of the value.

By contrast, references are the idea that two or more variables are pointing at the same value, such that modifying this shared value would be reflected by

access via any of those references. In JS, only object values (arrays, objects, functions, etc.) are treated as references.

Consider:

```
var myAddress = {  
  street: "123 JS Blvd",  
  city: "Austin",  
  state: "TX"  
};  
  
var yourAddress = myAddress;  
  
// I've got to move to a new house!  
myAddress.street = "456 TS Ave";  
  
console.log(yourAddress.street);  
// 456 TS Ave
```

Because the value assigned to `myAddress` is an object, it's held/assigned by reference, and thus the assignment to the `yourAddress` variable is a copy of the reference, not the object value itself. That's why the updated value assigned to the `myAddress.street` is reflected when we access `yourAddress.street`. `myAddress` and `yourAddress` have copies of the reference to the single shared object, so an update to one is an update to both.

Again, JS chooses the value-copy vs. reference-copy behavior based on the value type. Primitives are held by value, objects are held by reference. There's no way to override this in JS, in either direction.

So Many Function Forms

Recall this snippet from the “Functions” section in Chapter 2:

```
var awesomeFunction = function(coolThings) {  
  // ...  
  return amazingStuff;  
};
```

The function expression here is referred to as an *anonymous function expression*, since it has no name identifier between the `function` keyword and the `(..)` parameter list. This point confuses many JS developers because as of ES6, JS performs a “name inference” on an anonymous function:

```
awesomeFunction.name;  
// "awesomeFunction"
```

The `name` property of a function will reveal either its directly given name (in the case of a declaration) or its inferred name in the case of an anonymous function

expression. That value is generally used by developer tools when inspecting a function value or when reporting an error stack trace.

So even an anonymous function expression *might* get a name. However, name inference only happens in limited cases such as when the function expression is assigned (with `=`). If you pass a function expression as an argument to a function call, for example, no name inference occurs; the `name` property will be an empty string, and the developer console will usually report “(anonymous function)”.

Even if a name is inferred, **it’s still an anonymous function**. Why? Because the inferred name is a metadata string value, not an available identifier to refer to the function. An anonymous function doesn’t have an identifier to use to refer to itself from inside itself—for recursion, event unbinding, etc.

Compare the anonymous function expression form to:

```
// let awesomeFunction = ...
// const awesomeFunction = ...
var awesomeFunction = function someName(coolThings) {
  ...
  return amazingStuff;
};

awesomeFunction.name;
// "someName"
```

This function expression is a *named function expression*, since the identifier `someName` is directly associated with the function expression at compile time; the association with the identifier `awesomeFunction` still doesn’t happen until runtime at the time of that statement. Those two identifiers don’t have to match; sometimes it makes sense to have them be different, other times it’s better to have them be the same.

Notice also that the explicit function name, the identifier `someName`, takes precedence when assigning a *name* for the `name` property.

Should function expressions be named or anonymous? Opinions vary widely on this. Most developers tend to be unconcerned with using anonymous functions. They’re shorter, and unquestionably more common in the broad sphere of JS code out there.

In my opinion, if a function exists in your program, it has a purpose; otherwise, take it out! And if it has a purpose, it has a natural name that describes that purpose.

If a function has a name, you the code author should include that name in the code, so that the reader does not have to infer that name from reading and mentally executing that function’s source code. Even a trivial function body like `x * 2` has to be read to infer a name like “double” or “multBy2”; that brief extra mental work is unnecessary when you could just take a second to name the

function “double” or “multBy2” *once*, saving the reader that repeated mental work every time it’s read in the future.

There are, regrettably in some respects, many other function definition forms in JS as of early 2020 (maybe more in the future!).

Here are some more declaration forms:

```
// generator function declaration
function *two() { .. }

// async function declaration
async function three() { .. }

// async generator function declaration
async function *four() { .. }

// named function export declaration (ES6 modules)
export function five() { .. }
```

And here are some more of the (many!) function expression forms:

```
// IIFE
(function(){ .. })();
(function namedIIFE(){ .. })();

// asynchronous IIFE
(async function(){ .. })();
(async function namedAIIFE(){ .. })();

// arrow function expressions
var f;
f = () => 42;
f = x => x * 2;
f = (x) => x * 2;
f = (x,y) => x * y;
f = x => ({ x: x * 2 });
f = x => { return x * 2; };
f = async x => {
  var y = await doSomethingAsync(x);
  return y * 2;
};
someOperation( x => x * 2 );
// ..
```

Keep in mind that arrow function expressions are **syntactically anonymous**, meaning the syntax doesn’t provide a way to provide a direct name identifier for the function. The function expression may get an inferred name, but only

if it's one of the assignment forms, not in the (more common!) form of being passed as a function call argument (as in the last line of the snippet).

Since I don't think anonymous functions are a good idea to use frequently in your programs, I'm not a fan of using the `=>` arrow function form. This kind of function actually has a specific purpose (i.e., handling the `this` keyword lexically), but that doesn't mean we should use it for every function we write. Use the most appropriate tool for each job.

Functions can also be specified in class definitions and object literal definitions. They're typically referred to as "methods" when in these forms, though in JS this term doesn't have much observable difference over "function":

```
class SomethingKindaGreat {
  // class methods
  coolMethod() { .. }  // no commas!
  boringMethod() { .. }
}

var EntirelyDifferent = {
  // object methods
  coolMethod() { .. },  // commas!
  boringMethod() { .. },

  // (anonymous) function expression property
  oldSchool: function() { .. }
};
```

Phew! That's a lot of different ways to define functions.

There's no simple shortcut path here; you just have to build familiarity with all the function forms so you can recognize them in existing code and use them appropriately in the code you write. Study them closely and practice!

Coercive Conditional Comparison

Yes, that section name is quite a mouthful. But what are we talking about? We're talking about conditional expressions needing to perform coercion-oriented comparisons to make their decisions.

`if` and `? :-ternary` statements, as well as the test clauses in `while` and `for` loops, all perform an implicit value comparison. But what sort? Is it "strict" or "coercive"? Both, actually.

Consider:

```
var x = 1;

if (x) {
  // will run!
```

```
}
```

```
while (x) {
  // will run, once!
  x = false;
}
```

You might think of these (x) conditional expressions like this:

```
var x = 1;

if (x == true) {
  // will run!
}

while (x == true) {
  // will run, once!
  x = false;
}
```

In this specific case – the value of x being 1 – that mental model works, but it's not accurate more broadly. Consider:

```
var x = "hello";

if (x) {
  // will run!
}

if (x == true) {
  // won't run :(
}
```

Oops. So what is the `if` statement actually doing? This is the more accurate mental model:

```
var x = "hello";

if (Boolean(x) == true) {
  // will run
}

// which is the same as:

if (Boolean(x) === true) {
  // will run
}
```

Since the `Boolean(..)` function always returns a value of type boolean, the

`==` vs `===` in this snippet is irrelevant; they'll both do the same thing. But the important part is to see that before the comparison, a coercion occurs, from whatever type `x` currently is, to boolean.

You just can't get away from coercions in JS comparisons. Buckle down and learn them.

Prototypal “Classes”

In Chapter 3, we introduced prototypes and showed how we can link objects through a prototype chain.

Another way of wiring up such prototype linkages served as the (honestly, ugly) predecessor to the elegance of the ES6 `class` system (see Chapter 2, “Classes”), and is referred to as prototypal classes.

TIP:

While this style of code is quite uncommon in JS these days, it's still perplexingly rather common to be asked about it in job interviews!

Let's first recall the `Object.create(...)` style of coding:

```
var Classroom = {
  welcome() {
    console.log("Welcome, students!");
  }
};

var mathClass = Object.create(Classroom);

mathClass.welcome();
// Welcome, students!
```

Here, a `mathClass` object is linked via its prototype to a `Classroom` object. Through this linkage, the function call `mathClass.welcome()` is delegated to the method defined on `Classroom`.

The prototypal class pattern would have labeled this delegation behavior “inheritance,” and alternatively have defined it (with the same behavior) as:

```
function Classroom() {
  // ...
}

Classroom.prototype.welcome = function hello() {
  console.log("Welcome, students!");
};
```

```
var mathClass = new Classroom();

mathClass.welcome();
// Welcome, students!
```

All functions by default reference an empty object at a property named `prototype`. Despite the confusing naming, this is **not** the function's *prototype* (where the function is prototype linked to), but rather the prototype object to *link to* when other objects are created by calling the function with `new`.

We add a `welcome` property on that empty object (called `Classroom.prototype`), pointing at the `hello()` function.

Then `new Classroom()` creates a new object (assigned to `mathClass`), and prototype links it to the existing `Classroom.prototype` object.

Though `mathClass` does not have a `welcome()` property/function, it successfully delegates to the function `Classroom.prototype.welcome()`.

This “prototypal class” pattern is now strongly discouraged, in favor of using ES6's `class` mechanism:

```
class Classroom {
  constructor() {
    // ..
  }

  welcome() {
    console.log("Welcome, students!");
  }
}

var mathClass = new Classroom();

mathClass.welcome();
// Welcome, students!
```

Under the covers, the same prototype linkage is wired up, but this `class` syntax fits the class-oriented design pattern much more cleanly than “prototypal classes”.

You Don't Know JS Yet: Get Started - 2nd Edition

Appendix B: Practice, Practice, Practice!

In this appendix, we'll explore some exercises and their suggested solutions. These are just to *get you started* with practice over the concepts from the book.

Practicing Comparisons

Let's practice working with value types and comparisons (Chapter 4, Pillar 3) where coercion will need to be involved.

`scheduleMeeting(..)` should take a start time (in 24-hour format as a string "hh:mm") and a meeting duration (number of minutes). It should return `true` if the meeting falls entirely within the work day (according to the times specified in `dayStart` and `dayEnd`); return `false` if the meeting violates the work day bounds.

```
const dayStart = "07:30";
const dayEnd = "17:45";

function scheduleMeeting(startTime,durationMinutes) {
    // ...TODO...
}

scheduleMeeting("7:00",15);      // false
scheduleMeeting("07:15",30);    // false
scheduleMeeting("7:30",30);     // true
scheduleMeeting("11:30",60);    // true
scheduleMeeting("17:00",45);    // true
scheduleMeeting("17:30",30);    // false
scheduleMeeting("18:00",15);    // false
```

Try to solve this yourself first. Consider the usage of equality and relational comparison operators, and how coercion impacts this code. Once you have code that works, *compare* your solution(s) to the code in "Suggested Solutions" at the end of this appendix.

Practicing Closure

Now let's practice with closure (Chapter 4, Pillar 1).

The `range(..)` function takes a number as its first argument, representing the first number in a desired range of numbers. The second argument is also a number representing the end of the desired range (inclusive). If the second

argument is omitted, then another function should be returned that expects that argument.

```
function range(start,end) {  
    // ...TODO..  
}  
  
range(3,3);      // [3]  
range(3,8);      // [3,4,5,6,7,8]  
range(3,0);      // []  
  
var start3 = range(3);  
var start4 = range(4);  
  
start3(3);      // [3]  
start3(8);      // [3,4,5,6,7,8]  
start3(0);      // []  
  
start4(6);      // [4,5,6]
```

Try to solve this yourself first.

Once you have code that works, *compare* your solution(s) to the code in “Suggested Solutions” at the end of this appendix.

Practicing Prototypes

Finally, let's work on `this` and objects linked via prototype (Chapter 4, Pillar 2).

Define a slot machine with three reels that can individually `spin()`, and then `display()` the current contents of all the reels.

The basic behavior of a single reel is defined in the `reel` object below. But the slot machine needs individual reels—objects that delegate to `reel`, and which each have a `position` property.

A reel only *knows how to display()* its current slot symbol, but a slot machine typically shows three symbols per reel: the current slot (**position**), one slot above (**position - 1**), and one slot below (**position + 1**). So displaying the slot machine should end up displaying a 3 x 3 grid of slot symbols.

```
function randMax(max) {
    return Math.trunc(1E9 * Math.random()) % max;
}

var reel = {
    symbols: [
        " ", " ", " ", " ", " ", " ", " ", " ", " ", " "
    ]
}
```

```

],
spin() {
  if (this.position == null) {
    this.position = randMax(
      this.symbols.length - 1
    );
  }
  this.position = (
    this.position + 100 + randMax(100)
  ) % this.symbols.length;
},
display() {
  if (this.position == null) {
    this.position = randMax(
      this.symbols.length - 1
    );
  }
  return this.symbols[this.position];
}
};

var slotMachine = {
  reels: [
    // this slot machine needs 3 separate reels
    // hint: Object.create(..)
  ],
  spin() {
    this.reels.forEach(function spinReel(reel){
      reel.spin();
    });
  },
  display() {
    // TODO
  }
};

slotMachine.spin();
slotMachine.display();
//  /  /
//  /  /
//  /  /

slotMachine.spin();
slotMachine.display();
//  /  /
//  /  /

```

// | |

Try to solve this yourself first.

Hints:

- Use the `%` modulo operator for wrapping `position` as you access symbols circularly around a reel.
- Use `Object.create(..)` to create an object and prototype-link it to another object. Once linked, delegation allows the objects to share `this` context during method invocation.
- Instead of modifying the reel object directly to show each of the three positions, you can use another temporary object (`Object.create(..)` again) with its own `position`, to delegate from.

Once you have code that works, *compare* your solution(s) to the code in “Suggested Solutions” at the end of this appendix.

Suggested Solutions

Keep in mind that these suggested solutions are just that: suggestions. There are many different ways to solve these practice exercises. Compare your approach to what you see here, and consider the pros and cons of each.

Suggested solution for “Comparisons” (Pillar 3) practice:

```
const dayStart = "07:30";
const dayEnd = "17:45";

function scheduleMeeting(startTime,durationMinutes) {
    var [ , meetingStartHour, meetingStartMinutes ] =
        startTime.match(/^(\d{1,2}):(\d{2})$/) || [];
    durationMinutes = Number(durationMinutes);

    if (
        typeof meetingStartHour == "string" &&
        typeof meetingStartMinutes == "string"
    ) {
        let durationHours =
            Math.floor(durationMinutes / 60);
        durationMinutes =
            durationMinutes - (durationHours * 60);
        let meetingEndHour =
            Number(meetingStartHour) + durationHours;
        let meetingEndMinutes =
            Number(meetingStartMinutes) +
            durationMinutes;
    }
}
```

```

        if (meetingEndMinutes >= 60) {
            meetingEndHour = meetingEndHour + 1;
            meetingEndMinutes =
                meetingEndMinutes - 60;
        }

        // re-compose fully-qualified time strings
        // (to make comparison easier)
        let meetingStart = `${
            meetingStartHour.padStart(2,"0")
        }: ${
            meetingStartMinutes.padStart(2,"0")
        }`;
        let meetingEnd = `${
            String(meetingEndHour).padStart(2,"0")
        }: ${
            String(meetingEndMinutes).padStart(2,"0")
        }`;

        // NOTE: since expressions are all strings,
        // comparisons here are alphabetic, but it's
        // safe here since they're fully qualified
        // time strings (ie, "07:15" < "07:30")
        return (
            meetingStart >= dayStart &&
            meetingEnd <= dayEnd
        );
    }

    return false;
}

scheduleMeeting("7:00",15);      // false
scheduleMeeting("07:15",30);    // false
scheduleMeeting("7:30",30);     // true
scheduleMeeting("11:30",60);    // true
scheduleMeeting("17:00",45);    // true
scheduleMeeting("17:30",30);    // false
scheduleMeeting("18:00",15);    // false

```

Suggested solution for “Closure” (Pillar 1) practice:

```

function range(start,end) {
    start = Number(start) || 0;

```

```

if (end === undefined) {
    return function getEnd(end) {
        return getRange(start, end);
    };
}
else {
    end = Number(end) || 0;
    return getRange(start, end);
}

// *****

function getRange(start, end) {
    var ret = [];
    for (let i = start; i <= end; i++) {
        ret.push(i);
    }
    return ret;
}

range(3,3);    // [3]
range(3,8);    // [3,4,5,6,7,8]
range(3,0);    // []

var start3 = range(3);
var start4 = range(4);

start3(3);    // [3]
start3(8);    // [3,4,5,6,7,8]
start3(0);    // []

start4(6);    // [4,5,6]

```

Suggested solution for “Prototypes” (Pillar 2) practice:

```

function randMax(max) {
    return Math.trunc(1E9 * Math.random()) % max;
}

var reel = {
    symbols: [
        "", "", "", "", "", "", "", "", ""
    ]
}

```

```

],
spin() {
  if (this.position == null) {
    this.position = randMax(
      this.symbols.length - 1
    );
  }
  this.position = (
    this.position + 100 + randMax(100)
  ) % this.symbols.length;
},
display() {
  if (this.position == null) {
    this.position = randMax(
      this.symbols.length - 1
    );
  }
  return this.symbols[this.position];
}
};

var slotMachine = {
  reels: [
    Object.create(reel),
    Object.create(reel),
    Object.create(reel)
  ],
  spin() {
    this.reels.forEach(function spinReel(reel){
      reel.spin();
    });
  },
  display() {
    var lines = [];

    // display all 3 lines on the slot machine
    for (
      let linePos = -1; linePos <= 1; linePos++
    ) {
      let line = this.reels.map(
        function getSlot(reel){
          var slot = Object.create(reel);
          slot.position = (
            reel.symbols.length +
            reel.position +
            linePos
          );
          return slot;
        }
      );
      lines.push(line);
    }
    return lines;
  }
};

```

```

        ) % reel.symbols.length;
        return slot.display();
    }
);
lines.push(line.join(" | "));
}

return lines.join("\n");
}
};

slotMachine.spin();
slotMachine.display();
// | |
// | |
// | |
slotMachine.spin();
slotMachine.display();
// | |
// | |
// | |

```

That's it for this book. But now it's time to look for real projects to practice these ideas on. Just keep coding, because that's the best way to learn!

You Don't Know JS Yet: Scope & Closures - 2nd Edition

Foreword

If I look over the books on my bookshelf, I can clearly see which of these titles are well loved. Well loved in this sense meaning they are a little worn. Their binding is broken, their pages are tattered, there might even be a spilled drink smear or two. What's ironic to me is that the most loved of my books often *look* the least cared for, though honestly the opposite is true.

Scope and Closures (1st ed.) is one of my most loved books. It's small, but the binding is coming undone. The pages are worn and dog-eared. It's a bit rumpled. It's not a book I've read once. I've picked it up again and again in the many years since it was originally published.

For me, it's also been a benchmark for my own personal progression through JavaScript. When I first read it in 2014, I was familiar with the concepts but the depth of my understanding was admittedly not as deep as the thin volume.

Over the years, even though I wasn't necessarily feeling my own improvement on a day-to-day basis, each one of the concepts became more approachable. I'd smile to myself, realizing how far I'd come with the help of these guides. It became apparent there was an inverse correlation between how well I treated the book and how much I loved it.

When Kyle asked me to write the Foreword for the 2nd edition, I was floored. It's not often you're asked to write about a book that's been so formative for your own understanding and career, *Scope and Closures* in particular. I remember the day I first understood closures, the first time I used one well. The satisfaction was great, in part because the symmetry of the idea was compelling to me. Before I even picked this book up, I was already enamoured with closures. And yet, there's a difference between being able to execute code successfully and fully explore the concepts with any depth. This book took my base understanding and drew it out, helped me master it.

This book is deceptively short. It's helpful that it's small because it's dense with useful knowledge. Since it is compact, I'd suggest you give yourself time to absorb each page. Take your time with it. Treat the book with care, and by that I mean, wear it down.

Sarah Drasner Head of Developer Experience Netlify

You Don't Know JS Yet - 2nd Edition

Preface

Welcome to the 2nd edition of the widely acclaimed *You Don't Know JS* (**YDKJS**) book series: *You Don't Know JS Yet* (**YDKJSY**).

If you've read any of the 1st edition books, you can expect a refreshed approach in these new ones, with plenty of updated coverage of what's changed in JS over the last five years. But what I hope and believe you'll still *get* is the same commitment to respecting JS and digging into what really makes it tick.

If this is your first time reading these books, I'm glad you're here. Prepare for a deep and extensive journey into all the corners of JavaScript.

If you are new to programming or JS, be aware that these books are not intended as a gentle "intro to JavaScript." This material is, at times, complex and challenging, and goes much deeper than is typical for a first-time learner. You're welcome here no matter what your background is, but these books are written assuming you're already comfortable with JS and have at least 6–9 months experience with it.

The Parts

These books approach JavaScript intentionally opposite of how *The Good Parts* treats the language. No, that doesn't mean we're looking at *the bad parts*, but rather, exploring **all the parts**.

You may have been told, or felt yourself, that JS is a deeply flawed language that was poorly designed and inconsistently implemented. Many have asserted that it's the worst most popular language in the world; that nobody writes JS because they want to, only because they have to given its place at the center of the web. That's a ridiculous, unhealthy, and wholly condescending claim.

Millions of developers write JavaScript every day, and many of them appreciate and respect the language.

Like any great language, it has its brilliant parts as well as its scars. Even the creator of JavaScript himself, Brendan Eich, laments some of those parts as mistakes. But he's wrong: they weren't mistakes at all. JS is what it is today—the world's most ubiquitous and thus most influential programming language—precisely because of *all those parts*.

Don't buy the lie that you should only learn and use a small collection of *good parts* while avoiding all the bad stuff. Don't buy the "X is the new Y" snake oil, that some new feature of the language instantly relegates all usage of a previous feature as obsolete and ignorant. Don't listen when someone says your code isn't "modern" because it isn't yet using a stage-0 feature that was only proposed a few weeks ago!

Every part of JS is useful. Some parts are more useful than others. Some parts require you to be more careful and intentional.

I find it absurd to try to be a truly effective JavaScript developer while only using a small sliver of what the language has to offer. Can you imagine a construction worker with a toolbox full of tools, who only uses their hammer and scoffs at the screwdriver or tape measure as inferior? That's just silly.

My unreserved claim is that you should go about learning all parts of JavaScript, and where appropriate, use them! And if I may be so bold as to suggest: it's time to discard any JS books that tell you otherwise.

The Title?

So what's the title of the series all about?

I'm not trying to insult you with criticism about your current lack of knowledge or understanding of JavaScript. I'm not suggesting you can't or won't be able to learn JavaScript. I'm not boasting about secret advanced insider wisdom that I and only a select few possess.

Seriously, all those were real reactions to the original series title before folks even read the books. And they're baseless.

The primary point of the title "You Don't Know JS Yet" is to point out that most JS developers don't take the time to really understand how the code that they write works. They know *that* it works—that it produces a desired outcome. But they either don't understand exactly *how*, or worse, they have an inaccurate mental model for the *how* that falters on closer scrutiny.

I'm presenting a gentle but earnest challenge to you the reader, to set aside the assumptions you have about JS, and approach it with fresh eyes and an invigorated curiosity that leads you to ask *why* for every line of code you write. Why does it do what it does? Why is one way better or more appropriate than the other half-dozen ways you could have accomplished it? Why do all the "popular kids" say to do X with your code, but it turns out that Y might be a better choice?

I added "Yet" to the title, not only because it's the second edition, but because ultimately I want these books to challenge you in a hopeful rather than discouraging way.

But let me be clear: I don't think it's possible to ever fully *know* JS. That's not an achievement to be obtained, but a goal to strive after. You don't finish knowing everything about JS, you just keep learning more and more as you spend more time with the language. And the deeper you go, the more you revisit what you *knew* before, and you re-learn it from that more experienced perspective.

I encourage you to adopt a mindset around JavaScript, and indeed all of software

development, that you will never fully have mastered it, but that you can and should keep working to get closer to that end, a journey that will stretch for the entirety of your software development career, and beyond.

You can always know JS better than you currently do. That's what I hope these YDKJSY books represent.

The Mission

The case doesn't really need to be made for why developers should take JS seriously—I think it's already more than proven worthy of first-class status among the world's programming languages.

But a different, more important case still needs to be made, and these books rise to that challenge.

I've taught more than 5,000 developers from teams and companies all over the world, in more than 25 countries on six continents. And what I've seen is that far too often, what *counts* is generally just the result of the program, not how the program is written or how/why it works.

My experience not only as a developer but in teaching many other developers tells me: you will always be more effective in your development work if you more completely understand how your code works than you are solely *just* getting it to produce a desired outcome.

In other words, *good enough to work* is not, and should not be, *good enough*.

All developers regularly struggle with some piece of code not working correctly, and they can't figure out why. But far too often, JS developers will blame this on the language rather than admitting it's their own understanding that is falling short. These books serve as both the question and answer: why did it do *this*, and here's how to get it to do *that* instead.

My mission with YDKJSY is to empower every single JS developer to fully own the code they write, to understand it and to write with intention and clarity.

The Path

Some of you have started reading this book with the goal of completing all six books, back to back.

I would like to caution you to consider changing that plan.

It is not my intention that YDKJSY be read straight through. The material in these books is dense, because JavaScript is powerful, sophisticated, and in parts rather complex. Nobody can really hope to *download* all this information to their brains in a single pass and retain any significant amount of it. That's unreasonable, and it's foolish to try.

My suggestion is you take your time going through YDKJSY. Take one chapter, read it completely through start to finish, and then go back and re-read it section by section. Stop in between each section, and practice the code or ideas from that section. For larger concepts, it probably is a good idea to expect to spend several days digesting, re-reading, practicing, then digesting some more.

You could spend a week or two on each chapter, and a month or two on each book, and a year or more on the whole series, and you would still not be squeezing every ounce of YDKJSY out.

Don't binge these books; be patient and spread out your reading. Interleave reading with lots of practice on real code in your job or on projects you participate in. Wrestle with the opinions I've presented along the way, debate with others, and most of all, disagree with me! Run a study group or book club. Teach mini-workshops at your office. Write blog posts on what you've learned. Speak about these topics at local JS meetups.

It's never my goal to convince you to agree with my opinion, but to encourage you to own and be able to defend your opinions. You can't get *there* with an expedient read-through of these books. That's something that takes a long while to emerge, little by little, as you study and ponder and re-visit.

These books are meant to be a field-guide on your wanderings through JavaScript, from wherever you currently are with the language, to a place of deeper understanding. And the deeper you understand JS, the more questions you will ask and the more you will have to explore! That's what I find so exciting!

I'm so glad you're embarking on this journey, and I am so honored you would consider and consult these books along the way. It's time to start *getting to know JS*.

You Don't Know JS Yet: Scope & Closures - 2nd Edition

Chapter 1: What's the Scope?

By the time you've written your first few programs, you're likely getting somewhat comfortable with creating variables and storing values in them. Working with variables is one of the most foundational things we do in programming!

But you may not have considered very closely the underlying mechanisms used by the engine to organize and manage these variables. I don't mean how the memory is allocated on the computer, but rather: how does JS know which variables are accessible by any given statement, and how does it handle two variables of the same name?

The answers to questions like these take the form of well-defined rules called scope. This book will dig through all aspects of scope—how it works, what it's useful for, gotchas to avoid—and then point toward common scope patterns that guide the structure of programs.

Our first step is to uncover how the JS engine processes our program **before** it runs.

About This Book

Welcome to book 2 in the *You Don't Know JS Yet* series! If you already finished *Get Started* (the first book), you're in the right spot! If not, before you proceed I encourage you to *start there* for the best foundation.

Our focus will be the first of three pillars in the JS language: the scope system and its function closures, as well as the power of the module design pattern.

JS is typically classified as an interpreted scripting language, so it's assumed by most that JS programs are processed in a single, top-down pass. But JS is in fact parsed/compiled in a separate phase **before execution begins**. The code author's decisions on where to place variables, functions, and blocks with respect to each other are analyzed according to the rules of scope, during the initial parsing/compilation phase. The resulting scope structure is generally unaffected by runtime conditions.

JS functions are themselves first-class values; they can be assigned and passed around just like numbers or strings. But since these functions hold and access variables, they maintain their original scope no matter where in the program the functions are eventually executed. This is called closure.

Modules are a code organization pattern characterized by public methods that have privileged access (via closure) to hidden variables and functions in the internal scope of the module.

Compiled vs. Interpreted

You may have heard of *code compilation* before, but perhaps it seems like a mysterious black box where source code slides in one end and executable programs pop out the other.

It's not mysterious or magical, though. Code compilation is a set of steps that process the text of your code and turn it into a list of instructions the computer can understand. Typically, the whole source code is transformed at once, and those resulting instructions are saved as output (usually in a file) that can later be executed.

You also may have heard that code can be *interpreted*, so how is that different from being *compiled*?

Interpretation performs a similar task to compilation, in that it transforms your program into machine-understandable instructions. But the processing model is different. Unlike a program being compiled all at once, with interpretation the source code is transformed line by line; each line or statement is executed before immediately proceeding to processing the next line of the source code.

Fig. 1: Compiled vs. Interpreted Code

Figure 1 illustrates compilation vs. interpretation of programs.

Are these two processing models mutually exclusive? Generally, yes. However, the issue is more nuanced, because interpretation can actually take other forms than just operating line by line on source code text. Modern JS engines actually employ numerous variations of both compilation and interpretation in the handling of JS programs.

Recall that we surveyed this topic in Chapter 1 of the *Get Started* book. Our conclusion there is that JS is most accurately portrayed as a **compiled language**. For the benefit of readers here, the following sections will revisit and expand on that assertion.

Compiling Code

But first, why does it even matter whether JS is compiled or not?

Scope is primarily determined during compilation, so understanding how compilation and execution relate is key in mastering scope.

In classic compiler theory, a program is processed by a compiler in three basic stages:

1. **Tokenizing/Lexing:** breaking up a string of characters into meaningful (to the language) chunks, called tokens. For instance, consider the program: `var a = 2;`. This program would likely be broken up into the following tokens: `var`, `a`, `=`, `2`, and `;`. Whitespace may or may not be persisted as a token, depending on whether it's meaningful or not.

(The difference between tokenizing and lexing is subtle and academic, but it centers on whether or not these tokens are identified in a *stateless* or *stateful* way. Put simply, if the tokenizer were to invoke stateful parsing rules to figure out whether `a` should be considered a distinct token or just part of another token, *that* would be **lexing**.)

2. **Parsing:** taking a stream (array) of tokens and turning it into a tree of nested elements, which collectively represent the grammatical structure of the program. This is called an Abstract Syntax Tree (AST).

For example, the tree for `var a = 2;` might start with a top-level node called `VariableDeclaration`, with a child node called `Identifier` (whose value is `a`), and another child called `AssignmentExpression` which itself has a child called `NumericLiteral` (whose value is `2`).

3. **Code Generation:** taking an AST and turning it into executable code. This part varies greatly depending on the language, the platform it's targeting, and other factors.

The JS engine takes the just described AST for `var a = 2;` and turns it into a set of machine instructions to actually *create* a variable called `a` (including reserving memory, etc.), and then store a value into `a`.

NOTE:

The implementation details of a JS engine (utilizing system memory resources, etc.) is much deeper than we will dig here. We'll keep our focus on the observable behavior of our programs and let the JS engine manage those deeper system-level abstractions.

The JS engine is vastly more complex than *just* these three stages. In the process of parsing and code generation, there are steps to optimize the performance of the execution (i.e., collapsing redundant elements). In fact, code can even be re-compiled and re-optimized during the progression of execution.

So, I'm painting only with broad strokes here. But you'll see shortly why *these* details we *do* cover, even at a high level, are relevant.

JS engines don't have the luxury of an abundance of time to perform their work and optimizations, because JS compilation doesn't happen in a build step ahead of time, as with other languages. It usually must happen in mere microseconds (or less!) right before the code is executed. To ensure the fastest performance under these constraints, JS engines use all kinds of tricks (like JITs, which lazy compile and even hot re-compile); these are well beyond the "scope" of our discussion here.

Required: Two Phases

To state it as simply as possible, the most important observation we can make about processing of JS programs is that it occurs in (at least) two phases: parsing/compilation first, then execution.

The separation of a parsing/compilation phase from the subsequent execution phase is observable fact, not theory or opinion. While the JS specification does not require “compilation” explicitly, it requires behavior that is essentially only practical with a compile-then-execute approach.

There are three program characteristics you can observe to prove this to yourself: syntax errors, early errors, and hoisting.

Syntax Errors from the Start Consider this program:

```
var greeting = "Hello";  
  
console.log(greeting);  
  
greeting = ."Hi";  
// SyntaxError: unexpected token .
```

This program produces no output (“Hello” is not printed), but instead throws a `SyntaxError` about the unexpected `.` token right before the `“Hi”` string. Since the syntax error happens after the well-formed `console.log(..)` statement, if JS was executing top-down line by line, one would expect the `“Hello”` message being printed before the syntax error being thrown. That doesn’t happen.

In fact, the only way the JS engine could know about the syntax error on the third line, before executing the first and second lines, is by the JS engine first parsing the entire program before any of it is executed.

Early Errors Next, consider:

```
console.log("Howdy");  
  
saySomething("Hello", "Hi");  
// Uncaught SyntaxError: Duplicate parameter name not  
// allowed in this context  
  
function saySomething(greeting, greeting) {  
    "use strict";  
    console.log(greeting);  
}
```

The `“Howdy”` message is not printed, despite being a well-formed statement.

Instead, just like the snippet in the previous section, the `SyntaxError` here is thrown before the program is executed. In this case, it’s because strict-mode

(opted in for only the `saySomething(..)` function here) forbids, among many other things, functions to have duplicate parameter names; this has always been allowed in non-strict-mode.

The error thrown is not a syntax error in the sense of being a malformed string of tokens (like `."Hi"` prior), but in strict-mode is nonetheless required by the specification to be thrown as an “early error” before any execution begins.

But how does the JS engine know that the `greeting` parameter has been duplicated? How does it know that the `saySomething(..)` function is even in strict-mode while processing the parameter list (the `"use strict"` pragma appears only later, in the function body)?

Again, the only reasonable explanation is that the code must first be *fully* parsed before any execution occurs.

Hoisting Finally, consider:

```
function saySomething() {
  var greeting = "Hello";
  {
    greeting = "Howdy"; // error comes from here
    let greeting = "Hi";
    console.log(greeting);
  }
}

saySomething();
// ReferenceError: Cannot access 'greeting' before
// initialization
```

The noted `ReferenceError` occurs from the line with the statement `greeting = "Howdy"`. What’s happening is that the `greeting` variable for that statement belongs to the declaration on the next line, `let greeting = "Hi"`, rather than to the previous `var greeting = "Hello"` statement.

The only way the JS engine could know, at the line where the error is thrown, that the *next statement* would declare a block-scoped variable of the same name (`greeting`) is if the JS engine had already processed this code in an earlier pass, and already set up all the scopes and their variable associations. This processing of scopes and declarations can only accurately be accomplished by parsing the program before execution.

The `ReferenceError` here technically comes from `greeting = "Howdy"` accessing the `greeting` variable **too early**, a conflict referred to as the Temporal Dead Zone (TDZ). Chapter 5 will cover this in more detail.

WARNING:

It's often asserted that `let` and `const` declarations are not hoisted, as an explanation of the TDZ behavior just illustrated. But this is not accurate. We'll come back and explain both the hoisting and TDZ of `let`/`const` in Chapter 5.

Hopefully you're now convinced that JS programs are parsed before any execution begins. But does it prove they are compiled?

This is an interesting question to ponder. Could JS parse a program, but then execute that program by *interpreting* operations represented in the AST **without** first compiling the program? Yes, that is *possible*. But it's extremely unlikely, mostly because it would be extremely inefficient performance wise.

It's hard to imagine a production-quality JS engine going to all the trouble of parsing a program into an AST, but not then converting (aka, "compiling") that AST into the most efficient (binary) representation for the engine to then execute.

Many have endeavored to split hairs with this terminology, as there's plenty of nuance and "well, actually..." interjections floating around. But in spirit and in practice, what the engine is doing in processing JS programs is **much more alike compilation** than not.

Classifying JS as a compiled language is not concerned with the distribution model for its binary (or byte-code) executable representations, but rather in keeping a clear distinction in our minds about the phase where JS code is processed and analyzed; this phase observably and indisputably happens *before* the code starts to be executed.

We need proper mental models of how the JS engine treats our code if we want to understand JS and scope effectively.

Compiler Speak

With awareness of the two-phase processing of a JS program (compile, then execute), let's turn our attention to how the JS engine identifies variables and determines the scopes of a program as it is compiled.

First, let's examine a simple JS program to use for analysis over the next several chapters:

```
var students = [
  { id: 14, name: "Kyle" },
  { id: 73, name: "Suzy" },
  { id: 112, name: "Frank" },
  { id: 6, name: "Sarah" }
];
```

```

function getStudentName(studentID) {
  for (let student of students) {
    if (student.id == studentID) {
      return student.name;
    }
  }
}

var nextStudent = getStudentName(73);

console.log(nextStudent);
// Suzy

```

Other than declarations, all occurrences of variables/identifiers in a program serve in one of two “roles”: either they’re the *target* of an assignment or they’re the *source* of a value.

(When I first learned compiler theory while earning my computer science degree, we were taught the terms “LHS” (aka, *target*) and “RHS” (aka, *source*) for these roles, respectively. As you might guess from the “L” and the “R”, the acronyms mean “Left-Hand Side” and “Right-Hand Side”, as in left and right sides of an `=` assignment operator. However, assignment targets and sources don’t always literally appear on the left or right of an `=`, so it’s probably clearer to think in terms of *target* / *source* rather than *left* / *right*.)

How do you know if a variable is a *target*? Check if there is a value that is being assigned to it; if so, it’s a *target*. If not, then the variable is a *source*.

For the JS engine to properly handle a program’s variables, it must first label each occurrence of a variable as *target* or *source*. We’ll dig in now to how each role is determined.

Targets

What makes a variable a *target*? Consider:

```
students = [ // ..
```

This statement is clearly an assignment operation; remember, the `var students` part is handled entirely as a declaration at compile time, and is thus irrelevant during execution; we left it out for clarity and focus. Same with the `nextStudent = getStudentName(73)` statement.

But there are three other *target* assignment operations in the code that are perhaps less obvious. One of them:

```
for (let student of students) {
```

That statement assigns a value to `student` for each iteration of the loop. Another *target* reference:

```
getStudentName(73)
```

But how is that an assignment to a *target*? Look closely: the argument `73` is assigned to the parameter `studentID`.

And there's one last (subtle) *target* reference in our program. Can you spot it?

```
..  
..  
..
```

Did you identify this one?

```
function getStudentName(studentID) {
```

A `function` declaration is a special case of a *target* reference. You can think of it sort of like `var getStudentName = function(studentID)`, but that's not exactly accurate. An identifier `getStudentName` is declared (at compile time), but the `= function(studentID)` part is also handled at compilation; the association between `getStudentName` and the function is automatically set up at the beginning of the scope rather than waiting for an `=` assignment statement to be executed.

NOTE:

This automatic association of function and variable is referred to as “function hoisting”, and is covered in detail in Chapter 5.

Sources

So we've identified all five *target* references in the program. The other variable references must then be *source* references (because that's the only other option!).

In `for (let student of students)`, we said that `student` is a *target*, but `students` is a *source* reference. In the statement `if (student.id == studentID)`, both `student` and `studentID` are *source* references. `student` is also a *source* reference in `return student.name`.

In `getStudentName(73)`, `getStudentName` is a *source* reference (which we hope resolves to a function reference value). In `console.log(nextStudent)`, `console` is a *source* reference, as is `nextStudent`.

NOTE:

In case you were wondering, `id`, `name`, and `log` are all properties, not variable references.

What's the practical importance of understanding *targets* vs. *sources*? In Chapter 2, we'll revisit this topic and cover how a variable's role impacts its lookup (specifically, if the lookup fails).

Cheating: Runtime Scope Modifications

It should be clear by now that scope is determined as the program is compiled, and should not generally be affected by runtime conditions. However, in non-strict-mode, there are technically still two ways to cheat this rule, modifying a program's scopes during runtime.

Neither of these techniques *should* be used—they're both dangerous and confusing, and you should be using strict-mode (where they're disallowed) anyway. But it's important to be aware of them in case you run across them in some programs.

The `eval(..)` function receives a string of code to compile and execute on the fly during the program runtime. If that string of code has a `var` or `function` declaration in it, those declarations will modify the current scope that the `eval(..)` is currently executing in:

```
function badIdea() {
  eval("var oops = 'Ugh!';");
  console.log(oops);
}
badIdea(); // Ugh!
```

If the `eval(..)` had not been present, the `oops` variable in `console.log(oops)` would not exist, and would throw a `ReferenceError`. But `eval(..)` modifies the scope of the `badIdea()` function at runtime. This is bad for many reasons, including the performance hit of modifying the already compiled and optimized scope, every time `badIdea()` runs.

The second cheat is the `with` keyword, which essentially dynamically turns an object into a local scope—its properties are treated as identifiers in that new scope's block:

```
var badIdea = { oops: "Ugh!" };

with (badIdea) {
  console.log(oops); // Ugh!
}
```

The global scope was not modified here, but `badIdea` was turned into a scope at runtime rather than compile time, and its property `oops` becomes a variable in that scope. Again, this is a terrible idea, for performance and readability reasons.

At all costs, avoid `eval(..)` (at least, `eval(..)` creating declarations) and `with`. Again, neither of these cheats is available in strict-mode, so if you just

use strict-mode (you should!) then the temptation goes away!

Lexical Scope

We've demonstrated that JS's scope is determined at compile time; the term for this kind of scope is "lexical scope". "Lexical" is associated with the "lexing" stage of compilation, as discussed earlier in this chapter.

To narrow this chapter down to a useful conclusion, the key idea of "lexical scope" is that it's controlled entirely by the placement of functions, blocks, and variable declarations, in relation to one another.

If you place a variable declaration inside a function, the compiler handles this declaration as it's parsing the function, and associates that declaration with the function's scope. If a variable is block-scope declared (`let` / `const`), then it's associated with the nearest enclosing `{ ... }` block, rather than its enclosing function (as with `var`).

Furthermore, a reference (*target* or *source* role) for a variable must be resolved as coming from one of the scopes that are *lexically available* to it; otherwise the variable is said to be "undeclared" (which usually results in an error!). If the variable is not declared in the current scope, the next outer/enclosing scope will be consulted. This process of stepping out one level of scope nesting continues until either a matching variable declaration can be found, or the global scope is reached and there's nowhere else to go.

It's important to note that compilation doesn't actually *do anything* in terms of reserving memory for scopes and variables. None of the program has been executed yet.

Instead, compilation creates a map of all the lexical scopes that lays out what the program will need while it executes. You can think of this plan as inserted code for use at runtime, which defines all the scopes (aka, "lexical environments") and registers all the identifiers (variables) for each scope.

In other words, while scopes are identified during compilation, they're not actually created until runtime, each time a scope needs to run. In the next chapter, we'll sketch out the conceptual foundations for lexical scope.

You Don't Know JS Yet: Scope & Closures - 2nd Edition

Chapter 3: The Scope Chain

Chapters 1 and 2 laid down a concrete definition of *lexical scope* (and its parts) and illustrated helpful metaphors for its conceptual foundation. Before proceeding with this chapter, find someone else to explain (written or aloud), in your own words, what lexical scope is and why it's useful to understand.

That seems like a step you might skip, but I've found it really does help to take the time to reformulate these ideas as explanations to others. That helps our brains digest what we're learning!

Now it's time to dig into the nuts and bolts, so expect that things will get a lot more detailed from here forward. Stick with it, though, because these discussions really hammer home just how much we all *don't know* about scope, yet. Make sure to take your time with the text and all the code snippets provided.

To refresh the context of our running example, let's recall the color-coded illustration of the nested scope bubbles, from Chapter 2, Figure 2:

Fig. 2 (Ch. 2): Colored Scope Bubbles

The connections between scopes that are nested within other scopes is called the scope chain, which determines the path along which variables can be accessed. The chain is directed, meaning the lookup moves upward/outward only.

“Lookup” Is (Mostly) Conceptual

In Figure 2, notice the color of the `students` variable reference in the `for`-loop. How exactly did we determine that it's a RED(1) marble?

In Chapter 2, we described the runtime access of a variable as a “lookup,” where the *Engine* has to start by asking the current scope's *Scope Manager* if it knows about an identifier/variable, and proceeding upward/outward back through the chain of nested scopes (toward the global scope) until found, if ever. The lookup stops as soon as the first matching named declaration in a scope bucket is found.

The lookup process thus determined that `students` is a RED(1) marble, because we had not yet found a matching variable name as we traversed the scope chain, until we arrived at the final RED(1) global scope.

Similarly, `studentID` in the `if`-statement is determined to be a BLUE(2) marble.

This suggestion of a runtime lookup process works well for conceptual understanding, but it's not actually how things usually work in practice.

The color of a marble's bucket (aka, meta information of what scope a variable originates from) is *usually determined* during the initial compilation processing. Because lexical scope is pretty much finalized at that point, a marble's color will not change based on anything that can happen later during runtime.

Since the marble's color is known from compilation, and it's immutable, this information would likely be stored with (or at least accessible from) each variable's entry in the AST; that information is then used explicitly by the executable instructions that constitute the program's runtime.

In other words, *Engine* (from Chapter 2) doesn't need to lookup through a bunch of scopes to figure out which scope bucket a variable comes from. That information is already known! Avoiding the need for a runtime lookup is a key optimization benefit of lexical scope. The runtime operates more performantly without spending time on all these lookups.

But I said "...usually determined..." just a moment ago, with respect to figuring out a marble's color during compilation. So in what case would it ever *not* be known during compilation?

Consider a reference to a variable that isn't declared in any lexically available scopes in the current file—see *Get Started*, Chapter 1, which asserts that each file is its own separate program from the perspective of JS compilation. If no declaration is found, that's not *necessarily* an error. Another file (program) in the runtime may indeed declare that variable in the shared global scope.

So the ultimate determination of whether the variable was ever appropriately declared in some accessible bucket may need to be deferred to the runtime.

Any reference to a variable that's initially *undeclared* is left as an uncolored marble during that file's compilation; this color cannot be determined until other relevant file(s) have been compiled and the application runtime commences. That deferred lookup will eventually resolve the color to whichever scope the variable is found in (likely the global scope).

However, this lookup would only be needed once per variable at most, since nothing else during runtime could later change that marble's color.

The "Lookup Failures" section in Chapter 2 covers what happens if a marble is ultimately still uncolored at the moment its reference is runtime executed.

Shadowing

"Shadowing" might sound mysterious and a little bit sketchy. But don't worry, it's completely legit!

Our running example for these chapters uses different variable names across the scope boundaries. Since they all have unique names, in a way it wouldn't matter if all of them were just stored in one bucket (like RED(1)).

Where having different lexical scope buckets starts to matter more is when you have two or more variables, each in different scopes, with the same lexical names. A single scope cannot have two or more variables with the same name; such multiple references would be assumed as just one variable.

So if you need to maintain two or more variables of the same name, you must use separate (often nested) scopes. And in that case, it's very relevant how the different scope buckets are laid out.

Consider:

```
var studentName = "Suzy";  
  
function printStudent(studentName) {  
    studentName = studentName.toUpperCase();  
    console.log(studentName);  
}  
  
printStudent("Frank");  
// FRANK  
  
printStudent(studentName);  
// SUZY  
  
console.log(studentName);  
// Suzy
```

TIP:

Before you move on, take some time to analyze this code using the various techniques/metaphors we've covered in the book. In particular, make sure to identify the marble/bubble colors in this snippet. It's good practice!

The `studentName` variable on line 1 (the `var studentName = ...` statement) creates a RED(1) marble. The same named variable is declared as a BLUE(2) marble on line 3, the parameter in the `printStudent(..)` function definition.

What color marble will `studentName` be in the `studentName = studentName.toUpperCase()` assignment statement and the `console.log(studentName)` statement? All three `studentName` references will be BLUE(2).

With the conceptual notion of the “lookup,” we asserted that it starts with the current scope and works its way outward/upward, stopping as soon as a matching variable is found. The BLUE(2) `studentName` is found right away. The RED(1) `studentName` is never even considered.

This is a key aspect of lexical scope behavior, called *shadowing*. The BLUE(2) `studentName` variable (parameter) shadows the RED(1) `studentName`. So, the

parameter is shadowing the (shadowed) global variable. Repeat that sentence to yourself a few times to make sure you have the terminology straight!

That's why the re-assignment of `studentName` affects only the inner (parameter) variable: the BLUE(2) `studentName`, not the global RED(1) `studentName`.

When you choose to shadow a variable from an outer scope, one direct impact is that from that scope inward/downward (through any nested scopes) it's now impossible for any marble to be colored as the shadowed variable—(RED(1), in this case). In other words, any `studentName` identifier reference will correspond to that parameter variable, never the global `studentName` variable. It's lexically impossible to reference the global `studentName` anywhere inside of the `printStudent(...)` function (or from any nested scopes).

Global Unshadowing Trick

Please beware: leveraging the technique I'm about to describe is not very good practice, as it's limited in utility, confusing for readers of your code, and likely to invite bugs to your program. I'm covering it only because you may run across this behavior in existing programs, and understanding what's happening is critical to not getting tripped up.

It *is* possible to access a global variable from a scope where that variable has been shadowed, but not through a typical lexical identifier reference.

In the global scope (RED(1)), `var` declarations and `function` declarations also expose themselves as properties (of the same name as the identifier) on the *global object*—essentially an object representation of the global scope. If you've written JS for a browser environment, you probably recognize the global object as `window`. That's not *entirely* accurate, but it's good enough for our discussion. In the next chapter, we'll explore the global scope/object topic more.

Consider this program, specifically executed as a standalone .js file in a browser environment:

```
var studentName = "Suzy";

function printStudent(studentName) {
  console.log(studentName);
  console.log(window.studentName);
}

printStudent("Frank");
// "Frank"
// "Suzy"
```

Notice the `window.studentName` reference? This expression is accessing the global variable `studentName` as a property on `window` (which we're pretending

for now is synonymous with the global object). That's the only way to access a shadowed variable from inside a scope where the shadowing variable is present.

The `window.studentName` is a mirror of the global `studentName` variable, not a separate snapshot copy. Changes to one are still seen from the other, in either direction. You can think of `window.studentName` as a getter/setter that accesses the actual `studentName` variable. As a matter of fact, you can even *add* a variable to the global scope by creating/setting a property on the global object.

WARNING:

Remember: just because you *can* doesn't mean you *should*. Don't shadow a global variable that you need to access, and conversely, avoid using this trick to access a global variable that you've shadowed. And definitely don't confuse readers of your code by creating global variables as `window` properties instead of with formal declarations!

This little "trick" only works for accessing a global scope variable (not a shadowed variable from a nested scope), and even then, only one that was declared with `var` or `function`.

Other forms of global scope declarations do not create mirrored global object properties:

```
var one = 1;
let notOne = 2;
const notTwo = 3;
class notThree {}

console.log(window.one);      // 1
console.log(window.notOne);   // undefined
console.log(window.notTwo);   // undefined
console.log(window.notThree); // undefined
```

Variables (no matter how they're declared!) that exist in any other scope than the global scope are completely inaccessible from a scope where they've been shadowed:

```
var special = 42;

function lookingFor(special) {
  // The identifier `special` (parameter) in this
  // scope is shadowed inside keepLooking(), and
  // is thus inaccessible from that scope.

  function keepLooking() {
    var special = 3.141592;
```

```

        console.log(special);
        console.log(window.special);
    }

    keepLooking();
}

lookingFor(112358132134);
// 3.141592
// 42

```

The global RED(1) `special` is shadowed by the BLUE(2) `special` (parameter), and the BLUE(2) `special` is itself shadowed by the GREEN(3) `special` inside `keepLooking()`. We can still access the RED(1) `special` using the indirect reference `window.special`. But there's no way for `keepLooking()` to access the BLUE(2) `special` that holds the number 112358132134.

Copying Is Not Accessing

I've been asked the following "But what about...?" question dozens of times. Consider:

```

var special = 42;

function lookingFor(special) {
    var another = {
        special: special
    };

    function keepLooking() {
        var special = 3.141592;
        console.log(special);
        console.log(another.special); // Ooo, tricky!
        console.log(window.special);
    }

    keepLooking();
}

lookingFor(112358132134);
// 3.141592
// 112358132134
// 42

```

Oh! So does this `another` object technique disprove my claim that the `special` parameter is "completely inaccessible" from inside `keepLooking()`? No, the claim is still correct.

special: `special` is copying the value of the `special` parameter variable into another container (a property of the same name). Of course, if you put a value in another container, shadowing no longer applies (unless `another` was shadowed, too!). But that doesn't mean we're accessing the parameter `special`; it means we're accessing the copy of the value it had at that moment, by way of *another* container (object property). We cannot reassign the `BLUE(2)` `special` parameter to a different value from inside `keepLooking()`.

Another “But...!” you may be about to raise: what if I'd used objects or arrays as the values instead of the numbers (112358132134, etc.)? Would us having references to objects instead of copies of primitive values “fix” the inaccessibility?

No. Mutating the contents of the object value via a reference copy is **not** the same thing as lexically accessing the variable itself. We still can't reassign the `BLUE(2)` `special` parameter.

Illegal Shadowing

Not all combinations of declaration shadowing are allowed. `let` can shadow `var`, but `var` cannot shadow `let`:

```
function something() {
  var special = "JavaScript";

  {
    let special = 42;    // totally fine shadowing

    //
  }
}

function another() {
  //
  {

    let special = "JavaScript";

    {
      var special = "JavaScript";
      // ^^^ Syntax Error

      //
    }
  }
}
```

Notice in the `another()` function, the inner `var special` declaration is attempting to declare a function-wide `special`, which in and of itself is fine (as shown

by the `something()` function).

The syntax error description in this case indicates that `special` has already been defined, but that error message is a little misleading—again, no such error happens in `something()`, as shadowing is generally allowed just fine.

The real reason it's raised as a `SyntaxError` is because the `var` is basically trying to “cross the boundary” of (or hop over) the `let` declaration of the same name, which is not allowed.

That boundary-crossing prohibition effectively stops at each function boundary, so this variant raises no exception:

```
function another() {
  // ...

  {
    let special = "JavaScript";

    ajax("https://some.url",function callback(){
      // totally fine shadowing
      var special = "JavaScript";

      // ...
    });
  }
}
```

Summary: `let` (in an inner scope) can always shadow an outer scope's `var`. `var` (in an inner scope) can only shadow an outer scope's `let` if there is a function boundary in between.

Function Name Scope

As you've seen by now, a `function` declaration looks like this:

```
function askQuestion() {
  // ...
}
```

And as discussed in Chapters 1 and 2, such a `function` declaration will create an identifier in the enclosing scope (in this case, the global scope) named `askQuestion`.

What about this program?

```
var askQuestion = function(){
  // ...
};
```

The same is true for the variable `askQuestion` being created. But since it's a `function` expression—a function definition used as value instead of a standalone declaration—the function itself will not “hoist” (see Chapter 5).

One major difference between `function` declarations and `function` expressions is what happens to the name identifier of the function. Consider a named `function` expression:

```
var askQuestion = function ofTheTeacher(){  
  // ..  
};
```

We know `askQuestion` ends up in the outer scope. But what about the `ofTheTeacher` identifier? For formal `function` declarations, the name identifier ends up in the outer/enclosing scope, so it may be reasonable to assume that's the case here. But `ofTheTeacher` is declared as an identifier **inside the function itself**:

```
var askQuestion = function ofTheTeacher() {  
  console.log(ofTheTeacher);  
};  
  
askQuestion();  
// function ofTheTeacher()...  
  
console.log(ofTheTeacher);  
// ReferenceError: ofTheTeacher is not defined
```

NOTE:

Actually, `ofTheTeacher` is not exactly *in the scope of the function*. Appendix A, “Implied Scopes” will explain further.

Not only is `ofTheTeacher` declared inside the function rather than outside, but it's also defined as read-only:

```
var askQuestion = function ofTheTeacher() {  
  "use strict";  
  ofTheTeacher = 42;    // TypeError  
  
  //..  
};  
  
askQuestion();  
// TypeError
```

Because we used strict-mode, the assignment failure is reported as a `TypeError`; in non-strict-mode, such an assignment fails silently with no exception.

What about when a `function` expression has no name identifier?

```
var askQuestion = function(){
  // ..
};
```

A `function` expression with a name identifier is referred to as a “named function expression,” but one without a name identifier is referred to as an “anonymous function expression.” Anonymous function expressions clearly have no name identifier that affects either scope.

NOTE:

We’ll discuss named vs. anonymous `function` expressions in much more detail, including what factors affect the decision to use one or the other, in Appendix A.

Arrow Functions

ES6 added an additional `function` expression form to the language, called “arrow functions”:

```
var askQuestion = () => {
  // ..
};
```

The `=>` arrow function doesn’t require the word `function` to define it. Also, the `(...)` around the parameter list is optional in some simple cases. Likewise, the `{ ... }` around the function body is optional in some cases. And when the `{ ... }` are omitted, a return value is sent out without using a `return` keyword.

NOTE:

The attractiveness of `=>` arrow functions is often sold as “shorter syntax,” and that’s claimed to equate to objectively more readable code. This claim is dubious at best, and I believe outright misguided. We’ll dig into the “readability” of various function forms in Appendix A.

Arrow functions are lexically anonymous, meaning they have no directly related identifier that references the function. The assignment to `askQuestion` creates an inferred name of “`askQuestion`”, but that’s **not the same thing as being non-anonymous**:

```
var askQuestion = () => {
  // ..
};

askQuestion.name; // askQuestion
```

Arrow functions achieve their syntactic brevity at the expense of having to mentally juggle a bunch of variations for different forms/conditions. Just a few, for example:

```
() => 42;

id => id.toUpperCase();

(id, name) => ({ id, name });

(...args) => {
  return args[args.length - 1];
};
```

The real reason I bring up arrow functions is because of the common but incorrect claim that arrow functions somehow behave differently with respect to lexical scope from standard `function` functions.

This is incorrect.

Other than being anonymous (and having no declarative form), `=>` arrow functions have the same lexical scope rules as `function` functions do. An arrow function, with or without `{ .. }` around its body, still creates a separate, inner nested bucket of scope. Variable declarations inside this nested scope bucket behave the same as in a `function` scope.

Backing Out

When a function (declaration or expression) is defined, a new scope is created. The positioning of scopes nested inside one another creates a natural scope hierarchy throughout the program, called the scope chain. The scope chain controls variable access, directionally oriented upward and outward.

Each new scope offers a clean slate, a space to hold its own set of variables. When a variable name is repeated at different levels of the scope chain, shadowing occurs, which prevents access to the outer variable from that point inward.

As we step back out from these finer details, the next chapter shifts focus to the primary scope all JS programs include: the global scope.

You Don't Know JS Yet: Scope & Closures - 2nd Edition

Chapter 2: Illustrating Lexical Scope

In Chapter 1, we explored how scope is determined during code compilation, a model called “lexical scope.” The term “lexical” refers to the first stage of compilation (lexing/parsing).

To properly *reason* about our programs, it’s important to have a solid conceptual foundation of how scope works. If we rely on guesses and intuition, we may accidentally get the right answers some of the time, but many other times we’re far off. This isn’t a recipe for success.

Like way back in grade school math class, getting the right answer isn’t enough if we don’t show the correct steps to get there! We need to build accurate and helpful mental models as foundation moving forward.

This chapter will illustrate *scope* with several metaphors. The goal here is to *think* about how your program is handled by the JS engine in ways that more closely align with how the JS engine actually works.

Marbles, and Buckets, and Bubbles... Oh My!

One metaphor I’ve found effective in understanding scope is sorting colored marbles into buckets of their matching color.

Imagine you come across a pile of marbles, and notice that all the marbles are colored red, blue, or green. Let’s sort all the marbles, dropping the red ones into a red bucket, green into a green bucket, and blue into a blue bucket. After sorting, when you later need a green marble, you already know the green bucket is where to go to get it.

In this metaphor, the marbles are the variables in our program. The buckets are scopes (functions and blocks), which we just conceptually assign individual colors for our discussion purposes. The color of each marble is thus determined by which *color* scope we find the marble originally created in.

Let’s annotate the running program example from Chapter 1 with scope color labels:

```
// outer/global scope: RED

var students = [
  { id: 14, name: "Kyle" },
  { id: 73, name: "Suzy" },
  { id: 112, name: "Frank" },
  { id: 6, name: "Sarah" }
];
```

```

function getStudentName(studentID) {
  // function scope: BLUE

  for (let student of students) {
    // loop scope: GREEN

    if (student.id == studentID) {
      return student.name;
    }
  }
}

var nextStudent = getStudentName(73);
console.log(nextStudent); // Suzy

```

We've designated three scope colors with code comments: RED (outermost global scope), BLUE (scope of function `getStudentName(..)`), and GREEN (scope of/inside the `for` loop). But it still may be difficult to recognize the boundaries of these scope buckets when looking at a code listing.

Figure 2 helps visualize the boundaries of the scopes by drawing colored bubbles (aka, buckets) around each:

Fig. 2: Colored Scope Bubbles

1. **Bubble 1 (RED)** encompasses the global scope, which holds three identifiers/variables: `students` (line 1), `getStudentName` (line 8), and `nextStudent` (line 16).
2. **Bubble 2 (BLUE)** encompasses the scope of the function `getStudentName(..)` (line 8), which holds just one identifier/variable: the parameter `studentID` (line 8).
3. **Bubble 3 (GREEN)** encompasses the scope of the `for`-loop (line 9), which holds just one identifier/variable: `student` (line 9).

NOTE:

Technically, the parameter `studentID` is not exactly in the BLUE(2) scope. We'll unwind that confusion in "Implied Scopes" in Appendix A. For now, it's close enough to label `studentID` a BLUE(2) marble.

Scope bubbles are determined during compilation based on where the functions/blocks of scope are written, the nesting inside each other, and so on. Each scope bubble is entirely contained within its parent scope bubble—a scope is never partially in two different outer scopes.

Each marble (variable/identifier) is colored based on which bubble (bucket) it's declared in, not the color of the scope it may be accessed from (e.g., `students` on line 9 and `studentID` on line 10).

NOTE:

Remember we asserted in Chapter 1 that `id`, `name`, and `log` are all properties, not variables; in other words, they're not marbles in buckets, so they don't get colored based on any the rules we're discussing in this book. To understand how such property accesses are handled, see the third book in the series, *Objects & Classes*.

As the JS engine processes a program (during compilation), and finds a declaration for a variable, it essentially asks, "Which *color* scope (bubble or bucket) am I currently in?" The variable is designated as that same *color*, meaning it belongs to that bucket/bubble.

The GREEN(3) bucket is wholly nested inside of the BLUE(2) bucket, and similarly the BLUE(2) bucket is wholly nested inside the RED(1) bucket. Scopes can nest inside each other as shown, to any depth of nesting as your program needs.

References (non-declarations) to variables/identifiers are allowed if there's a matching declaration either in the current scope, or any scope above/outside the current scope, but not with declarations from lower/nested scopes.

An expression in the RED(1) bucket only has access to RED(1) marbles, **not** BLUE(2) or GREEN(3). An expression in the BLUE(2) bucket can reference either BLUE(2) or RED(1) marbles, **not** GREEN(3). And an expression in the GREEN(3) bucket has access to RED(1), BLUE(2), and GREEN(3) marbles.

We can conceptualize the process of determining these non-declaration marble colors during runtime as a lookup. Since the `students` variable reference in the `for`-loop statement on line 9 is not a declaration, it has no color. So we ask the current BLUE(2) scope bucket if it has a marble matching that name. Since it doesn't, the lookup continues with the next outer/containing scope: RED(1). The RED(1) bucket has a marble of the name `students`, so the loop-statement's `students` variable reference is determined to be a RED(1) marble.

The `if (student.id == studentID)` statement on line 10 is similarly determined to reference a GREEN(3) marble named `student` and a BLUE(2) marble `studentID`.

NOTE:

The JS engine doesn't generally determine these marble colors during runtime; the "lookup" here is a rhetorical device to help you understand the concepts. During compilation, most or all variable references will match already-known scope buckets, so their color is already determined, and stored with each marble reference to avoid unnecessary lookups as the program runs. More on this nuance in Chapter 3.

The key take-aways from marbles & buckets (and bubbles!):

- Variables are declared in specific scopes, which can be thought of as colored marbles from matching-color buckets.
- Any variable reference that appears in the scope where it was declared, or appears in any deeper nested scopes, will be labeled a marble of that same color—unless an intervening scope “shadows” the variable declaration; see “Shadowing” in Chapter 3.
- The determination of colored buckets, and the marbles they contain, happens during compilation. This information is used for variable (marble color) “lookups” during code execution.

A Conversation Among Friends

Another useful metaphor for the process of analyzing variables and the scopes they come from is to imagine various conversations that occur inside the engine as code is processed and then executed. We can “listen in” on these conversations to get a better conceptual foundation for how scopes work.

Let's now meet the members of the JS engine that will have conversations as they process our program:

- *Engine*: responsible for start-to-finish compilation and execution of our JavaScript program.
- *Compiler*: one of *Engine*'s friends; handles all the dirty work of parsing and code-generation (see previous section).
- *Scope Manager*: another friend of *Engine*; collects and maintains a lookup list of all the declared variables/identifiers, and enforces a set of rules as to how these are accessible to currently executing code.

For you to *fully understand* how JavaScript works, you need to begin to *think* like *Engine* (and friends) think, ask the questions they ask, and answer their questions likewise.

To explore these conversations, recall again our running program example:

```

var students = [
  { id: 14, name: "Kyle" },
  { id: 73, name: "Suzy" },
  { id: 112, name: "Frank" },
  { id: 6, name: "Sarah" }
];

function getStudentName(studentID) {
  for (let student of students) {
    if (student.id == studentID) {
      return student.name;
    }
  }
}

var nextStudent = getStudentName(73);

console.log(nextStudent);
// Suzy

```

Let's examine how JS is going to process that program, specifically starting with the first statement. The array and its contents are just basic JS value literals (and thus unaffected by any scoping concerns), so our focus here will be on the `var students = [...]` declaration and initialization-assignment parts.

We typically think of that as a single statement, but that's not how our friend *Engine* sees it. In fact, JS treats these as two distinct operations, one which *Compiler* will handle during compilation, and the other which *Engine* will handle during execution.

The first thing *Compiler* will do with this program is perform lexing to break it down into tokens, which it will then parse into a tree (AST).

Once *Compiler* gets to code generation, there's more detail to consider than may be obvious. A reasonable assumption would be that *Compiler* will produce code for the first statement such as: “Allocate memory for a variable, label it `students`, then stick a reference to the array into that variable.” But that's not the whole story.

Here's the steps *Compiler* will follow to handle that statement:

1. Encountering `var students`, *Compiler* will ask *Scope Manager* to see if a variable named `students` already exists for that particular scope bucket. If so, *Compiler* would ignore this declaration and move on. Otherwise, *Compiler* will produce code that (at execution time) asks *Scope Manager* to create a new variable called `students` in that scope bucket.
2. *Compiler* then produces code for *Engine* to later execute, to handle the `students = []` assignment. The code *Engine* runs will first ask *Scope*

Manager if there is a variable called `students` accessible in the current scope bucket. If not, *Engine* keeps looking elsewhere (see “Nested Scope” below). Once *Engine* finds a variable, it assigns the reference of the `[..]` array to it.

In conversational form, the first phase of compilation for the program might play out between *Compiler* and *Scope Manager* like this:

Compiler: Hey, *Scope Manager* (of the global scope), I found a formal declaration for an identifier called `students`, ever heard of it?

(Global) Scope Manager: Nope, never heard of it, so I just created it for you.

Compiler: Hey, *Scope Manager*, I found a formal declaration for an identifier called `getStudentName`, ever heard of it?

(Global) Scope Manager: Nope, but I just created it for you.

Compiler: Hey, *Scope Manager*, `getStudentName` points to a function, so we need a new scope bucket.

(Function) Scope Manager: Got it, here’s the scope bucket.

Compiler: Hey, *Scope Manager* (of the function), I found a formal parameter declaration for `studentID`, ever heard of it?

(Function) Scope Manager: Nope, but now it’s created in this scope.

Compiler: Hey, *Scope Manager* (of the function), I found a `for`-loop that will need its own scope bucket.

...

The conversation is a question-and-answer exchange, where **Compiler** asks the current *Scope Manager* if an encountered identifier declaration has already been encountered. If “no,” *Scope Manager* creates that variable in that scope. If the answer is “yes,” then it’s effectively skipped over since there’s nothing more for that *Scope Manager* to do.

Compiler also signals when it runs across functions or block scopes, so that a new scope bucket and *Scope Manager* can be instantiated.

Later, when it comes to execution of the program, the conversation will shift to *Engine* and *Scope Manager*, and might play out like this:

Engine: Hey, *Scope Manager* (of the global scope), before we begin, can you look up the identifier `getStudentName` so I can assign this function to it?

(Global) Scope Manager: Yep, here’s the variable.

Engine: Hey, *Scope Manager*, I found a *target* reference for `students`, ever heard of it?

(Global) Scope Manager: Yes, it was formally declared for this scope, so here it is.

Engine: Thanks, I'm initializing `students` to `undefined`, so it's ready to use.

Hey, *Scope Manager* (of the global scope), I found a *target* reference for `nextStudent`, ever heard of it?

(Global) Scope Manager: Yes, it was formally declared for this scope, so here it is.

Engine: Thanks, I'm initializing `nextStudent` to `undefined`, so it's ready to use.

Hey, *Scope Manager* (of the global scope), I found a *source* reference for `getStudentName`, ever heard of it?

(Global) Scope Manager: Yes, it was formally declared for this scope. Here it is.

Engine: Great, the value in `getStudentName` is a function, so I'm going to execute it.

Engine: Hey, *Scope Manager*, now we need to instantiate the function's scope.

...

This conversation is another question-and-answer exchange, where *Engine* first asks the current *Scope Manager* to look up the hoisted `getStudentName` identifier, so as to associate the function with it. *Engine* then proceeds to ask *Scope Manager* about the *target* reference for `students`, and so on.

To review and summarize how a statement like `var students = [...]` is processed, in two distinct steps:

1. *Compiler* sets up the declaration of the scope variable (since it wasn't previously declared in the current scope).
2. While *Engine* is executing, to process the assignment part of the statement, *Engine* asks *Scope Manager* to look up the variable, initializes it to `undefined` so it's ready to use, and then assigns the array value to it.

Nested Scope

When it comes time to execute the `getStudentName()` function, *Engine* asks for a *Scope Manager* instance for that function's scope, and it will then proceed to look up the parameter (`studentID`) to assign the `73` argument value to, and so on.

The function scope for `getStudentName(..)` is nested inside the global scope. The block scope of the `for`-loop is similarly nested inside that function scope. Scopes can be lexically nested to any arbitrary depth as the program defines.

Each scope gets its own *Scope Manager* instance each time that scope is executed (one or more times). Each scope automatically has all its identifiers registered at the start of the scope being executed (this is called “variable hoisting”; see Chapter 5).

At the beginning of a scope, if any identifier came from a `function` declaration, that variable is automatically initialized to its associated function reference. And if any identifier came from a `var` declaration (as opposed to `let/const`), that variable is automatically initialized to `undefined` so that it can be used; otherwise, the variable remains uninitialized (aka, in its “TDZ,” see Chapter 5) and cannot be used until its full declaration-and-initialization are executed.

In the `for (let student of students) {` statement, `students` is a *source* reference that must be looked up. But how will that lookup be handled, since the scope of the function will not find such an identifier?

To explain, let’s imagine that bit of conversation playing out like this:

Engine: Hey, *Scope Manager* (for the function), I have a *source* reference for `students`, ever heard of it?

(Function) Scope Manager: Nope, never heard of it. Try the next outer scope.

Engine: Hey, *Scope Manager* (for the global scope), I have a *source* reference for `students`, ever heard of it?

(Global) Scope Manager: Yep, it was formally declared, here it is.

...

One of the key aspects of lexical scope is that any time an identifier reference cannot be found in the current scope, the next outer scope in the nesting is consulted; that process is repeated until an answer is found or there are no more scopes to consult.

Lookup Failures

When *Engine* exhausts all *lexically available* scopes (moving outward) and still cannot resolve the lookup of an identifier, an error condition then exists. However, depending on the mode of the program (strict-mode or not) and the role of the variable (i.e., *target* vs. *source*; see Chapter 1), this error condition will be handled differently.

Undefined Mess If the variable is a *source*, an unresolved identifier lookup is considered an undeclared (unknown, missing) variable, which always results in

a `ReferenceError` being thrown. Also, if the variable is a *target*, and the code at that moment is running in strict-mode, the variable is considered undeclared and similarly throws a `ReferenceError`.

The error message for an undeclared variable condition, in most JS environments, will look like, “Reference Error: XYZ is not defined.” The phrase “not defined” seems almost identical to the word “undefined,” as far as the English language goes. But these two are very different in JS, and this error message unfortunately creates a persistent confusion.

“Not defined” really means “not declared”—or, rather, “undeclared,” as in a variable that has no matching formal declaration in any *lexically available* scope. By contrast, “undefined” really means a variable was found (declared), but the variable otherwise has no other value in it at the moment, so it defaults to the `undefined` value.

To perpetuate the confusion even further, JS’s `typeof` operator returns the string `"undefined"` for variable references in either state:

```
var studentName;
typeof studentName;    // "undefined"

typeof doesntExist;  // "undefined"
```

These two variable references are in very different conditions, but JS sure does muddy the waters. The terminology mess is confusing and terribly unfortunate. Unfortunately, JS developers just have to pay close attention to not mix up *which kind* of “undefined” they’re dealing with!

Global... What!? If the variable is a *target* and strict-mode is not in effect, a confusing and surprising legacy behavior kicks in. The troublesome outcome is that the global scope’s *Scope Manager* will just create an **accidental global variable** to fulfill that target assignment!

Consider:

```
function getStudentName() {
  // assignment to an undeclared variable :(
  nextStudent = "Suzy";
}

getStudentName();

console.log(nextStudent);
// "Suzy" -- oops, an accidental-global variable!
```

Here’s how that *conversation* will proceed:

Engine: Hey, *Scope Manager* (for the function), I have a *target* reference for `nextStudent`, ever heard of it?

(Function) Scope Manager: Nope, never heard of it. Try the next outer scope.

Engine: Hey, *Scope Manager* (for the global scope), I have a *target* reference for `nextStudent`, ever heard of it?

(Global) Scope Manager: Nope, but since we're in non-strict-mode, I helped you out and just created a global variable for you, here it is!

Yuck.

This sort of accident (almost certain to lead to bugs eventually) is a great example of the beneficial protections offered by strict-mode, and why it's such a bad idea *not* to be using strict-mode. In strict-mode, the **Global Scope Manager** would instead have responded:

(Global) Scope Manager: Nope, never heard of it. Sorry, I've got to throw a `ReferenceError`.

Assigning to a never-declared variable *is* an error, so it's right that we would receive a `ReferenceError` here.

Never rely on accidental global variables. Always use strict-mode, and always formally declare your variables. You'll then get a helpful `ReferenceError` if you ever mistakenly try to assign to a not-declared variable.

Building On Metaphors

To visualize nested scope resolution, I prefer yet another metaphor, an office building, as in Figure 3:

Fig. 3: Scope “Building”

The building represents our program's nested scope collection. The first floor of the building represents the currently executing scope. The top level of the building is the global scope.

You resolve a *target* or *source* variable reference by first looking on the current floor, and if you don't find it, taking the elevator to the next floor (i.e., an outer scope), looking there, then the next, and so on. Once you get to the top floor (the global scope), you either find what you're looking for, or you don't. But you have to stop regardless.

Continue the Conversation

By this point, you should be developing richer mental models for what scope is and how the JS engine determines and uses it from your code.

Before *continuing*, go find some code in one of your projects and run through these conversations. Seriously, actually speak out loud. Find a friend and

practice each role with them. If either of you find yourself confused or tripped up, spend more time reviewing this material.

As we move (up) to the next (outer) chapter, we'll explore how the lexical scopes of a program are connected in a chain.

You Don't Know JS Yet: Scope & Closures - 2nd Edition

Chapter 4: Around the Global Scope

Chapter 3 mentioned the “global scope” several times, but you may still be wondering why a program’s outermost scope is all that important in modern JS. The vast majority of work is now done inside of functions and modules rather than globally.

Is it good enough to just assert, “Avoid using the global scope,” and be done with it?

The global scope of a JS program is a rich topic, with much more utility and nuance than you would likely assume. This chapter first explores how the global scope is (still) useful and relevant to writing JS programs today, then looks at differences in where and *how to access* the global scope in different JS environments.

Fully understanding the global scope is critical in your mastery of using lexical scope to structure your programs.

Why Global Scope?

It’s likely no surprise to readers that most applications are composed of multiple (sometimes many!) individual JS files. So how exactly do all those separate files get stitched together in a single runtime context by the JS engine?

With respect to browser-executed applications, there are three main ways.

First, if you’re directly using ES modules (not transpiling them into some other module-bundle format), these files are loaded individually by the JS environment. Each module then `imports` references to whichever other modules it needs to access. The separate module files cooperate with each other exclusively through these shared imports, without needing any shared outer scope.

Second, if you’re using a bundler in your build process, all the files are typically concatenated together before delivery to the browser and JS engine, which then only processes one big file. Even with all the pieces of the application co-located in a single file, some mechanism is necessary for each piece to register a *name* to be referred to by other pieces, as well as some facility for that access to occur.

In some build setups, the entire contents of the file are wrapped in a single enclosing scope, such as a wrapper function, universal module (UMD—see Appendix A), etc. Each piece can register itself for access from other pieces by way of local variables in that shared scope. For example:

```
(function wrappingOuterScope(){
  var moduleOne = (function one(){
```

```

    // ..
})();

var moduleTwo = (function two(){
    // ..

    function callModuleOne() {
        moduleOne.someMethod();
    }

    // ..
})();
})();

```

As shown, the `moduleOne` and `moduleTwo` local variables inside the `wrappingOuterScope()` function scope are declared so that these modules can access each other for their cooperation.

While the scope of `wrappingOuterScope()` is a function and not the full environment global scope, it does act as a sort of “application-wide scope,” a bucket where all the top-level identifiers can be stored, though not in the real global scope. It’s kind of like a stand-in for the global scope in that respect.

And finally, the third way: whether a bundler tool is used for an application, or whether the (non-ES module) files are simply loaded in the browser individually (via `<script>` tags or other dynamic JS resource loading), if there is no single surrounding scope encompassing all these pieces, the **global scope** is the only way for them to cooperate with each other:

A bundled file of this sort often looks something like this:

```

var moduleOne = (function one(){
    // ..
})();
var moduleTwo = (function two(){
    // ..

    function callModuleOne() {
        moduleOne.someMethod();
    }

    // ..
})();

```

Here, since there is no surrounding function scope, these `moduleOne` and `moduleTwo` declarations are simply dropped into the global scope. This is effectively the same as if the files hadn’t been concatenated, but loaded separately:

module1.js:

```
var moduleOne = (function one(){
  // ...
})();
```

module2.js:

```
var moduleTwo = (function two(){
  // ...

  function callModuleOne() {
    moduleOne.someMethod();
  }

  // ...
})();
```

If these files are loaded separately as normal standalone .js files in a browser environment, each top-level variable declaration will end up as a global variable, since the global scope is the only shared resource between these two separate files—they’re independent programs, from the perspective of the JS engine.

In addition to (potentially) accounting for where an application’s code resides during runtime, and how each piece is able to access the other pieces to cooperate, the global scope is also where:

- JS exposes its built-ins:
 - primitives: `undefined`, `null`, `Infinity`, `NaN`
 - natives: `Date()`, `Object()`, `String()`, etc.
 - global functions: `eval()`, `parseInt()`, etc.
 - namespaces: `Math`, `Atomics`, `JSON`
 - friends of JS: `Intl`, `WebAssembly`
- The environment hosting the JS engine exposes its own built-ins:
 - `console` (and its methods)
 - the DOM (`window`, `document`, etc)
 - timers (`setTimeout(..)`, etc)
 - web platform APIs: `navigator`, `history`, `geolocation`, `WebRTC`, etc.

These are just some of the many *globals* your programs will interact with.

NOTE:

Node also exposes several elements “globally,” but they’re technically not in the *global* scope: `require()`, `__dirname`, `module`, `URL`, and so on.

Most developers agree that the global scope shouldn’t just be a dumping ground for every variable in your application. That’s a mess of bugs just waiting to

happen. But it's also undeniable that the global scope is an important *glue* for practically every JS application.

Where Exactly is this Global Scope?

It might seem obvious that the global scope is located in the outermost portion of a file; that is, not inside any function or other block. But it's not quite as simple as that.

Different JS environments handle the scopes of your programs, especially the global scope, differently. It's quite common for JS developers to harbor misconceptions without even realizing it.

Browser “Window”

With respect to treatment of the global scope, the most *pure* environment JS can be run in is as a standalone .js file loaded in a web page environment in a browser. I don't mean “pure” as in nothing automatically added—lots may be added!—but rather in terms of minimal intrusion on the code or interference with its expected global scope behavior.

Consider this .js file:

```
var studentName = "Kyle";

function hello() {
    console.log(`Hello, ${studentName}!`);
}

hello();
// Hello, Kyle!
```

This code may be loaded in a web page environment using an inline <script> tag, a <script src=..> script tag in the markup, or even a dynamically created <script> DOM element. In all three cases, the `studentName` and `hello` identifiers are declared in the global scope.

That means if you access the global object (commonly, `window` in the browser), you'll find properties of those same names there:

```
var studentName = "Kyle";

function hello() {
    console.log(`Hello, ${window.studentName}!`);
}

window.hello();
// Hello, Kyle!
```

That's the default behavior one would expect from a reading of the JS specification: the outer scope *is* the global scope and `studentName` is legitimately created as global variable.

That's what I mean by *pure*. But unfortunately, that won't always be true of all JS environments you encounter, and that's often surprising to JS developers.

Globals Shadowing Globals Recall the discussion of shadowing (and global unshadowing) from Chapter 3, where one variable declaration can override and prevent access to a declaration of the same name from an outer scope.

An unusual consequence of the difference between a global variable and a global property of the same name is that, within just the global scope itself, a global object can be shadowed by a global variable:

```
window.something = 42;

let something = "Kyle";

console.log(something);
// Kyle

console.log(window.something);
// 42
```

The `let` declaration adds a `something` global variable but not a global object property (see Chapter 3). The effect then is that the `something` lexical identifier shadows the `something` global object property.

It's almost certainly a bad idea to create a divergence between the global object and the global scope. Readers of your code will almost certainly be tripped up.

A simple way to avoid this gotcha with global declarations: always use `var` for globals. Reserve `let` and `const` for block scopes (see "Scoping with Blocks" in Chapter 6).

DOM Globals I asserted that a browser-hosted JS environment has the most *pure* global scope behavior we'll see. However, it's not entirely *pure*.

One surprising behavior in the global scope you may encounter with browser-based JS applications: a DOM element with an `id` attribute automatically creates a global variable that references it.

Consider this markup:

```
<ul id="my-todo-list">
  <li id="first">Write a book</li>
  ..
</ul>
```

And the JS for that page could include:

```
first;  
// <li id="first">..</li>  
  
window["my-todo-list"];  
// <ul id="my-todo-list">..</ul>
```

If the `id` value is a valid lexical name (like `first`), the lexical variable is created. If not, the only way to access that global is through the global object (`window[...]`).

The auto-registration of all `id`-bearing DOM elements as global variables is an old legacy browser behavior that nevertheless must remain because so many old sites still rely on it. My advice is never to use these global variables, even though they will always be silently created.

What's in a (Window) Name? Another global scope oddity in browser-based JS:

```
var name = 42;  
  
console.log(name, typeof name);  
// "42" string
```

`window.name` is a pre-defined “global” in a browser context; it’s a property on the global object, so it seems like a normal global variable (yet it’s anything but “normal”).

We used `var` for our declaration, which **does not** shadow the pre-defined `name` global property. That means, effectively, the `var` declaration is ignored, since there’s already a global scope object property of that name. As we discussed earlier, had we used `let name`, we would have shadowed `window.name` with a separate global `name` variable.

But the truly surprising behavior is that even though we assigned the number 42 to `name` (and thus `window.name`), when we then retrieve its value, it’s a string “42”! In this case, the weirdness is because `name` is actually a pre-defined getter/setter on the `window` object, which insists on its value being a string value. Yikes!

With the exception of some rare corner cases like DOM element ID’s and `window.name`, JS running as a standalone file in a browser page has some of the most *pure* global scope behavior we will encounter.

Web Workers

Web Workers are a web platform extension on top of browser-JS behavior, which allows a JS file to run in a completely separate thread (operating system wise) from the thread that’s running the main JS program.

Since these Web Worker programs run on a separate thread, they're restricted in their communications with the main application thread, to avoid/limit race conditions and other complications. Web Worker code does not have access to the DOM, for example. Some web APIs are, however, made available to the worker, such as `navigator`.

Since a Web Worker is treated as a wholly separate program, it does not share the global scope with the main JS program. However, the browser's JS engine is still running the code, so we can expect similar *purity* of its global scope behavior. Since there is no DOM access, the `window` alias for the global scope doesn't exist.

In a Web Worker, the global object reference is typically made using `self`:

```
var studentName = "Kyle";
let studentID = 42;

function hello() {
  console.log(`Hello, ${self.studentName}`);
}

self.hello();
// Hello, Kyle!

self.studentID;
// undefined
```

Just as with main JS programs, `var` and `function` declarations create mirrored properties on the global object (aka, `self`), where other declarations (`let`, etc) do not.

So again, the global scope behavior we're seeing here is about as *pure* as it gets for running JS programs; perhaps it's even more *pure* since there's no DOM to muck things up!

Developer Tools Console/REPL

Recall from Chapter 1 in *Get Started* that Developer Tools don't create a completely adherent JS environment. They do process JS code, but they also lean in favor of the UX interaction being most friendly to developers (aka, developer experience, or DX).

In some cases, favoring DX when typing in short JS snippets, over the normal strict steps expected for processing a full JS program, produces observable differences in code behavior between programs and tools. For example, certain error conditions applicable to a JS program may be relaxed and not displayed when the code is entered into a developer tool.

With respect to our discussions here about scope, such observable differences in

behavior may include:

- The behavior of the global scope
- Hoisting (see Chapter 5)
- Block-scoping declarators (`let` / `const`, see Chapter 6) when used in the outermost scope

Although it might seem, while using the console/REPL, that statements entered in the outermost scope are being processed in the real global scope, that's not quite accurate. Such tools typically emulate the global scope position to an extent; it's emulation, not strict adherence. These tool environments prioritize developer convenience, which means that at times (such as with our current discussions regarding scope), observed behavior may deviate from the JS specification.

The take-away is that Developer Tools, while optimized to be convenient and useful for a variety of developer activities, are **not** suitable environments to determine or verify explicit and nuanced behaviors of an actual JS program context.

ES Modules (ESM)

ES6 introduced first-class support for the module pattern (covered in Chapter 8). One of the most obvious impacts of using ESM is how it changes the behavior of the observable top-level scope in a file.

Recall this code snippet from earlier (which we'll adjust to ESM format by using the `export` keyword):

```
var studentName = "Kyle";  
  
function hello() {  
    console.log(`Hello, ${studentName}!`);  
}  
  
hello();  
// Hello, Kyle!  
  
export hello;
```

If that code is in a file that's loaded as an ES module, it will still run exactly the same. However, the observable effects, from the overall application perspective, will be different.

Despite being declared at the top level of the (module) file, in the outermost obvious scope, `studentName` and `hello` are not global variables. Instead, they are module-wide, or if you prefer, “module-global.”

However, in a module there's no implicit "module-wide scope object" for these top-level declarations to be added to as properties, as there is when declarations appear in the top-level of non-module JS files. This is not to say that global variables cannot exist or be accessed in such programs. It's just that global variables don't get *created* by declaring variables in the top-level scope of a module.

The module's top-level scope is descended from the global scope, almost as if the entire contents of the module were wrapped in a function. Thus, all variables that exist in the global scope (whether they're on the global object or not!) are available as lexical identifiers from inside the module's scope.

ESM encourages a minimization of reliance on the global scope, where you import whatever modules you may need for the current module to operate. As such, you less often see usage of the global scope or its global object.

However, as noted earlier, there are still plenty of JS and web globals that you will continue to access from the global scope, whether you realize it or not!

Node

One aspect of Node that often catches JS developers off-guard is that Node treats every single .js file that it loads, including the main one you start the Node process with, as a *module* (ES module or CommonJS module, see Chapter 8). The practical effect is that the top level of your Node programs **is never actually the global scope**, the way it is when loading a non-module file in the browser.

As of time of this writing, Node has recently added support for ES modules. But additionally, Node has from its beginning supported a module format referred to as "CommonJS", which looks like this:

```
var studentName = "Kyle";  
  
function hello() {  
    console.log(`Hello, ${studentName}!`);  
}  
  
hello();  
// Hello, Kyle!  
  
module.exports.hello = hello;
```

Before processing, Node effectively wraps such code in a function, so that the `var` and `function` declarations are contained in that wrapping function's scope, **not** treated as global variables.

Envision the preceding code as being seen by Node as this (illustrative, not actual):

```

function Module(module, require, __dirname, ...) {
  var studentName = "Kyle";

  function hello() {
    console.log(`Hello, ${studentName}!`);
  }

  hello();
  // Hello, Kyle!

  module.exports.hello = hello;
}

```

Node then essentially invokes the added `Module(..)` function to run your module. You can clearly see here why `studentName` and `hello` identifiers are not global, but rather declared in the module scope.

As noted earlier, Node defines a number of “globals” like `require()`, but they’re not actually identifiers in the global scope (nor properties of the global object). They’re injected in the scope of every module, essentially a bit like the parameters listed in the `Module(..)` function declaration.

So how do you define actual global variables in Node? The only way to do so is to add properties to another of Node’s automatically provided “globals,” which is ironically called `global`. `global` is a reference to the real global scope object, somewhat like using `window` in a browser JS environment.

Consider:

```

global.studentName = "Kyle";

function hello() {
  console.log(`Hello, ${studentName}!`);
}

hello();
// Hello, Kyle!

module.exports.hello = hello;

```

Here we add `studentName` as a property on the `global` object, and then in the `console.log(..)` statement we’re able to access `studentName` as a normal global variable.

Remember, the identifier `global` is not defined by JS; it’s specifically defined by Node.

Global This

Reviewing the JS environments we've looked at so far, a program may or may not:

- Declare a global variable in the top-level scope with `var` or `function` declarations—or `let`, `const`, and `class`.
- Also add global variables declarations as properties of the global scope object if `var` or `function` are used for the declaration.
- Refer to the global scope object (for adding or retrieving global variables, as properties) with `window`, `self`, or `global`.

I think it's fair to say that global scope access and behavior is more complicated than most developers assume, as the preceding sections have illustrated. But the complexity is never more obvious than in trying to nail down a universally applicable reference to the global scope object.

Yet another “trick” for obtaining a reference to the global scope object looks like:

```
const theGlobalScopeObject =
  (new Function("return this"))();
```

NOTE:

A function can be dynamically constructed from code stored in a string value with the `Function()` constructor, similar to `eval(..)` (see “Cheating: Runtime Scope Modifications” in Chapter 1). Such a function will automatically be run in non-strict-mode (for legacy reasons) when invoked with the normal () function invocation as shown; its `this` will point at the global object. See the third book in the series, *Objects & Classes*, for more information on determining `this` bindings.

So, we have `window`, `self`, `global`, and this ugly `new Function(..)` trick. That's a lot of different ways to try to get at this global object. Each has its pros and cons.

Why not introduce yet another!?!?

As of ES2020, JS has finally defined a standardized reference to the global scope object, called `globalThis`. So, subject to the recency of the JS engines your code runs in, you can use `globalThis` in place of any of those other approaches.

We could even attempt to define a cross-environment polyfill that's safer across pre-`globalThis` JS environments, such as:

```
const theGlobalScopeObject =
  (typeof globalThis != "undefined") ? globalThis :
  (typeof global != "undefined") ? global :
```

```
(typeof window != "undefined") ? window :  
(typeof self != "undefined") ? self :  
(new Function("return this"))();
```

Phew! That's certainly not ideal, but it works if you find yourself needing a reliable global scope reference.

(The proposed name `globalThis` was fairly controversial while the feature was being added to JS. Specifically, I and many others felt the “this” reference in its name was misleading, since the reason you reference this object is to access to the global scope, never to access some sort of global/default `this` binding. There were many other names considered, but for a variety of reasons ruled out. Unfortunately, the name chosen ended up as a last resort. If you plan to interact with the global scope object in your programs, to reduce confusion, I strongly recommend choosing a better name, such as (the laughably long but accurate!) `theGlobalScopeObject` used here.)

Globally Aware

The global scope is present and relevant in every JS program, even though modern patterns for organizing code into modules de-emphasizes much of the reliance on storing identifiers in that namespace.

Still, as our code proliferates more and more beyond the confines of the browser, it's especially important we have a solid grasp on the differences in how the global scope (and global scope object!) behave across different JS environments.

With the big picture of global scope now sharper in focus, the next chapter again descends into the deeper details of lexical scope, examining how and when variables can be used.

You Don't Know JS Yet: Scope & Closures - 2nd Edition

Chapter 5: The (Not So) Secret Lifecycle of Variables

By now you should have a decent grasp of the nesting of scopes, from the global scope downward—called a program’s scope chain.

But just knowing which scope a variable comes from is only part of the story. If a variable declaration appears past the first statement of a scope, how will any references to that identifier *before* the declaration behave? What happens if you try to declare the same variable twice in a scope?

JS’s particular flavor of lexical scope is rich with nuance in how and when variables come into existence and become available to the program.

When Can I Use a Variable?

At what point does a variable become available to use within its scope? There may seem to be an obvious answer: *after* the variable has been declared/created. Right? Not quite.

Consider:

```
greeting();
// Hello!

function greeting() {
  console.log("Hello!");
}
```

This code works fine. You may have seen or even written code like it before. But did you ever wonder how or why it works? Specifically, why can you access the identifier `greeting` from line 1 (to retrieve and execute a function reference), even though the `greeting()` function declaration doesn’t occur until line 4?

Recall Chapter 1 points out that all identifiers are registered to their respective scopes during compile time. Moreover, every identifier is *created* at the beginning of the scope it belongs to, **every time that scope is entered**.

The term most commonly used for a variable being visible from the beginning of its enclosing scope, even though its declaration may appear further down in the scope, is called **hoisting**.

But hoisting alone doesn’t fully answer the question. We can see an identifier called `greeting` from the beginning of the scope, but why can we **call** the `greeting()` function before it’s been declared?

In other words, how does the variable `greeting` have any value (the function reference) assigned to it, from the moment the scope starts running? The answer is a special characteristic of formal `function` declarations, called *function hoisting*. When a `function` declaration's name identifier is registered at the top of its scope, it's additionally auto-initialized to that function's reference. That's why the function can be called throughout the entire scope!

One key detail is that both *function hoisting* and `var`-flavored *variable hoisting* attach their name identifiers to the nearest enclosing **function scope** (or, if none, the global scope), not a block scope.

NOTE:

Declarations with `let` and `const` still hoist (see the TDZ discussion later in this chapter). But these two declaration forms attach to their enclosing block rather than just an enclosing function as with `var` and `function` declarations. See “Scoping with Blocks” in Chapter 6 for more information.

Hoisting: Declaration vs. Expression

Function hoisting only applies to formal `function` declarations (specifically those which appear outside of blocks—see “FiB” in Chapter 6), not to `function` expression assignments. Consider:

```
greeting();
// TypeError

var greeting = function greeting() {
  console.log("Hello!");
};
```

Line 1 (`greeting();`) throws an error. But the *kind* of error thrown is very important to notice. A `TypeError` means we're trying to do something with a value that is not allowed. Depending on your JS environment, the error message would say something like, “‘undefined’ is not a function,” or more helpfully, “‘greeting’ is not a function.”

Notice that the error is **not** a `ReferenceError`. JS isn't telling us that it couldn't find `greeting` as an identifier in the scope. It's telling us that `greeting` was found but doesn't hold a function reference at that moment. Only functions can be invoked, so attempting to invoke some non-function value results in an error.

But what does `greeting` hold, if not the function reference?

In addition to being hoisted, variables declared with `var` are also automatically initialized to `undefined` at the beginning of their scope—again, the nearest enclosing function, or the global. Once initialized, they're available to be used (assigned to, retrieved from, etc.) throughout the whole scope.

So on that first line, `greeting` exists, but it holds only the default `undefined` value. It's not until line 4 that `greeting` gets assigned the function reference.

Pay close attention to the distinction here. A `function` declaration is hoisted **and initialized to its function value** (again, called *function hoisting*). A `var` variable is also hoisted, and then auto-initialized to `undefined`. Any subsequent `function` expression assignments to that variable don't happen until that assignment is processed during runtime execution.

In both cases, the name of the identifier is hoisted. But the function reference association isn't handled at initialization time (beginning of the scope) unless the identifier was created in a formal `function` declaration.

Variable Hoisting

Let's look at another example of *variable hoisting*:

```
greeting = "Hello!";
console.log(greeting);
// Hello!

var greeting = "Howdy!";
```

Though `greeting` isn't declared until line 5, it's available to be assigned to as early as line 1. Why?

There's two necessary parts to the explanation:

- the identifier is hoisted,
- **and** it's automatically initialized to the value `undefined` from the top of the scope.

NOTE:

Using *variable hoisting* of this sort probably feels unnatural, and many readers might rightly want to avoid relying on it in their programs. But should all hoisting (including *function hoisting*) be avoided? We'll explore these different perspectives on hoisting in more detail in Appendix A.

Hoisting: Yet Another Metaphor

Chapter 2 was full of metaphors (to illustrate scope), but here we are faced with yet another: hoisting itself. Rather than hoisting being a concrete execution step the JS engine performs, it's more useful to think of hoisting as a visualization of various actions JS takes in setting up the program **before execution**.

The typical assertion of what hoisting means: *lifting*—like lifting a heavy weight upward—any identifiers all the way to the top of a scope. The explanation often asserted is that the JS engine will actually *rewrite* that program before execution, so that it looks more like this:

```

var greeting;           // hoisted declaration
greeting = "Hello!";   // the original line 1
console.log(greeting); // Hello!
greeting = "Howdy!";   // `var` is gone!

```

The hoisting (metaphor) proposes that JS pre-processes the original program and re-arranges it a bit, so that all the declarations have been moved to the top of their respective scopes, before execution. Moreover, the hoisting metaphor asserts that **function** declarations are, in their entirety, hoisted to the top of each scope. Consider:

```

studentName = "Suzy";
greeting();
// Hello Suzy!

function greeting() {
  console.log(`Hello ${ studentName }!`);
}
var studentName;

```

The “rule” of the hoisting metaphor is that function declarations are hoisted first, then variables are hoisted immediately after all the functions. Thus, the hoisting story suggests that program is *re-arranged* by the JS engine to look like this:

```

function greeting() {
  console.log(`Hello ${ studentName }!`);
}
var studentName;

studentName = "Suzy";
greeting();
// Hello Suzy!

```

This hoisting metaphor is convenient. Its benefit is allowing us to hand wave over the magical look-ahead pre-processing necessary to find all these declarations buried deep in scopes and somehow move (hoist) them to the top; we can just think about the program as if it’s executed by the JS engine in a **single pass**, top-down.

Single-pass definitely seems more straightforward than Chapter 1’s assertion of a two-phase processing.

Hoisting as a mechanism for re-ordering code may be an attractive simplification, but it’s not accurate. The JS engine doesn’t actually re-arrange the code. It can’t magically look ahead and find declarations; the only way to accurately find them, as well as all the scope boundaries in the program, would be to fully parse the code.

Guess what parsing is? The first phase of the two-phase processing! There's no magical mental gymnastics that gets around that fact.

So if the hoisting metaphor is (at best) inaccurate, what should we do with the term? I think it's still useful—indeed, even members of TC39 regularly use it!—but I don't think we should claim it's an actual re-arrangement of source code.

WARNING:

Incorrect or incomplete mental models often still seem sufficient because they can occasionally lead to accidental right answers. But in the long run it's harder to accurately analyze and predict outcomes if your thinking isn't particularly aligned with how the JS engine works.

I assert that hoisting *should* be used to refer to the **compile-time operation** of generating runtime instructions for the automatic registration of a variable at the beginning of its scope, each time that scope is entered.

That's a subtle but important shift, from hoisting as a runtime behavior to its proper place among compile-time tasks.

Re-declaration?

What do you think happens when a variable is declared more than once in the same scope? Consider:

```
var studentName = "Frank";
console.log(studentName);
// Frank

var studentName;
console.log(studentName); // ???
```

What do you expect to be printed for that second message? Many believe the second `var studentName` has re-declared the variable (and thus “reset” it), so they expect `undefined` to be printed.

But is there such a thing as a variable being “re-declared” in the same scope? No.

If you consider this program from the perspective of the hoisting metaphor, the code would be re-arranged like this for execution purposes:

```
var studentName;
var studentName; // clearly a pointless no-op!

studentName = "Frank";
console.log(studentName);
```

```
// Frank

console.log(studentName);
// Frank
```

Since hoisting is actually about registering a variable at the beginning of a scope, there's nothing to be done in the middle of the scope where the original program actually had the second `var studentName` statement. It's just a no-op(eration), a pointless statement.

TIP:

In the style of the conversation narrative from Chapter 2, *Compiler* would find the second `var` declaration statement and ask the *Scope Manager* if it had already seen a `studentName` identifier; since it had, there wouldn't be anything else to do.

It's also important to point out that `var studentName;` doesn't mean `var studentName = undefined;`, as most assume. Let's prove they're different by considering this variation of the program:

```
var studentName = "Frank";
console.log(studentName); // Frank

var studentName;
console.log(studentName); // Frank <-- still!

// let's add the initialization explicitly
var studentName = undefined;
console.log(studentName); // undefined <-- see!?
```

See how the explicit `= undefined` initialization produces a different outcome than assuming it happens implicitly when omitted? In the next section, we'll revisit this topic of initialization of variables from their declarations.

A repeated `var` declaration of the same identifier name in a scope is effectively a do-nothing operation. Here's another illustration, this time across a function of the same name:

```
var greeting;

function greeting() {
    console.log("Hello!");
}

// basically, a no-op
var greeting;
```

```
typeof greeting;           // "function"  
  
var greeting = "Hello!";  
  
typeof greeting;           // "string"
```

The first `greeting` declaration registers the identifier to the scope, and because it's a `var` the auto-initialization will be `undefined`. The `function` declaration doesn't need to re-register the identifier, but because of *function hoisting* it overrides the auto-initialization to use the function reference. The second `var` `greeting` by itself doesn't do anything since `greeting` is already an identifier and *function hoisting* already took precedence for the auto-initialization.

Actually assigning "Hello!" to `greeting` changes its value from the initial `function greeting()` to the string; `var` itself doesn't have any effect.

What about repeating a declaration within a scope using `let` or `const`?

```
let studentName = "Frank";  
  
console.log(studentName);  
  
let studentName = "Suzy";
```

This program will not execute, but instead immediately throw a `SyntaxError`. Depending on your JS environment, the error message will indicate something like: "studentName has already been declared." In other words, this is a case where attempted "re-declaration" is explicitly not allowed!

It's not just that two declarations involving `let` will throw this error. If either declaration uses `let`, the other can be either `let` or `var`, and the error will still occur, as illustrated with these two variations:

```
var studentName = "Frank";  
  
let studentName = "Suzy";  
and:  
let studentName = "Frank";
```

```
var studentName = "Suzy";
```

In both cases, a `SyntaxError` is thrown on the *second* declaration. In other words, the only way to "re-declare" a variable is to use `var` for all (two or more) of its declarations.

But why disallow it? The reason for the error is not technical per se, as `var` "re-declaration" has always been allowed; clearly, the same allowance could have been made for `let`.

It's really more of a "social engineering" issue. "Re-declaration" of variables is seen by some, including many on the TC39 body, as a bad habit that can lead to program bugs. So when ES6 introduced `let`, they decided to prevent "re-declaration" with an error.

NOTE:

This is of course a stylistic opinion, not really a technical argument. Many developers agree with the position, and that's probably in part why TC39 included the error (as well as `let` conforming to `const`). But a reasonable case could have been made that staying consistent with `var`'s precedent was more prudent, and that such opinion-enforcement was best left to opt-in tooling like linters. In Appendix A, we'll explore whether `var` (and its associated behavior, like "re-declaration") can still be useful in modern JS.

When *Compiler* asks *Scope Manager* about a declaration, if that identifier has already been declared, and if either/both declarations were made with `let`, an error is thrown. The intended signal to the developer is "Stop relying on sloppy re-declaration!"

Constants?

The `const` keyword is more constrained than `let`. Like `let`, `const` cannot be repeated with the same identifier in the same scope. But there's actually an overriding technical reason why that sort of "re-declaration" is disallowed, unlike `let` which disallows "re-declaration" mostly for stylistic reasons.

The `const` keyword requires a variable to be initialized, so omitting an assignment from the declaration results in a `SyntaxError`:

```
const empty; // SyntaxError

const declarations create variables that cannot be re-assigned:

const studentName = "Frank";
console.log(studentName);
// Frank

studentName = "Suzy"; // TypeError
```

The `studentName` variable cannot be re-assigned because it's declared with a `const`.

WARNING:

The error thrown when re-assigning `studentName` is a `TypeError`, not a `SyntaxError`. The subtle distinction here is actually pretty important, but unfortunately far too easy to miss. Syntax errors represent faults in the program that stop it from even starting execution. Type errors represent faults that arise during program execution. In the preceding snippet, "Frank" is printed out before we process the re-assignment of `studentName`, which then throws the error.

So if `const` declarations cannot be re-assigned, and `const` declarations always require assignments, then we have a clear technical reason why `const` must disallow any "re-declarations": any `const` "re-declaration" would also necessarily be a `const` re-assignment, which can't be allowed!

```
const studentName = "Frank";
```

```
// obviously this must be an error
const studentName = "Suzy";
```

Since `const` "re-declaration" must be disallowed (on those technical grounds), TC39 essentially felt that `let` "re-declaration" should be disallowed as well, for consistency. It's debatable if this was the best choice, but at least we have the reasoning behind the decision.

Loops

So it's clear from our previous discussion that JS doesn't really want us to "re-declare" our variables within the same scope. That probably seems like a straightforward admonition, until you consider what it means for repeated execution of declaration statements in loops. Consider:

```
var keepGoing = true;
while (keepGoing) {
    let value = Math.random();
    if (value > 0.5) {
        keepGoing = false;
    }
}
```

Is `value` being "re-declared" repeatedly in this program? Will we get errors thrown? No.

All the rules of scope (including "re-declaration" of `let`-created variables) are applied *per scope instance*. In other words, each time a scope is entered during execution, everything resets.

Each loop iteration is its own new scope instance, and within each scope instance, `value` is only being declared once. So there's no attempted "re-declaration," and

thus no error. Before we consider other loop forms, what if the `value` declaration in the previous snippet were changed to a `var`?

```
var keepGoing = true;
while (keepGoing) {
  var value = Math.random();
  if (value > 0.5) {
    keepGoing = false;
  }
}
```

Is `value` being “re-declared” here, especially since we know `var` allows it? No. Because `var` is not treated as a block-scoping declaration (see Chapter 6), it attaches itself to the global scope. So there’s just one `value` variable, in the same scope as `keepGoing` (global scope, in this case). No “re-declaration” here, either!

One way to keep this all straight is to remember that `var`, `let`, and `const` keywords are effectively *removed* from the code by the time it starts to execute. They’re handled entirely by the compiler.

If you mentally erase the declarator keywords and then try to process the code, it should help you decide if and when (re-)declarations might occur.

What about “re-declaration” with other loop forms, like `for`-loops?

```
for (let i = 0; i < 3; i++) {
  let value = i * 10;
  console.log(`#${i}: ${value}`);
}
// 0: 0
// 1: 10
// 2: 20
```

It should be clear that there’s only one `value` declared per scope instance. But what about `i`? Is it being “re-declared”?

To answer that, consider what scope `i` is in. It might seem like it would be in the outer (in this case, global) scope, but it’s not. It’s in the scope of `for`-loop body, just like `value` is. In fact, you could sorta think about that loop in this more verbose equivalent form:

```
{
  // a fictional variable for illustration
  let $$i = 0;

  for (/* nothing */; $$i < 3; $$i++) {
    // here's our actual loop `i`!
    let i = $$i;
```

```

        let value = i * 10;
        console.log(`#${ i }: ${ value }`);
    }
    // 0: 0
    // 1: 10
    // 2: 20
}

```

Now it should be clear: the `i` and `value` variables are both declared exactly once **per scope instance**. No “re-declaration” here.

What about other `for`-loop forms?

```

for (let index in students) {
    // this is fine
}

for (let student of students) {
    // so is this
}

```

Same thing with `for..in` and `for..of` loops: the declared variable is treated as *inside* the loop body, and thus is handled per iteration (aka, per scope instance). No “re-declaration.”

OK, I know you’re thinking that I sound like a broken record at this point. But let’s explore how `const` impacts these looping constructs. Consider:

```

var keepGoing = true;
while (keepGoing) {
    // ooo, a shiny constant!
    const value = Math.random();
    if (value > 0.5) {
        keepGoing = false;
    }
}

```

Just like the `let` variant of this program we saw earlier, `const` is being run exactly once within each loop iteration, so it’s safe from “re-declaration” troubles. But things get more complicated when we talk about `for`-loops.

`for..in` and `for..of` are fine to use with `const`:

```

for (const index in students) {
    // this is fine
}

for (const student of students) {
    // this is also fine
}

```

But not the general `for`-loop:

```
for (const i = 0; i < 3; i++) {  
  // oops, this is going to fail with  
  // a Type Error after the first iteration  
}
```

What's wrong here? We could use `let` just fine in this construct, and we asserted that it creates a new `i` for each loop iteration scope, so it doesn't even seem to be a "re-declaration."

Let's mentally "expand" that loop like we did earlier:

```
{  
  // a fictional variable for illustration  
  const $$i = 0;  
  
  for ( ; $$i < 3; $$i++) {  
    // here's our actual loop `i`!  
    const i = $$i;  
    // ...  
  }  
}
```

Do you spot the problem? Our `i` is indeed just created once inside the loop. That's not the problem. The problem is the conceptual `$$i` that must be incremented each time with the `$$i++` expression. That's **re-assignment** (not "re-declaration"), which isn't allowed for constants.

Remember, this "expanded" form is only a conceptual model to help you intuit the source of the problem. You might wonder if JS could have effectively made the `const $$i = 0` instead into `let $ii = 0`, which would then allow `const` to work with our classic `for`-loop? It's possible, but then it could have introduced potentially surprising exceptions to `for`-loop semantics.

For example, it would have been a rather arbitrary (and likely confusing) nuanced exception to allow `i++` in the `for`-loop header to skirt strictness of the `const` assignment, but not allow other re-assignments of `i` inside the loop iteration, as is sometimes useful.

The straightforward answer is: `const` can't be used with the classic `for`-loop form because of the required re-assignment.

Interestingly, if you don't do re-assignment, then it's valid:

```
var keepGoing = true;  
  
for (const i = 0; keepGoing; /* nothing here */ ) {  
  keepGoing = (Math.random() > 0.5);  
  // ...  
}
```

That works, but it's pointless. There's no reason to declare `i` in that position with a `const`, since the whole point of such a variable in that position is **to be used for counting iterations**. Just use a different loop form, like a `while` loop, or use a `let`!

Uninitialized Variables (aka, TDZ)

With `var` declarations, the variable is “hoisted” to the top of its scope. But it's also automatically initialized to the `undefined` value, so that the variable can be used throughout the entire scope.

However, `let` and `const` declarations are not quite the same in this respect.

Consider:

```
console.log(studentName);  
// ReferenceError  
  
let studentName = "Suzy";
```

The result of this program is that a `ReferenceError` is thrown on the first line. Depending on your JS environment, the error message may say something like: “Cannot access `studentName` before initialization.”

NOTE:

The error message as seen here used to be much more vague or misleading. Thankfully, several of us in the community were successfully able to lobby for JS engines to improve this error message so it more accurately tells you what's wrong!

That error message is quite indicative of what's wrong: `studentName` exists on line 1, but it's not been initialized, so it cannot be used yet. Let's try this:

```
studentName = "Suzy"; // let's try to initialize it!  
// ReferenceError  
  
console.log(studentName);  
  
let studentName;
```

Oops. We still get the `ReferenceError`, but now on the first line where we're trying to assign to (aka, initialize!) this so-called “uninitialized” variable `studentName`. What's the deal?

The real question is, how do we initialize an uninitialized variable? For `let/const`, the **only way** to do so is with an assignment attached to a declaration statement. An assignment by itself is insufficient! Consider:

```
let studentName = "Suzy";
console.log(studentName); // Suzy
```

Here, we are initializing the `studentName` (in this case, to "Suzy" instead of `undefined`) by way of the `let` declaration statement form that's coupled with an assignment.

Alternatively:

```
// ..

let studentName;
// or:
// let studentName = undefined;

// ..

studentName = "Suzy";

console.log(studentName);
// Suzy
```

NOTE:

That's interesting! Recall from earlier, we said that `var studentName;` is *not* the same as `var studentName = undefined;`, but here with `let`, they behave the same. The difference comes down to the fact that `var studentName` automatically initializes at the top of the scope, where `let studentName` does not.

Remember that we've asserted a few times so far that *Compiler* ends up removing any `var/let/const` declarators, replacing them with the instructions at the top of each scope to register the appropriate identifiers.

So if we analyze what's going on here, we see that an additional nuance is that *Compiler* is also adding an instruction in the middle of the program, at the point where the variable `studentName` was declared, to handle that declaration's auto-initialization. We cannot use the variable at any point prior to that initialization occurring. The same goes for `const` as it does for `let`.

The term coined by TC39 to refer to this *period of time* from the entering of a scope to where the auto-initialization of the variable occurs is: Temporal Dead Zone (TDZ).

The TDZ is the time window where a variable exists but is still uninitialized, and therefore cannot be accessed in any way. Only the execution of the instructions left by *Compiler* at the point of the original declaration can do that initialization.

After that moment, the TDZ is done, and the variable is free to be used for the rest of the scope.

A `var` also technically has a TDZ, but it's zero in length and thus unobservable to our programs! Only `let` and `const` have an observable TDZ.

By the way, “temporal” in TDZ does indeed refer to *time* not *position in code*. Consider:

```
askQuestion();
// ReferenceError

let studentName = "Suzy";

function askQuestion() {
  console.log(` ${studentName}, do you know?`);
}
```

Even though positionally the `console.log(..)` referencing `studentName` comes *after* the `let studentName` declaration, timing wise the `askQuestion()` function is invoked *before* the `let` statement is encountered, while `studentName` is still in its TDZ! Hence the error.

There's a common misconception that TDZ means `let` and `const` do not hoist. This is an inaccurate, or at least slightly misleading, claim. They definitely hoist.

The actual difference is that `let/const` declarations do not automatically initialize at the beginning of the scope, the way `var` does. The *debate* then is if the auto-initialization is *part of* hoisting, or not? I think auto-registration of a variable at the top of the scope (i.e., what I call “hoisting”) and auto-initialization at the top of the scope (to `undefined`) are distinct operations and shouldn't be lumped together under the single term “hoisting.”

We've already seen that `let` and `const` don't auto-initialize at the top of the scope. But let's prove that `let` and `const` *do* hoist (auto-register at the top of the scope), courtesy of our friend shadowing (see “Shadowing” in Chapter 3):

```
var studentName = "Kyle";

{
  console.log(studentName);
  // ???

  // ..

  let studentName = "Suzy";

  console.log(studentName);
```

```
// Suzy
}
```

What's going to happen with the first `console.log(..)` statement? If `let studentName` didn't hoist to the top of the scope, then the first `console.log(..)` *should* print "Kyle", right? At that moment, it would seem, only the outer `studentName` exists, so that's the variable `console.log(..)` should access and print.

But instead, the first `console.log(..)` throws a TDZ error, because in fact, the inner scope's `studentName` **was** hoisted (auto-registered at the top of the scope). What **didn't** happen (yet!) was the auto-initialization of that inner `studentName`; it's still uninitialized at that moment, hence the TDZ violation!

So to summarize, TDZ errors occur because `let/const` declarations *do* hoist their declarations to the top of their scopes, but unlike `var`, they defer the auto-initialization of their variables until the moment in the code's sequencing where the original declaration appeared. This window of time (hint: temporal), whatever its length, is the TDZ.

How can you avoid TDZ errors?

My advice: always put your `let` and `const` declarations at the top of any scope. Shrink the TDZ window to zero (or near zero) length, and then it'll be moot.

But why is TDZ even a thing? Why didn't TC39 dictate that `let/const` auto-initialize the way `var` does? Just be patient, we'll come back to explore the *why* of TDZ in Appendix A.

Finally Initialized

Working with variables has much more nuance than it seems at first glance. *Hoisting*, *(re)declaration*, and the *TDZ* are common sources of confusion for developers, especially those who have worked in other languages before coming to JS. Before moving on, make sure your mental model is fully grounded on these aspects of JS scope and variables.

Hoisting is generally cited as an explicit mechanism of the JS engine, but it's really more a metaphor to describe the various ways JS handles variable declarations during compilation. But even as a metaphor, hoisting offers useful structure for thinking about the life-cycle of a variable—when it's created, when it's available to use, when it goes away.

Declaration and re-declaration of variables tend to cause confusion when thought of as runtime operations. But if you shift to compile-time thinking for these operations, the quirks and *shadows* diminish.

The TDZ (temporal dead zone) error is strange and frustrating when encountered. Fortunately, TDZ is relatively straightforward to avoid if you're always careful to place `let/const` declarations at the top of any scope.

As you successfully navigate these twists and turns of variable scope, the next chapter will lay out the factors that guide our decisions to place our declarations in various scopes, especially nested blocks.

You Don't Know JS Yet: Scope & Closures - 2nd Edition

Chapter 6: Limiting Scope Exposure

So far our focus has been explaining the mechanics of how scopes and variables work. With that foundation now firmly in place, our attention raises to a higher level of thinking: decisions and patterns we apply across the whole program.

To begin, we're going to look at how and why we should be using different levels of scope (functions and blocks) to organize our program's variables, specifically to reduce scope over-exposure.

Least Exposure

It makes sense that functions define their own scopes. But why do we need blocks to create scopes as well?

Software engineering articulates a fundamental discipline, typically applied to software security, called “The Principle of Least Privilege” (POLP). ¹ And a variation of this principle that applies to our current discussion is typically labeled as “Least Exposure” (POLE).

POLP expresses a defensive posture to software architecture: components of the system should be designed to function with least privilege, least access, least exposure. If each piece is connected with minimum-necessary capabilities, the overall system is stronger from a security standpoint, because a compromise or failure of one piece has a minimized impact on the rest of the system.

If POLP focuses on system-level component design, the POLE *Exposure* variant focuses on a lower level; we'll apply it to how scopes interact with each other.

In following POLE, what do we want to minimize the exposure of? Simply: the variables registered in each scope.

Think of it this way: why shouldn't you just place all the variables of your program out in the global scope? That probably immediately feels like a bad idea, but it's worth considering why that is. When variables used by one part of the program are exposed to another part of the program, via scope, there are three main hazards that often arise:

- **Naming Collisions:** if you use a common and useful variable/function name in two different parts of the program, but the identifier comes from one shared scope (like the global scope), then name collision occurs, and it's very likely that bugs will occur as one part uses the variable/function in a way the other part doesn't expect.

¹ *Principle of Least Privilege*, https://en.wikipedia.org/wiki/Principle_of_least_privilege, 3 March 2020.

For example, imagine if all your loops used a single global `i` index variable, and then it happens that one loop in a function is running during an iteration of a loop from another function, and now the shared `i` variable gets an unexpected value.

- **Unexpected Behavior:** if you expose variables/functions whose usage is otherwise *private* to a piece of the program, it allows other developers to use them in ways you didn't intend, which can violate expected behavior and cause bugs.

For example, if your part of the program assumes an array contains all numbers, but someone else's code accesses and modifies the array to include booleans and strings, your code may then misbehave in unexpected ways.

Worse, exposure of *private* details invites those with mal-intent to try to work around limitations you have imposed, to do things with your part of the software that shouldn't be allowed.

- **Unintended Dependency:** if you expose variables/functions unnecessarily, it invites other developers to use and depend on those otherwise *private* pieces. While that doesn't break your program today, it creates a refactoring hazard in the future, because now you cannot as easily refactor that variable or function without potentially breaking other parts of the software that you don't control.

For example, if your code relies on an array of numbers, and you later decide it's better to use some other data structure instead of an array, you now must take on the liability of adjusting other affected parts of the software.

POLE, as applied to variable/function scoping, essentially says, default to exposing the bare minimum necessary, keeping everything else as private as possible. Declare variables in as small and deeply nested of scopes as possible, rather than placing everything in the global (or even outer function) scope.

If you design your software accordingly, you have a much greater chance of avoiding (or at least minimizing) these three hazards.

Consider:

```
function diff(x,y) {
  if (x > y) {
    let tmp = x;
    x = y;
    y = tmp;
  }

  return y - x;
}
```

```
diff(3,7);      // 4
diff(7,5);      // 2
```

In this `diff(..)` function, we want to ensure that `y` is greater than or equal to `x`, so that when we subtract (`y - x`), the result is 0 or larger. If `x` is initially larger (the result would be negative!), we swap `x` and `y` using a `tmp` variable, to keep the result positive.

In this simple example, it doesn't seem to matter whether `tmp` is inside the `if` block or whether it belongs at the function level—it certainly shouldn't be a global variable! However, following the POLE principle, `tmp` should be as hidden in scope as possible. So we block scope `tmp` (using `let`) to the `if` block.

Hiding in Plain (Function) Scope

It should now be clear why it's important to hide our variable and function declarations in the lowest (most deeply nested) scopes possible. But how do we do so?

We've already seen the `let` and `const` keywords, which are block scoped declarations; we'll come back to them in more detail shortly. But first, what about hiding `var` or `function` declarations in scopes? That can easily be done by wrapping a `function` scope around a declaration.

Let's consider an example where `function` scoping can be useful.

The mathematical operation “factorial” (notated as “6!”) is the multiplication of a given integer against all successively lower integers down to 1—actually, you can stop at 2 since multiplying 1 does nothing. In other words, “6!” is the same as “6 * 5!”, which is the same as “6 * 5 * 4!”, and so on. Because of the nature of the math involved, once any given integer's factorial (like “4!”) has been calculated, we shouldn't need to do that work again, as it'll always be the same answer.

So if you naively calculate factorial for 6, then later want to calculate factorial for 7, you might unnecessarily re-calculate the factorials of all the integers from 2 up to 6. If you're willing to trade memory for speed, you can solve that wasted computation by caching each integer's factorial as it's calculated:

```
var cache = {};  
  
function factorial(x) {  
  if (x < 2) return 1;  
  if (!(x in cache)) {  
    cache[x] = x * factorial(x - 1);  
  }  
  return cache[x];  
}
```

```

factorial(6);
// 720

cache;
// {
//   "2": 2,
//   "3": 6,
//   "4": 24,
//   "5": 120,
//   "6": 720
// }

factorial(7);
// 5040

```

We're storing all the computed factorials in `cache` so that across multiple calls to `factorial(..)`, the previous computations remain. But the `cache` variable is pretty obviously a *private* detail of how `factorial(..)` works, not something that should be exposed in an outer scope—especially not the global scope.

NOTE:

`factorial(..)` here is recursive—a call to itself is made from inside—but that's just for brevity of code sake; a non-recursive implementation would yield the same scoping analysis with respect to `cache`.

However, fixing this over-exposure issue is not as simple as hiding the `cache` variable inside `factorial(..)`, as it might seem. Since we need `cache` to survive multiple calls, it must be located in a scope outside that function. So what can we do?

Define another middle scope (between the outer/global scope and the inside of `factorial(..)`) for `cache` to be located:

```

// outer/global scope

function hideTheCache() {
  // "middle scope", where we hide `cache`
  var cache = {};

  return factorial;

  // *****

  function factorial(x) {
    // inner scope

```

```

        if (x < 2) return 1;
        if (!(x in cache)) {
            cache[x] = x * factorial(x - 1);
        }
        return cache[x];
    }
}

var factorial = hideTheCache();

factorial(6);
// 720

factorial(7);
// 5040

```

The `hideTheCache()` function serves no other purpose than to create a scope for `cache` to persist in across multiple calls to `factorial(..)`. But for `factorial(..)` to have access to `cache`, we have to define `factorial(..)` inside that same scope. Then we return the function reference, as a value from `hideTheCache()`, and store it in an outer scope variable, also named `factorial`. Now as we call `factorial(..)` (multiple times!), its persistent `cache` stays hidden yet accessible only to `factorial(..)!`

OK, but... it's going to be tedious to define (and name!) a `hideTheCache(..)` function scope each time such a need for variable/function hiding occurs, especially since we'll likely want to avoid name collisions with this function by giving each occurrence a unique name. Ugh.

NOTE:

The illustrated technique—caching a function's computed output to optimize performance when repeated calls of the same inputs are expected—is quite common in the Functional Programming (FP) world, canonically referred to as “memoization”; this caching relies on closure (see Chapter 7). Also, there are memory usage concerns (addressed in “A Word About Memory” in Appendix B). FP libraries will usually provide an optimized and vetted utility for memoization of functions, which would take the place of `hideTheCache(..)` here. Memoization is beyond the *scope* (pun intended!) of our discussion, but see my *Functional-Light JavaScript* book for more information.

Rather than defining a new and uniquely named function each time one of those scope-only-for-the-purpose-of-hiding-a-variable situations occurs, a perhaps better solution is to use a function expression:

```

var factorial = (function hideTheCache() {
    var cache = {};

```

```

function factorial(x) {
  if (x < 2) return 1;
  if (!(x in cache)) {
    cache[x] = x * factorial(x - 1);
  }
  return cache[x];
}

return factorial;
})();

factorial(6);
// 720

factorial(7);
// 5040

```

Wait! This is still using a function to create the scope for hiding `cache`, and in this case, the function is still named `hideTheCache`, so how does that solve anything?

Recall from “Function Name Scope” (in Chapter 3), what happens to the name identifier from a `function` expression. Since `hideTheCache(..)` is defined as a `function` expression instead of a `function` declaration, its name is in its own scope—essentially the same scope as `cache`—rather than in the outer/global scope.

That means we can name every single occurrence of such a function expression the exact same name, and never have any collision. More appropriately, we can name each occurrence semantically based on whatever it is we’re trying to hide, and not worry that whatever name we choose is going to collide with any other `function` expression scope in the program.

In fact, we *could* just leave off the name entirely—thus defining an “anonymous `function` expression” instead. But Appendix A will discuss the importance of names even for such scope-only functions.

Invoking Function Expressions Immediately

There’s another important bit in the previous factorial recursive program that’s easy to miss: the line at the end of the `function` expression that contains `)();`.

Notice that we surrounded the entire `function` expression in a set of `(..)`, and then on the end, we added that second `()` parentheses set; that’s actually calling the `function` expression we just defined. Moreover, in this case, the first set of surrounding `(..)` around the function expression is not strictly necessary (more on that in a moment), but we used them for readability sake

anyway.

So, in other words, we're defining a `function` expression that's then immediately invoked. This common pattern has a (very creative!) name: Immediately Invoked Function Expression (IIFE).

An IIFE is useful when we want to create a scope to hide variables/functions. Since it's an expression, it can be used in **any** place in a JS program where an expression is allowed. An IIFE can be named, as with `hideTheCache()`, or (much more commonly!) unnamed/anonymous. And it can be standalone or, as before, part of another statement—`hideTheCache()` returns the `factorial()` function reference which is then = assigned to the variable `factorial`.

For comparison, here's an example of a standalone IIFE:

```
// outer scope

(function(){
  // inner hidden scope
})();
```



```
// more outer scope
```

Unlike earlier with `hideTheCache()`, where the outer surrounding (...) were noted as being an optional stylistic choice, for a standalone IIFE they're **required**; they distinguish the `function` as an expression, not a statement. For consistency, however, always surround an IIFE `function` with (...).

NOTE:

Technically, the surrounding (...) aren't the only syntactic way to ensure the `function` in an IIFE is treated by the JS parser as a function expression. We'll look at some other options in Appendix A.

Function Boundaries Beware that using an IIFE to define a scope can have some unintended consequences, depending on the code around it. Because an IIFE is a full function, the function boundary alters the behavior of certain statements/constructs.

For example, a `return` statement in some piece of code would change its meaning if an IIFE is wrapped around it, because now the `return` would refer to the IIFE's function. Non-arrow function IIFEs also change the binding of a `this` keyword—more on that in the *Objects & Classes* book. And statements like `break` and `continue` won't operate across an IIFE function boundary to control an outer loop or block.

So, if the code you need to wrap a scope around has `return`, `this`, `break`, or `continue` in it, an IIFE is probably not the best approach. In that case, you might look to create the scope with a block instead of a function.

Scoping with Blocks

You should by this point feel fairly comfortable with the merits of creating scopes to limit identifier exposure.

So far, we looked at doing this via `function` (i.e., IIFE) scope. But let's now consider using `let` declarations with nested blocks. In general, any `{ ... }` curly-brace pair which is a statement will act as a block, but **not necessarily** as a scope.

A block only becomes a scope if necessary, to contain its block-scoped declarations (i.e., `let` or `const`). Consider:

```
{  
  // not necessarily a scope (yet)  
  
  // ...  
  
  // now we know the block needs to be a scope  
  let thisIsNowAScope = true;  
  
  for (let i = 0; i < 5; i++) {  
    // this is also a scope, activated each  
    // iteration  
    if (i % 2 == 0) {  
      // this is just a block, not a scope  
      console.log(i);  
    }  
  }  
}  
// 0 2 4
```

Not all `{ ... }` curly-brace pairs create blocks (and thus are eligible to become scopes):

- Object literals use `{ ... }` curly-brace pairs to delimit their key-value lists, but such object values are **not** scopes.
- `class` uses `{ ... }` curly-braces around its body definition, but this is not a block or scope.
- A `function` uses `{ ... }` around its body, but this is not technically a block—it's a single statement for the function body. It *is*, however, a (function) scope.
- The `{ ... }` curly-brace pair on a `switch` statement (around the set of `case` clauses) does not define a block/scope.

Other than such non-block examples, a `{ ... }` curly-brace pair can define a block attached to a statement (like an `if` or `for`), or stand alone by itself—see

the outermost `{ ... }` curly brace pair in the previous snippet. An explicit block of this sort—if it has no declarations, it’s not actually a scope—serves no operational purpose, though it can still be useful as a semantic signal.

Explicit standalone `{ ... }` blocks have always been valid JS syntax, but since they couldn’t be a scope prior to ES6’s `let`/`const`, they are quite rare. However, post ES6, they’re starting to catch on a little bit.

In most languages that support block scoping, an explicit block scope is an extremely common pattern for creating a narrow slice of scope for one or a few variables. So following the POLE principle, we should embrace this pattern more widespread in JS as well; use (explicit) block scoping to narrow the exposure of identifiers to the minimum practical.

An explicit block scope can be useful even inside of another block (whether the outer block is a scope or not).

For example:

```
if (somethingHappened) {  
  // this is a block, but not a scope  
  
  {  
    // this is both a block and an  
    // explicit scope  
    let msg = somethingHappened.message();  
    notifyOthers(msg);  
  }  
  
  // ...  
  
  recoverFromSomething();  
}
```

Here, the `{ ... }` curly-brace pair **inside** the `if` statement is an even smaller inner explicit block scope for `msg`, since that variable is not needed for the entire `if` block. Most developers would just block-scope `msg` to the `if` block and move on. And to be fair, when there’s only a few lines to consider, it’s a toss-up judgement call. But as code grows, these over-exposure issues become more pronounced.

So does it matter enough to add the extra `{ ... }` pair and indentation level? I think you should follow POLE and always (within reason!) define the smallest block for each variable. So I recommend using the extra explicit block scope as shown.

Recall the discussion of TDZ errors from “Uninitialized Variables (TDZ)” (Chapter 5). My suggestion there was: to minimize the risk of TDZ errors with `let`/`const` declarations, always put those declarations at the top of their scope.

If you find yourself placing a `let` declaration in the middle of a scope, first think, “Oh, no! TDZ alert!” If this `let` declaration isn’t needed in the first half of that block, you should use an inner explicit block scope to further narrow its exposure!

Another example with an explicit block scope:

```
function getNextMonthStart(dateStr) {
  var nextMonth, year;

  {
    let curMonth;
    [ , year, curMonth ] = dateStr.match(
      /(\d{4})-(\d{2})-(\d{2})/
    ) || [];
    nextMonth = (Number(curMonth) % 12) + 1;
  }

  if (nextMonth == 1) {
    year++;
  }

  return `${year}-${String(nextMonth).padStart(2, "0")}-01`;
}

getNextMonthStart("2019-12-25"); // 2020-01-01
```

Let’s first identify the scopes and their identifiers:

1. The outer/global scope has one identifier, the function `getNextMonthStart(...)`.
2. The function scope for `getNextMonthStart(...)` has three: `dateStr` (parameter), `nextMonth`, and `year`.
3. The `{ .. }` curly-brace pair defines an inner block scope that includes one variable: `curMonth`.

So why put `curMonth` in an explicit block scope instead of just alongside `nextMonth` and `year` in the top-level function scope? Because `curMonth` is only needed for those first two statements; at the function scope level it’s over-exposed.

This example is small, so the hazards of over-exposing `curMonth` are pretty limited. But the benefits of the POLE principle are best achieved when you adopt the mindset of minimizing scope exposure by default, as a habit. If you follow the principle consistently even in the small cases, it will serve you more as your programs grow.

Let’s now look at an even more substantial example:

```

function sortNamesByLength(names) {
  var buckets = [];

  for (let firstName of names) {
    if (buckets[firstName.length] == null) {
      buckets[firstName.length] = [];
    }
    buckets[firstName.length].push(firstName);
  }

  // a block to narrow the scope
  {
    let sortedNames = [];

    for (let bucket of buckets) {
      if (bucket) {
        // sort each bucket alphanumerically
        bucket.sort();

        // append the sorted names to our
        // running list
        sortedNames = [
          ...sortedNames,
          ...bucket
        ];
      }
    }
  }

  return sortedNames;
}
}

sortNamesByLength([
  "Sally",
  "Suzy",
  "Frank",
  "John",
  "Jennifer",
  "Scott"
]);
// [ "John", "Suzy", "Frank", "Sally",
//   "Scott", "Jennifer" ]

```

There are six identifiers declared across five different scopes. Could all of these variables have existed in the single outer/global scope? Technically, yes, since they're all uniquely named and thus have no name collisions. But this would

be really poor code organization, and would likely lead to both confusion and future bugs.

We split them out into each inner nested scope as appropriate. Each variable is defined at the innermost scope possible for the program to operate as desired.

`sortedNames` could have been defined in the top-level function scope, but it's only needed for the second half of this function. To avoid over-exposing that variable in a higher level scope, we again follow POLE and block-scope it in the inner explicit block scope.

`var` and `let`

Next, let's talk about the declaration `var buckets`. That variable is used across the entire function (except the final `return` statement). Any variable that is needed across all (or even most) of a function should be declared so that such usage is obvious.

NOTE:

The parameter `names` isn't used across the whole function, but there's no way limit the scope of a parameter, so it behaves as a function-wide declaration regardless.

So why did we use `var` instead of `let` to declare the `buckets` variable? There's both semantic and technical reasons to choose `var` here.

Stylistically, `var` has always, from the earliest days of JS, signaled “variable that belongs to a whole function.” As we asserted in “Lexical Scope” (Chapter 1), `var` attaches to the nearest enclosing function scope, no matter where it appears. That's true even if `var` appears inside a block:

```
function diff(x,y) {
  if (x > y) {
    var tmp = x;    // `tmp` is function-scoped
    x = y;
    y = tmp;
  }

  return y - x;
}
```

Even though `var` is inside a block, its declaration is function-scoped (to `diff(..)`), not block-scoped.

While you can declare `var` inside a block (and still have it be function-scoped), I would recommend against this approach except in a few specific cases (discussed in Appendix A). Otherwise, `var` should be reserved for use in the top-level scope of a function.

Why not just use `let` in that same location? Because `var` is visually distinct from `let` and therefore signals clearly, “this variable is function-scoped.” Using `let` in the top-level scope, especially if not in the first few lines of a function, and when all the other declarations in blocks use `let`, does not visually draw attention to the difference with the function-scoped declaration.

In other words, I feel `var` better communicates function-scoped than `let` does, and `let` both communicates (and achieves!) block-scoping where `var` is insufficient. As long as your programs are going to need both function-scoped and block-scoped variables, the most sensible and readable approach is to use both `var` and `let` together, each for their own best purpose.

There are other semantic and operational reasons to choose `var` or `let` in different scenarios. We’ll explore the case for `var` and `let` in more detail in Appendix A.

WARNING:

My recommendation to use both `var` and `let` is clearly controversial and contradicts the majority. It’s far more common to hear assertions like, “var is broken, let fixes it” and, “never use var, let is the replacement.” Those opinions are valid, but they’re merely opinions, just like mine. `var` is not factually broken or deprecated; it has worked since early JS and it will continue to work as long as JS is around.

Where To `let`?

My advice to reserve `var` for (mostly) only a top-level function scope means that most other declarations should use `let`. But you may still be wondering how to decide where each declaration in your program belongs?

POLE already guides you on those decisions, but let’s make sure we explicitly state it. The way to decide is not based on which keyword you want to use. The way to decide is to ask, “What is the most minimal scope exposure that’s sufficient for this variable?”

Once that is answered, you’ll know if a variable belongs in a block scope or the function scope. If you decide initially that a variable should be block-scoped, and later realize it needs to be elevated to be function-scoped, then that dictates a change not only in the location of that variable’s declaration, but also the declarator keyword used. The decision-making process really should proceed like that.

If a declaration belongs in a block scope, use `let`. If it belongs in the function scope, use `var` (again, just my opinion).

But another way to sort of visualize this decision making is to consider the pre-ES6 version of a program. For example, let’s recall `diff(..)` from earlier:

```

function diff(x,y) {
  var tmp;

  if (x > y) {
    tmp = x;
    x = y;
    y = tmp;
  }

  return y - x;
}

```

In this version of `diff(..)`, `tmp` is clearly declared in the function scope. Is that appropriate for `tmp`? I would argue, no. `tmp` is only needed for those few statements. It's not needed for the `return` statement. It should therefore be block-scoped.

Prior to ES6, we didn't have `let` so we couldn't *actually* block-scope it. But we could do the next-best thing in signaling our intent:

```

function diff(x,y) {
  if (x > y) {
    // `tmp` is still function-scoped, but
    // the placement here semantically
    // signals block-scoping
    var tmp = x;
    x = y;
    y = tmp;
  }

  return y - x;
}

```

Placing the `var` declaration for `tmp` inside the `if` statement signals to the reader of the code that `tmp` belongs to that block. Even though JS doesn't enforce that scoping, the semantic signal still has benefit for the reader of your code.

Following this perspective, you can find any `var` that's inside a block of this sort and switch it to `let` to enforce the semantic signal already being sent. That's proper usage of `let` in my opinion.

Another example that was historically based on `var` but which should now pretty much always use `let` is the `for` loop:

```

for (var i = 0; i < 5; i++) {
  // do something
}

```

No matter where such a loop is defined, the `i` should basically always be used only inside the loop, in which case POLE dictates it should be declared with

`let` instead of `var`:

```
for (let i = 0; i < 5; i++) {  
    // do something  
}
```

Almost the only case where switching a `var` to a `let` in this way would “break” your code is if you were relying on accessing the loop’s iterator (`i`) outside/after the loop, such as:

```
for (var i = 0; i < 5; i++) {  
    if (checkValue(i)) {  
        break;  
    }  
}  
  
if (i < 5) {  
    console.log("The loop stopped early!");  
}
```

This usage pattern is not terribly uncommon, but most feel it smells like poor code structure. A preferable approach is to use another outer-scoped variable for that purpose:

```
var lastI;  
  
for (let i = 0; i < 5; i++) {  
    lastI = i;  
    if (checkValue(i)) {  
        break;  
    }  
}  
  
if (lastI < 5) {  
    console.log("The loop stopped early!");  
}
```

`lastI` is needed across this whole scope, so it’s declared with `var`. `i` is only needed in (each) loop iteration, so it’s declared with `let`.

What’s the Catch?

So far we’ve asserted that `var` and parameters are function-scoped, and `let/const` signal block-scoped declarations. There’s one little exception to call out: the `catch` clause.

Since the introduction of `try..catch` back in ES3 (in 1999), the `catch` clause has used an additional (little-known) block-scoping declaration capability:

```

try {
  doesntExist();
}

catch (err) {
  console.log(err);
  // ReferenceError: 'doesntExist' is not defined
  // ^^^ message printed from the caught exception

  let onlyHere = true;
  var outerVariable = true;
}

console.log(outerVariable);      // true

console.log(err);
// ReferenceError: 'err' is not defined
// ^^^ this is another thrown (uncaught) exception

```

The `err` variable declared by the `catch` clause is block-scoped to that block. This `catch` clause block can hold other block-scoped declarations via `let`. But a `var` declaration inside this block still attaches to the outer function/global scope.

ES2019 (recently, at the time of writing) changed `catch` clauses so their declaration is optional; if the declaration is omitted, the `catch` block is no longer (by default) a scope; it's still a block, though!

So if you need to react to the condition *that an exception occurred* (so you can gracefully recover), but you don't care about the error value itself, you can omit the `catch` declaration:

```

try {
  doOptionOne();
}

catch { // catch-declaration omitted
  doOptionTwoInstead();
}

```

This is a small but delightful simplification of syntax for a fairly common use case, and may also be slightly more performant in removing an unnecessary scope!

Function Declarations in Blocks (FiB)

We've seen now that declarations using `let` or `const` are block-scoped, and `var` declarations are function-scoped. So what about `function` declarations that appear directly inside blocks? As a feature, this is called "FiB."

We typically think of `function` declarations like they're the equivalent of a `var` declaration. So are they function-scoped like `var` is?

No and yes. I know... that's confusing. Let's dig in:

```
if (false) {
  function ask() {
    console.log("Does this run?");
  }
}
ask();
```

What do you expect for this program to do? Three reasonable outcomes:

1. The `ask()` call might fail with a `ReferenceError` exception, because the `ask` identifier is block-scoped to the `if` block scope and thus isn't available in the outer/global scope.
2. The `ask()` call might fail with a `TypeError` exception, because the `ask` identifier exists, but it's `undefined` (since the `if` statement doesn't run) and thus not a callable function.
3. The `ask()` call might run correctly, printing out the "Does it run?" message.

Here's the confusing part: depending on which JS environment you try that code snippet in, you may get different results! This is one of those few crazy areas where existing legacy behavior betrays a predictable outcome.

The JS specification says that `function` declarations inside of blocks are block-scoped, so the answer should be (1). However, most browser-based JS engines (including v8, which comes from Chrome but is also used in Node) will behave as (2), meaning the identifier is scoped outside the `if` block but the function value is not automatically initialized, so it remains `undefined`.

Why are browser JS engines allowed to behave contrary to the specification? Because these engines already had certain behaviors around FiB before ES6 introduced block scoping, and there was concern that changing to adhere to the specification might break some existing website JS code. As such, an exception was made in Appendix B of the JS specification, which allows certain deviations for browser JS engines (only!).

NOTE:

You wouldn't typically categorize Node as a browser JS environment, since it usually runs on a server. But Node's v8 engine is shared with Chrome (and Edge) browsers. Since v8 is first a browser JS engine, it adopts this Appendix B exception, which then means that the browser exceptions are extended to Node.

One of the most common use cases for placing a `function` declaration in a block is to conditionally define a function one way or another (like with an `if..else` statement) depending on some environment state. For example:

```
if (typeof Array.isArray != "undefined") {
  function isArray(a) {
    return Array.isArray(a);
  }
} else {
  function isArray(a) {
    return Object.prototype.toString.call(a)
      == "[object Array]";
  }
}
```

It's tempting to structure code this way for performance reasons, since the `typeof Array.isArray` check is only performed once, as opposed to defining just one `isArray(..)` and putting the `if` statement inside it—the check would then run unnecessarily on every call.

WARNING:

In addition to the risks of FiB deviations, another problem with conditional-definition of functions is it's harder to debug such a program. If you end up with a bug in the `isArray(..)` function, you first have to figure out *which* `isArray(..)` implementation is actually running! Sometimes, the bug is that the wrong one was applied because the conditional check was incorrect! If you define multiple versions of a function, that program is always harder to reason about and maintain.

In addition to the previous snippets, several other FiB corner cases are lurking; such behaviors in various browsers and non-browser JS environments (JS engines that aren't browser based) will likely vary. For example:

```
if (true) {
  function ask() {
    console.log("Am I called?");
  }
}

if (true) {
  function ask() {
    console.log("Or what about me?");
  }
}
```

```

for (let i = 0; i < 5; i++) {
  function ask() {
    console.log("Or is it one of these?");
  }
}

ask();

function ask() {
  console.log("Wait, maybe, it's this one?");
}

```

Recall that function hoisting as described in “When Can I Use a Variable?” (in Chapter 5) might suggest that the final `ask()` in this snippet, with “Wait, maybe...” as its message, would hoist above the call to `ask()`. Since it’s the last function declaration of that name, it should “win,” right? Unfortunately, no.

It’s not my intention to document all these weird corner cases, nor to try to explain why each of them behaves a certain way. That information is, in my opinion, arcane legacy trivia.

My real concern with FiB is, what advice can I give to ensure your code behaves predictably in all circumstances?

As far as I’m concerned, the only practical answer to avoiding the vagaries of FiB is to simply avoid FiB entirely. In other words, never place a `function` declaration directly inside any block. Always place `function` declarations anywhere in the top-level scope of a function (or in the global scope).

So for the earlier `if..else` example, my suggestion is to avoid conditionally defining functions if at all possible. Yes, it may be slightly less performant, but this is the better overall approach:

```

function isArray(a) {
  if (typeof Array.isArray != "undefined") {
    return Array.isArray(a);
  }
  else {
    return Object.prototype.toString.call(a)
      == "[object Array]";
  }
}

```

If that performance hit becomes a critical path issue for your application, I suggest you consider this approach:

```

var isArray = function isArray(a) {
  return Array.isArray(a);
};

```

```
// override the definition, if you must
if (typeof Array.isArray == "undefined") {
  isArray = function isArray(a) {
    return Object.prototype.toString.call(a)
      == "[object Array]";
  };
}
```

It's important to notice that here I'm placing a **function expression**, not a declaration, inside the **if** statement. That's perfectly fine and valid, for **function expressions** to appear inside blocks. Our discussion about FiB is about avoiding **function declarations** in blocks.

Even if you test your program and it works correctly, the small benefit you may derive from using FiB style in your code is far outweighed by the potential risks in the future for confusion by other developers, or variances in how your code runs in other JS environments.

FiB is not worth it, and should be avoided.

Blocked Over

The point of lexical scoping rules in a programming language is so we can appropriately organize our program's variables, both for operational as well as semantic code communication purposes.

And one of the most important organizational techniques is to ensure that no variable is over-exposed to unnecessary scopes (POLE). Hopefully you now appreciate block scoping much more deeply than before.

Hopefully by now you feel like you're standing on much more solid ground with understanding lexical scope. From that base, the next chapter jumps into the weighty topic of closure.

You Don't Know JS Yet: Scope & Closures - 2nd Edition

Chapter 7: Using Closures

Up to this point, we've focused on the ins and outs of lexical scope, and how that affects the organization and usage of variables in our programs.

Our attention again shifts broader in abstraction, to the historically somewhat daunting topic of closure. Don't worry! You don't need an advanced computer science degree to make sense of it. Our broad goal in this book is not merely to understand scope, but to more effectively use it in the structure of our programs; closure is central to that effort.

Recall the main conclusion of Chapter 6: the *least exposure* principle (POLE) encourages us to use block (and function) scoping to limit the scope exposure of variables. This helps keep code understandable and maintainable, and helps avoid many scoping pitfalls (i.e., name collision, etc.).

Closure builds on this approach: for variables we need to use over time, instead of placing them in larger outer scopes, we can encapsulate (more narrowly scope) them but still preserve access from inside functions, for broader use. Functions *remember* these referenced scoped variables via closure.

We already saw an example of this kind of closure in the previous chapter (`factorial(..)` in Chapter 6), and you've almost certainly already used it in your own programs. If you've ever written a callback that accesses variables outside its own scope... guess what!? That's closure.

Closure is one of the most important language characteristics ever invented in programming—it underlies major programming paradigms, including Functional Programming (FP), modules, and even a bit of class-oriented design. Getting comfortable with closure is required for mastering JS and effectively leveraging many important design patterns throughout your code.

Addressing all aspects of closure requires a daunting mountain of discussion and code throughout this chapter. Make sure to take your time and ensure you're comfortable with each bit before moving onto the next.

See the Closure

Closure is originally a mathematical concept, from lambda calculus. But I'm not going to list out math formulas or use a bunch of notation and jargon to define it.

Instead, I'm going to focus on a practical perspective. We'll start by defining closure in terms of what we can observe in different behavior of our programs, as opposed to if closure was not present in JS. However, later in this chapter, we're going to flip closure around to look at it from an *alternative perspective*.

Closure is a behavior of functions and only functions. If you aren't dealing with a function, closure does not apply. An object cannot have closure, nor does a class have closure (though its functions/methods might). Only functions have closure.

For closure to be observed, a function must be invoked, and specifically it must be invoked in a different branch of the scope chain from where it was originally defined. A function executing in the same scope it was defined would not exhibit any observably different behavior with or without closure being possible; by the observational perspective and definition, that is not closure.

Let's look at some code, annotated with its relevant scope bubble colors (see Chapter 2):

```
// outer/global scope: RED(1)

function lookupStudent(studentID) {
  // function scope: BLUE(2)

  var students = [
    { id: 14, name: "Kyle" },
    { id: 73, name: "Suzy" },
    { id: 112, name: "Frank" },
    { id: 6, name: "Sarah" }
  ];

  return function greetStudent(greeting){
    // function scope: GREEN(3)

    var student = students.find(
      student => student.id == studentID
    );

    return `${greeting}, ${student.name}!`;
  };
}

var chosenStudents = [
  lookupStudent(6),
  lookupStudent(112)
];

// accessing the function's name:
chosenStudents[0].name;
// greetStudent

chosenStudents[0]("Hello");
```

```
// Hello, Sarah!  
  
chosenStudents[1]("Howdy");  
// Howdy, Frank!
```

The first thing to notice about this code is that the `lookupStudent(..)` outer function creates and returns an inner function called `greetStudent(..)`. `lookupStudent(..)` is called twice, producing two separate instances of its inner `greetStudent(..)` function, both of which are saved into the `chosenStudents` array.

We verify that's the case by checking the `.name` property of the returned function saved in `chosenStudents[0]`, and it's indeed an instance of the inner `greetStudent(..)`.

After each call to `lookupStudent(..)` finishes, it would seem like all its inner variables would be discarded and GC'd (garbage collected). The inner function is the only thing that seems to be returned and preserved. But here's where the behavior differs in ways we can start to observe.

While `greetStudent(..)` does receive a single argument as the parameter named `greeting`, it also makes reference to both `students` and `studentID`, identifiers which come from the enclosing scope of `lookupStudent(..)`. Each of those references from the inner function to the variable in an outer scope is called a *closure*. In academic terms, each instance of `greetStudent(..)` *closes over* the outer variables `students` and `studentID`.

So what do those closures do here, in a concrete, observable sense?

Closure allows `greetStudent(..)` to continue to access those outer variables even after the outer scope is finished (when each call to `lookupStudent(..)` completes). Instead of the instances of `students` and `studentID` being GC'd, they stay around in memory. At a later time when either instance of the `greetStudent(..)` function is invoked, those variables are still there, holding their current values.

If JS functions did not have closure, the completion of each `lookupStudent(..)` call would immediately tear down its scope and GC the `students` and `studentID` variables. When we later called one of the `greetStudent(..)` functions, what would then happen?

If `greetStudent(..)` tried to access what it thought was a `BLUE(2)` marble, but that marble did not actually exist (anymore), the reasonable assumption is we should get a `ReferenceError`, right?

But we don't get an error. The fact that the execution of `chosenStudents[0]("Hello")` works and returns us the message "Hello, Sarah!", means it was still able to access the `students` and `studentID` variables. This is a direct observation of closure!

Pointed Closure

Actually, we glossed over a little detail in the previous discussion which I'm guessing many readers missed!

Because of how terse the syntax for `=>` arrow functions is, it's easy to forget that they still create a scope (as asserted in "Arrow Functions" in Chapter 3). The `student => student.id == studentID` arrow function is creating another scope bubble inside the `greetStudent(..)` function scope.

Building on the metaphor of colored buckets and bubbles from Chapter 2, if we were creating a colored diagram for this code, there's a fourth scope at this innermost nesting level, so we'd need a fourth color; perhaps we'd pick ORANGE(4) for that scope:

```
var student = students.find(
  student =>
    // function scope: ORANGE(4)
    student.id == studentID
);
```

The BLUE(2) `studentID` reference is actually inside the ORANGE(4) scope rather than the GREEN(3) scope of `greetStudent(..)`; also, the `student` parameter of the arrow function is ORANGE(4), shadowing the GREEN(3) `student`.

The consequence here is that this arrow function passed as a callback to the array's `find(..)` method has to hold the closure over `studentID`, rather than `greetStudent(..)` holding that closure. That's not too big of a deal, as everything still works as expected. It's just important not to skip over the fact that even tiny arrow functions can get in on the closure party.

Adding Up Closures

Let's examine one of the canonical examples often cited for closure:

```
function adder(num1) {
  return function addTo(num2){
    return num1 + num2;
  };
}

var add10To = adder(10);
var add42To = adder(42);

add10To(15);    // 25
add42To(9);    // 51
```

Each instance of the inner `addTo(..)` function is closing over its own `num1` variable (with values 10 and 42, respectively), so those `num1`'s don't go away just

because `adder(..)` finishes. When we later invoke one of those inner `addTo(..)` instances, such as the `add10To(15)` call, its closed-over `num1` variable still exists and still holds the original 10 value. The operation is thus able to perform $10 + 15$ and return the answer 25.

An important detail might have been too easy to gloss over in that previous paragraph, so let's reinforce it: closure is associated with an instance of a function, rather than its single lexical definition. In the preceding snippet, there's just one inner `addTo(..)` function defined inside `adder(..)`, so it might seem like that would imply a single closure.

But actually, every time the outer `adder(..)` function runs, a *new* inner `addTo(..)` function instance is created, and for each new instance, a new closure. So each inner function instance (labeled `add10To(..)` and `add42To(..)` in our program) has its own closure over its own instance of the scope environment from that execution of `adder(..)`.

Even though closure is based on lexical scope, which is handled at compile time, closure is observed as a runtime characteristic of function instances.

Live Link, Not a Snapshot

In both examples from the previous sections, we **read the value from a variable** that was held in a closure. That makes it feel like closure might be a snapshot of a value at some given moment. Indeed, that's a common misconception.

Closure is actually a live link, preserving access to the full variable itself. We're not limited to merely reading a value; the closed-over variable can be updated (re-assigned) as well! By closing over a variable in a function, we can keep using that variable (read and write) as long as that function reference exists in the program, and from anywhere we want to invoke that function. This is why closure is such a powerful technique used widely across so many areas of programming!

Figure 4 depicts the function instances and scope links:

Fig. 4: Visualizing Closures

As shown in Figure 4, each call to `adder(..)` creates a new BLUE(2) scope containing a `num1` variable, as well as a new instance of `addTo(..)` function as a GREEN(3) scope. Notice that the function instances (`add10To(..)` and `add42To(..)`) are present in and invoked from the RED(1) scope.

Now let's examine an example where the closed-over variable is updated:

```
function makeCounter() {  
  var count = 0;  
  
  return function getCurrent() {
```

```

        count = count + 1;
        return count;
    };
}

var hits = makeCounter();

// later

hits();      // 1

// later

hits();      // 2
hits();      // 3

```

The `count` variable is closed over by the inner `getCurrent()` function, which keeps it around instead of it being subjected to GC. The `hits()` function calls access *and* update this variable, returning an incrementing count each time.

Though the enclosing scope of a closure is typically from a function, that's not actually required; there only needs to be an inner function present inside an outer scope:

```

var hits;
{  // an outer scope (but not a function)
    let count = 0;
    hits = function getCurrent(){
        count = count + 1;
        return count;
    };
}
hits();      // 1
hits();      // 2
hits();      // 3

```

NOTE:

I deliberately defined `getCurrent()` as a `function` expression instead of a `function` declaration. This isn't about closure, but with the dangerous quirks of FiB (Chapter 6).

Because it's so common to mistake closure as value-oriented instead of variable-oriented, developers sometimes get tripped up trying to use closure to snapshot-preserve a value from some moment in time. Consider:

```
var studentName = "Frank";
```

```

var greeting = function hello() {
  // we are closing over `studentName`,
  // not "Frank"
  console.log(
    `Hello, ${studentName}!`
  );
}

// later

studentName = "Suzy";

// later

greeting();
// Hello, Suzy!

```

By defining `greeting()` (aka, `hello()`) when `studentName` holds the value `"Frank"` (before the re-assignment to `"Suzy"`), the mistaken assumption is often that the closure will capture `"Frank"`. But `greeting()` is closed over the variable `studentName`, not its value. Whenever `greeting()` is invoked, the current value of the variable (`"Suzy"`, in this case) is reflected.

The classic illustration of this mistake is defining functions inside a loop:

```

var keeps = [];

for (var i = 0; i < 3; i++) {
  keeps[i] = function keepI(){
    // closure over `i`
    return i;
}
}

keeps[0](); // 3 -- WHY!?
keeps[1](); // 3
keeps[2](); // 3

```

NOTE:

This kind of closure illustration typically uses a `setTimeout(..)` or some other callback like an event handler, inside the loop. I've simplified the example by storing function references in an array, so that we don't need to consider asynchronous timing in our analysis. The closure principle is the same, regardless.

You might have expected the `keeps[0]()` invocation to return `0`, since that

function was created during the first iteration of the loop when `i` was 0. But again, that assumption stems from thinking of closure as value-oriented rather than variable-oriented.

Something about the structure of a `for`-loop can trick us into thinking that each iteration gets its own new `i` variable; in fact, this program only has one `i` since it was declared with `var`.

Each saved function returns 3, because by the end of the loop, the single `i` variable in the program has been assigned 3. Each of the three functions in the `keeps` array do have individual closures, but they're all closed over that same shared `i` variable.

Of course, a single variable can only ever hold one value at any given moment. So if you want to preserve multiple values, you need a different variable for each.

How could we do that in the loop snippet? Let's create a new variable for each iteration:

```
var keeps = [];

for (var i = 0; i < 3; i++) {
  // new `j` created each iteration, which gets
  // a copy of the value of `i` at this moment
  let j = i;

  // the `i` here isn't being closed over, so
  // it's fine to immediately use its current
  // value in each loop iteration
  keeps[i] = function keepEachJ(){
    // close over `j`, not `i`!
    return j;
  };
}
keeps[0](); // 0
keeps[1](); // 1
keeps[2](); // 2
```

Each function is now closed over a separate (new) variable from each iteration, even though all of them are named `j`. And each `j` gets a copy of the value of `i` at that point in the loop iteration; that `j` never gets re-assigned. So all three functions now return their expected values: 0, 1, and 2!

Again remember, even if we were using asynchrony in this program, such as passing each inner `keepEachJ()` function into `setTimeout(..)` or some event handler subscription, the same kind of closure behavior would still be observed.

Recall the “Loops” section in Chapter 5, which illustrates how a `let` declaration in a `for` loop actually creates not just one variable for the loop, but actually

creates a new variable for *each iteration* of the loop. That trick/quirk is exactly what we need for our loop closures:

```
var keeps = [];  
  
for (let i = 0; i < 3; i++) {  
    // the `let i` gives us a new `i` for  
    // each iteration, automatically!  
    keeps[i] = function keepEachI(){  
        return i;  
    };  
}  
keeps[0](); // 0  
keeps[1](); // 1  
keeps[2](); // 2
```

Since we're using `let`, three `i`'s are created, one for each loop, so each of the three closures *just work* as expected.

Common Closures: Ajax and Events

Closure is most commonly encountered with callbacks:

```
function lookupStudentRecord(studentID) {  
    ajax(  
        `https://some.api/student/${ studentID }`,  
        function onRecord(record) {  
            console.log(  
                `${ record.name } (${ studentID })`  
            );  
        }  
    );  
}  
  
lookupStudentRecord(114);  
// Frank (114)
```

The `onRecord(..)` callback is going to be invoked at some point in the future, after the response from the Ajax call comes back. This invocation will happen from the internals of the `ajax(..)` utility, wherever that comes from. Furthermore, when that happens, the `lookupStudentRecord(..)` call will long since have completed.

Why then is `studentID` still around and accessible to the callback? Closure.

Event handlers are another common usage of closure:

```
function listenForClicks(btn,label) {  
    btn.addEventListener("click",function onClick(){
```

```

        console.log(
            `The ${ label } button was clicked!`
        );
    });
}

var submitBtn = document.getElementById("submit-btn");

listenForClicks(submitBtn, "Checkout");

```

The `label` parameter is closed over by the `onClick(..)` event handler callback. When the button is clicked, `label` still exists to be used. This is closure.

What If I Can't See It?

You've probably heard this common adage:

If a tree falls in the forest but nobody is around to hear it, does it make a sound?

It's a silly bit of philosophical gymnastics. Of course from a scientific perspective, sound waves are created. But the real point: *does it matter* if the sound happens?

Remember, the emphasis in our definition of closure is observability. If a closure exists (in a technical, implementation, or academic sense) but it cannot be observed in our programs, *does it matter*? No.

To reinforce this point, let's look at some examples that are *not* observably based on closure.

For example, invoking a function that makes use of lexical scope lookup:

```

function say(myName) {
    var greeting = "Hello";
    output();

    function output() {
        console.log(
            `${ greeting }, ${ myName }!`
        );
    }
}

say("Kyle");
// Hello, Kyle!

```

The inner function `output()` accesses the variables `greeting` and `myName` from its enclosing scope. But the invocation of `output()` happens in that same scope,

where of course `greeting` and `myName` are still available; that's just lexical scope, not closure.

Any lexically scoped language whose functions didn't support closure would still behave this same way.

In fact, global scope variables essentially cannot be (observably) closed over, because they're always accessible from everywhere. No function can ever be invoked in any part of the scope chain that is not a descendant of the global scope.

Consider:

```
var students = [
  { id: 14, name: "Kyle" },
  { id: 73, name: "Suzy" },
  { id: 112, name: "Frank" },
  { id: 6, name: "Sarah" }
];

function getFirstStudent() {
  return function firstStudent(){
    return students[0].name;
  };
}

var student = getFirstStudent();

student();
// Kyle
```

The inner `firstStudent()` function does reference `students`, which is a variable outside its own scope. But since `students` happens to be from the global scope, no matter where that function is invoked in the program, its ability to access `students` is nothing more special than normal lexical scope.

All function invocations can access global variables, regardless of whether closure is supported by the language or not. Global variables don't need to be closed over.

Variables that are merely present but never accessed don't result in closure:

```
function lookupStudent(studentID) {
  return function nobody(){
    var msg = "Nobody's here yet.";
    console.log(msg);
  };
}

var student = lookupStudent(112);
```

```
student();
// Nobody's here yet.
```

The inner function `nobody()` doesn't close over any outer variables—it only uses its own variable `msg`. Even though `studentID` is present in the enclosing scope, `studentID` is not referred to by `nobody()`. The JS engine doesn't need to keep `studentID` around after `lookupStudent(..)` has finished running, so GC wants to clean up that memory!

Whether JS functions support closure or not, this program would behave the same. Therefore, no observed closure here.

If there's no function invocation, closure can't be observed:

```
function greetStudent(studentName) {
  return function greeting(){
    console.log(
      `Hello, ${ studentName }!`
    );
  };
}

greetStudent("Kyle");

// nothing else happens
```

This one's tricky, because the outer function definitely does get invoked. But the inner function is the one that *could* have had closure, and yet it's never invoked; the returned function here is just thrown away. So even if technically the JS engine created closure for a brief moment, it was not observed in any meaningful way in this program.

A tree may have fallen... but we didn't hear it, so we don't care.

Observable Definition

We're now ready to define closure:

Closure is observed when a function uses variable(s) from outer scope(s) even while running in a scope where those variable(s) wouldn't be accessible.

The key parts of this definition are:

- Must be a function involved
- Must reference at least one variable from an outer scope
- Must be invoked in a different branch of the scope chain from the variable(s)

This observation-oriented definition means we shouldn't dismiss closure as some indirect, academic trivia. Instead, we should look and plan for the direct, concrete effects closure has on our program behavior.

The Closure Lifecycle and Garbage Collection (GC)

Since closure is inherently tied to a function instance, its closure over a variable lasts as long as there is still a reference to that function.

If ten functions all close over the same variable, and over time nine of these function references are discarded, the lone remaining function reference still preserves that variable. Once that final function reference is discarded, the last closure over that variable is gone, and the variable itself is GC'd.

This has an important impact on building efficient and performant programs. Closure can unexpectedly prevent the GC of a variable that you're otherwise done with, which leads to run-away memory usage over time. That's why it's important to discard function references (and thus their closures) when they're not needed anymore.

Consider:

```
function manageBtnClickEvents(btn) {
  var clickHandlers = [];

  return function listener(cb){
    if (cb) {
      let clickHandler =
        function onClick(evt){
          console.log("clicked!");
          cb(evt);
        };
      clickHandlers.push(clickHandler);
      btn.addEventListener(
        "click",
        clickHandler
      );
    }
    else {
      // passing no callback unsubscribes
      // all click handlers
      for (let handler of clickHandlers) {
        btn.removeEventListener(
          "click",
          handler
        );
      }
    }
  }
}
```

```

        clickHandlers = [];
    }
}
};

// var mySubmitBtn = ..
var onSubmit = manageBtnClickEvents(mySubmitBtn);

onSubmit(function checkout(evt){
    // handle checkout
});

onSubmit(function trackAction(evt){
    // log action to analytics
});

// later, unsubscribe all handlers:
onSubmit();

```

In this program, the inner `onClick(..)` function holds a closure over the passed in `cb` (the provided event callback). That means the `checkout()` and `trackAction()` function expression references are held via closure (and cannot be GC'd) for as long as these event handlers are subscribed.

When we call `onSubmit()` with no input on the last line, all event handlers are unsubscribed, and the `clickHandlers` array is emptied. Once all click handler function references are discarded, the closures of `cb` references to `checkout()` and `trackAction()` are discarded.

When considering the overall health and efficiency of the program, unsubscribing an event handler when it's no longer needed can be even more important than the initial subscription!

Per Variable or Per Scope?

Another question we need to tackle: should we think of closure as applied only to the referenced outer variable(s), or does closure preserve the entire scope chain with all its variables?

In other words, in the previous event subscription snippet, is the inner `onClick(..)` function closed over only `cb`, or is it also closed over `clickHandler`, `clickHandlers`, and `btn`?

Conceptually, closure is **per variable** rather than *per scope*. Ajax callbacks, event handlers, and all other forms of function closures are typically assumed to close over only what they explicitly reference.

But the reality is more complicated than that.

Another program to consider:

```
function manageStudentGrades(studentRecords) {
  var grades = studentRecords.map(getGrade);

  return addGrade;

  // *****

  function getGrade(record){
    return record.grade;
  }

  function sortAndTrimGradesList() {
    // sort by grades, descending
    grades.sort(function desc(g1,g2){
      return g2 - g1;
    });

    // only keep the top 10 grades
    grades = grades.slice(0,10);
  }

  function addGrade(newGrade) {
    grades.push(newGrade);
    sortAndTrimGradesList();
    return grades;
  }
}

var addNextGrade = manageStudentGrades([
  { id: 14, name: "Kyle", grade: 86 },
  { id: 73, name: "Suzy", grade: 87 },
  { id: 112, name: "Frank", grade: 75 },
  // ...many more records...
  { id: 6, name: "Sarah", grade: 91 }
]);

// later

addNextGrade(81);
addNextGrade(68);
// [ ..., ..., ... ]
```

The outer function `manageStudentGrades(..)` takes a list of student records, and returns an `addGrade(..)` function reference, which we externally label `addNextGrade(..)`. Each time we call `addNextGrade(..)` with a new grade,

we get back a current list of the top 10 grades, sorted numerically descending (see `sortAndTrimGradesList()`).

From the end of the original `manageStudentGrades(..)` call, and between the multiple `addNextGrade(..)` calls, the `grades` variable is preserved inside `addGrade(..)` via closure; that's how the running list of top grades is maintained. Remember, it's a closure over the variable `grades` itself, not the array it holds.

That's not the only closure involved, however. Can you spot other variables being closed over?

Did you spot that `addGrade(..)` references `sortAndTrimGradesList`? That means it's also closed over that identifier, which happens to hold a reference to the `sortAndTrimGradesList()` function. That second inner function has to stay around so that `addGrade(..)` can keep calling it, which also means any variables *it* closes over stick around—though, in this case, nothing extra is closed over there.

What else is closed over?

Consider the `getGrade` variable (and its function); is it closed over? It's referenced in the outer scope of `manageStudentGrades(..)` in the `.map(getGrade)` call. But it's not referenced in `addGrade(..)` or `sortAndTrimGradesList()`.

What about the (potentially) large list of student records we pass in as `studentRecords`? Is that variable closed over? If it is, the array of student records is never getting GC'd, which leads to this program holding onto a larger amount of memory than we might assume. But if we look closely again, none of the inner functions reference `studentRecords`.

According to the *per variable* definition of closure, since `getGrade` and `studentRecords` are *not* referenced by the inner functions, they're not closed over. They should be freely available for GC right after the `manageStudentGrades(..)` call completes.

Indeed, try debugging this code in a recent JS engine, like v8 in Chrome, placing a breakpoint inside the `addGrade(..)` function. You may notice that the inspector **does not** list the `studentRecords` variable. That's proof, debugging-wise anyway, that the engine does not maintain `studentRecords` via closure. Phew!

But how reliable is this observation as proof? Consider this (rather contrived!) program:

```
function storeStudentInfo(id, name, grade) {
  return function getInfo(whichValue) {
    // warning:
    // using `eval(..)` is a bad idea!
    var val = eval(whichValue);
    return val;
```

```

    };
}

var info = storeStudentInfo(73, "Suzy", 87);

info("name");
// Suzy

info("grade");
// 87

```

Notice that the inner function `getInfo(..)` is not explicitly closed over any of `id`, `name`, or `grade` variables. And yet, calls to `info(..)` seem to still be able to access the variables, albeit through use of the `eval(..)` lexical scope cheat (see Chapter 1).

So all the variables were definitely preserved via closure, despite not being explicitly referenced by the inner function. So does that disprove the *per variable* assertion in favor of *per scope*? Depends.

Many modern JS engines do apply an *optimization* that removes any variables from a closure scope that aren't explicitly referenced. However, as we see with `eval(..)`, there are situations where such an optimization cannot be applied, and the closure scope continues to contain all its original variables. In other words, closure must be *per scope*, implementation wise, and then an optional optimization trims down the scope to only what was closed over (a similar outcome as *per variable* closure).

Even as recent as a few years ago, many JS engines did not apply this optimization; it's possible your websites may still run in such browsers, especially on older or lower-end devices. That means it's possible that long-lived closures such as event handlers may be holding onto memory much longer than we would have assumed.

And the fact that it's an optional optimization in the first place, rather than a requirement of the specification, means that we shouldn't just casually over-assume its applicability.

In cases where a variable holds a large value (like an object or array) and that variable is present in a closure scope, if you don't need that value anymore and don't want that memory held, it's safer (memory usage) to manually discard the value rather than relying on closure optimization/GC.

Let's apply a *fix* to the earlier `manageStudentGrades(..)` example to ensure the potentially large array held in `studentRecords` is not caught up in a closure scope unnecessarily:

```

function manageStudentGrades(studentRecords) {
  var grades = studentRecords.map(getGrade);

```

```

    // unset `studentRecords` to prevent unwanted
    // memory retention in the closure
    studentRecords = null;

    return addGrade;
    // ...
}

```

We're not removing `studentRecords` from the closure scope; that we cannot control. We're ensuring that even if `studentRecords` remains in the closure scope, that variable is no longer referencing the potentially large array of data; the array can be GC'd.

Again, in many cases JS might automatically optimize the program to the same effect. But it's still a good habit to be careful and explicitly make sure we don't keep any significant amount of device memory tied up any longer than necessary.

As a matter of fact, we also technically don't need the function `getGrade()` anymore after the `.map(getGrade)` call completes. If profiling our application showed this was a critical area of excess memory use, we could possibly eek out a tiny bit more memory by freeing up that reference so its value isn't tied up either. That's likely unnecessary in this toy example, but this is a general technique to keep in mind if you're optimizing the memory footprint of your application.

The takeaway: it's important to know where closures appear in our programs, and what variables are included. We should manage these closures carefully so we're only holding onto what's minimally needed and not wasting memory.

An Alternative Perspective

Reviewing our working definition for closure, the assertion is that functions are “first-class values” that can be passed around the program, just like any other value. Closure is the link-association that connects that function to the scope/variables outside of itself, no matter where that function goes.

Let's recall a code example from earlier in this chapter, again with relevant scope bubble colors annotated:

```

// outer/global scope: RED(1)

function adder(num1) {
    // function scope: BLUE(2)

    return function addTo(num2){
        // function scope: GREEN(3)

        return num1 + num2;
    }
}

```

```

    };
}

var add10To = adder(10);
var add42To = adder(42);

add10To(15);    // 25
add42To(9);    // 51

```

Our current perspective suggests that wherever a function is passed and invoked, closure preserves a hidden link back to the original scope to facilitate the access to the closed-over variables. Figure 4, repeated here for convenience, illustrates this notion:

Fig. 4 (repeat): Visualizing Closures

But there's another way of thinking about closure, and more precisely the nature of functions being *passed around*, that may help deepen the mental models.

This alternative model de-emphasizes "functions as first-class values," and instead embraces how functions (like all non-primitive values) are held by reference in JS, and assigned/passed by reference-copy—see Appendix A of the *Get Started* book for more information.

Instead of thinking about the inner function instance of `addTo(..)` moving to the outer RED(1) scope via the `return` and assignment, we can envision that function instances actually just stay in place in their own scope environment, of course with their scope-chain intact.

What gets *sent* to the RED(1) scope is **just a reference** to the in-place function instance, rather than the function instance itself. Figure 5 depicts the inner function instances remaining in place, pointed to by the RED(1) `addTo10` and `addTo42` references, respectively:

Fig. 5: Visualizing Closures (Alternative)

As shown in Figure 5, each call to `adder(..)` still creates a new BLUE(2) scope containing a `num1` variable, as well as an instance of the GREEN(3) `addTo(..)` scope. But what's different from Figure 4 is, now these GREEN(3) instances remain in place, naturally nested inside of their BLUE(2) scope instances. The `addTo10` and `addTo42` references are moved to the RED(1) outer scope, not the function instances themselves.

When `addTo10(15)` is called, the `addTo(..)` function instance (still in place in its original BLUE(2) scope environment) is invoked. Since the function instance itself never moved, of course it still has natural access to its scope chain. Same with the `addTo42(9)` call—nothing special here beyond lexical scope.

So what then *is* closure, if not the *magic* that lets a function maintain a link to its original scope chain even as that function moves around in other scopes? In

this alternative model, functions stay in place and keep accessing their original scope chain just like they always could.

Closure instead describes the *magic* of **keeping alive a function instance**, along with its whole scope environment and chain, for as long as there's at least one reference to that function instance floating around in any other part of the program.

That definition of closure is less observational and a bit less familiar-sounding compared to the traditional academic perspective. But it's nonetheless still useful, because the benefit is that we simplify explanation of closure to a straightforward combination of references and in-place function instances.

The previous model (Figure 4) is not *wrong* at describing closure in JS. It's just more conceptually inspired, an academic perspective on closure. By contrast, the alternative model (Figure 5) could be described as a bit more implementation focused, how JS actually works.

Both perspectives/models are useful in understanding closure, but the reader may find one a little easier to hold than the other. Whichever you choose, the observable outcomes in our program are the same.

NOTE:

This alternative model for closure does affect whether we classify synchronous callbacks as examples of closure or not. More on this nuance in Appendix A.

Why Closure?

Now that we have a well-rounded sense of what closure is and how it works, let's explore some ways it can improve the code structure and organization of an example program.

Imagine you have a button on a page that when clicked, should retrieve and send some data via an Ajax request. Without using closure:

```
var APIendpoints = {
  studentIDs:
    "https://some.api/register-students",
  // ..
};

var data = {
  studentIDs: [ 14, 73, 112, 6 ],
  // ..
};

function makeRequest(evt) {
```

```

    var btn = evt.target;
    var recordKind = btn.dataset.kind;
    ajax(
        APIendpoints[recordKind],
        data[recordKind]
    );
}

// <button data-kind="studentIDs">
//   Register Students
// </button>
btn.addEventListener("click",makeRequest);

```

The `makeRequest(..)` utility only receives an `evt` object from a click event. From there, it has to retrieve the `data-kind` attribute from the target button element, and use that value to lookup both a URL for the API endpoint as well as what data should be included in the Ajax request.

This works OK, but it's unfortunate (inefficient, more confusing) that the event handler has to read a DOM attribute each time it's fired. Why couldn't an event handler *remember* this value? Let's try using closure to improve the code:

```

var APIendpoints = {
    studentIDs:
        "https://some.api/register-students",
    // ..
};

var data = {
    studentIDs: [ 14, 73, 112, 6 ],
    // ..
};

function setupButtonHandler(btn) {
    var recordKind = btn.dataset.kind;

    btn.addEventListener(
        "click",
        function makeRequest(evt){
            ajax(
                APIendpoints[recordKind],
                data[recordKind]
            );
        }
    );
}

```

```

// <button data-kind="studentIDs">
//   Register Students
// </button>

setupButtonHandler(btn);

```

With the `setupButtonHandler(..)` approach, the `data-kind` attribute is retrieved once and assigned to the `recordKind` variable at initial setup. `recordKind` is then closed over by the inner `makeRequest(..)` click handler, and its value is used on each event firing to look up the URL and data that should be sent.

NOTE:

`evt` is still passed to `makeRequest(..)`, though in this case we're not using it anymore. It's still listed, for consistency with the previous snippet.

By placing `recordKind` inside `setupButtonHandler(..)`, we limit the scope exposure of that variable to a more appropriate subset of the program; storing it globally would have been worse for code organization and readability. Closure lets the inner `makeRequest()` function instance *remember* this variable and access whenever it's needed.

Building on this pattern, we could have looked up both the URL and data once, at setup:

```

function setupButtonHandler(btn) {
  var recordKind = btn.dataset.kind;
  var requestURL = APIendpoints[recordKind];
  var requestData = data[recordKind];

  btn.addEventListener(
    "click",
    function makeRequest(evt){
      ajax(requestURL, requestData);
    }
  );
}

```

Now `makeRequest(..)` is closed over `requestURL` and `requestData`, which is a little bit cleaner to understand, and also slightly more performant.

Two similar techniques from the Functional Programming (FP) paradigm that rely on closure are partial application and currying. Briefly, with these techniques, we alter the *shape* of functions that require multiple inputs so some inputs are provided up front, and other inputs are provided later; the initial inputs are remembered via closure. Once all inputs have been provided, the underlying action is performed.

By creating a function instance that encapsulates some information inside (via closure), the function-with-stored-information can later be used directly without needing to re-provide that input. This makes that part of the code cleaner, and also offers the opportunity to label partially applied functions with better semantic names.

Adapting partial application, we can further improve the preceding code:

```
function defineHandler(requestURL,requestData) {
  return function makeRequest(evt){
    ajax(requestURL,requestData);
  };
}

function setupButtonHandler(btn) {
  var recordKind = btn.dataset.kind;
  var handler = defineHandler(
    APIendpoints[recordKind],
    data[recordKind]
  );
  btn.addEventListener("click",handler);
}
```

The `requestURL` and `requestData` inputs are provided ahead of time, resulting in the `makeRequest(..)` partially applied function, which we locally label `handler`. When the event eventually fires, the final input (`evt`, even though it's ignored) is passed to `handler()`, completing its inputs and triggering the underlying Ajax request.

Behavior-wise, this program is pretty similar to the previous one, with the same type of closure. But by isolating the creation of `makeRequest(..)` in a separate utility (`defineHandler(..)`), we make that definition more reusable across the program. We also explicitly limit the closure scope to only the two variables needed.

Closer to Closure

As we close down a dense chapter, take some deep breaths let it all sink in. Seriously, that's a lot of information for anyone to consume!

We explored two models for mentally tackling closure:

- Observational: closure is a function instance remembering its outer variables even as that function is passed to and **invoked in** other scopes.
- Implementational: closure is a function instance and its scope environment preserved in-place while any references to it are passed around and **invoked from** other scopes.

Summarizing the benefits to our programs:

- Closure can improve efficiency by allowing a function instance to remember previously determined information instead of having to compute it each time.
- Closure can improve code readability, bounding scope-exposure by encapsulating variable(s) inside function instances, while still making sure the information in those variables is accessible for future use. The resultant narrower, more specialized function instances are cleaner to interact with, since the preserved information doesn't need to be passed in every invocation.

Before you move on, take some time to restate this summary *in your own words*, explaining what closure is and why it's helpful in your programs. The main book text concludes with a final chapter that builds on top of closure with the module pattern.

You Don't Know JS Yet: Scope & Closures - 2nd Edition

Chapter 8: The Module Pattern

In this chapter, we wrap up the main text of the book by exploring one of the most important code organization patterns in all of programming: the module. As we'll see, modules are inherently built from what we've already covered: the payoff for your efforts in learning lexical scope and closure.

We've examined every angle of lexical scope, from the breadth of the global scope down through nested block scopes, into the intricacies of the variable lifecycle. Then we leveraged lexical scope to understand the full power of closure.

Take a moment to reflect on how far you've come in this journey so far; you've taken big steps in getting to know JS more deeply!

The central theme of this book has been that understanding and mastering scope and closure is key in properly structuring and organizing our code, especially the decisions on where to store information in variables.

Our goal in this final chapter is to appreciate how modules embody the importance of these topics, elevating them from abstract concepts to concrete, practical improvements in building programs.

Encapsulation and Least Exposure (POLE)

Encapsulation is often cited as a principle of object-oriented (OO) programming, but it's more fundamental and broadly applicable than that. The goal of encapsulation is the bundling or co-location of information (data) and behavior (functions) that together serve a common purpose.

Independent of any syntax or code mechanisms, the spirit of encapsulation can be realized in something as simple as using separate files to hold bits of the overall program with common purpose. If we bundle everything that powers a list of search results into a single file called "search-list.js", we're encapsulating that part of the program.

The recent trend in modern front-end programming to organize applications around Component architecture pushes encapsulation even further. For many, it feels natural to consolidate everything that constitutes the search results list—even beyond code, including presentational markup and styling—into a single unit of program logic, something tangible we can interact with. And then we label that collection the "SearchList" component.

Another key goal is the control of visibility of certain aspects of the encapsulated data and functionality. Recall from Chapter 6 the *least exposure* principle (POLE), which seeks to defensively guard against various *dangers* of scope

over-exposure; these affect both variables and functions. In JS, we most often implement visibility control through the mechanics of lexical scope.

The idea is to group alike program bits together, and selectively limit programmatic access to the parts we consider *private* details. What's not considered *private* is then marked as *public*, accessible to the whole program.

The natural effect of this effort is better code organization. It's easier to build and maintain software when we know where things are, with clear and obvious boundaries and connection points. It's also easier to maintain quality if we avoid the pitfalls of over-exposed data and functionality.

These are some of the main benefits of organizing JS programs into modules.

What Is a Module?

A module is a collection of related data and functions (often referred to as methods in this context), characterized by a division between hidden *private* details and *public* accessible details, usually called the “public API.”

A module is also stateful: it maintains some information over time, along with functionality to access and update that information.

NOTE:

A broader concern of the module pattern is fully embracing system-level modularization through loose-coupling and other program architecture techniques. That's a complex topic well beyond the bounds of our discussion, but is worth further study beyond this book.

To get a better sense of what a module is, let's compare some module characteristics to useful code patterns that aren't quite modules.

Namespaces (Stateless Grouping)

If you group a set of related functions together, without data, then you don't really have the expected encapsulation a module implies. The better term for this grouping of *stateless* functions is a namespace:

```
// namespace, not module
var Utils = {
  cancelEvt(evt) {
    evt.preventDefault();
    evt.stopPropagation();
    evt.stopImmediatePropagation();
  },
  wait(ms) {
    return new Promise(function c(res){
```

```

        setTimeout(res,ms);
    });
},
isValidEmail(email) {
    return /^[^@]+@[^@.]+\\.[^@.]+/.test(email);
}
};

```

Utils here is a useful collection of utilities, yet they're all state-independent functions. Gathering functionality together is generally good practice, but that doesn't make this a module. Rather, we've defined a Utils namespace and organized the functions under it.

Data Structures (Stateful Grouping)

Even if you bundle data and stateful functions together, if you're not limiting the visibility of any of it, then you're stopping short of the POLE aspect of encapsulation; it's not particularly helpful to label that a module.

Consider:

```

// data structure, not module
var Student = {
    records: [
        { id: 14, name: "Kyle", grade: 86 },
        { id: 73, name: "Suzy", grade: 87 },
        { id: 112, name: "Frank", grade: 75 },
        { id: 6, name: "Sarah", grade: 91 }
    ],
    getName(studentID) {
        var student = this.records.find(
            student => student.id == studentID
        );
        return student.name;
    }
};

Student.getName(73);
// Suzy

```

Since records is publicly accessible data, not hidden behind a public API, Student here isn't really a module.

Student does have the data-and-functionality aspect of encapsulation, but not the visibility-control aspect. It's best to label this an instance of a data structure.

Modules (Stateful Access Control)

To embody the full spirit of the module pattern, we not only need grouping and state, but also access control through visibility (private vs. public).

Let's turn `Student` from the previous section into a module. We'll start with a form I call the "classic module," which was originally referred to as the "revealing module" when it first emerged in the early 2000s. Consider:

```
var Student = (function defineStudent(){
  var records = [
    { id: 14, name: "Kyle", grade: 86 },
    { id: 73, name: "Suzy", grade: 87 },
    { id: 112, name: "Frank", grade: 75 },
    { id: 6, name: "Sarah", grade: 91 }
  ];

  var publicAPI = {
    getName
  };

  return publicAPI;

  // *****

  function getName(studentID) {
    var student = records.find(
      student => student.id == studentID
    );
    return student.name;
  }
})();

Student.getName(73); // Suzy
```

`Student` is now an instance of a module. It features a public API with a single method: `getName(..)`. This method is able to access the private hidden `records` data.

WARNING:

I should point out that the explicit student data being hard-coded into this module definition is just for our illustration purposes. A typical module in your program will receive this data from an outside source, typically loaded from databases, JSON data files, Ajax calls, etc. The data is then injected into the module instance typically through method(s) on the module's public API.

How does the classic module format work?

Notice that the instance of the module is created by the `defineStudent()` IIFE being executed. This IIFE returns an object (named `publicAPI`) that has a property on it referencing the inner `getName(..)` function.

Naming the object `publicAPI` is stylistic preference on my part. The object can be named whatever you like (JS doesn't care), or you can just return an object directly without assigning it to any internal named variable. More on this choice in Appendix A.

From the outside, `Student.getName(..)` invokes this exposed inner function, which maintains access to the inner `records` variable via closure.

You don't *have* to return an object with a function as one of its properties. You could just return a function directly, in place of the object. That still satisfies all the core bits of a classic module.

By virtue of how lexical scope works, defining variables and functions inside your outer module definition function makes everything *by default* private. Only properties added to the public API object returned from the function will be exported for external public use.

The use of an IIFE implies that our program only ever needs a single central instance of the module, commonly referred to as a "singleton." Indeed, this specific example is simple enough that there's no obvious reason we'd need anything more than just one instance of the `Student` module.

Module Factory (Multiple Instances) But if we did want to define a module that supported multiple instances in our program, we can slightly tweak the code:

```
// factory function, not singleton IIFE
function defineStudent() {
  var records = [
    { id: 14, name: "Kyle", grade: 86 },
    { id: 73, name: "Suzy", grade: 87 },
    { id: 112, name: "Frank", grade: 75 },
    { id: 6, name: "Sarah", grade: 91 }
  ];

  var publicAPI = {
    getName
  };

  return publicAPI;
}

// ****

function getName(studentID) {
```

```

        var student = records.find(
            student => student.id == studentID
        );
        return student.name;
    }
}

var fullTime = defineStudent();
fullTime.getName(73); // Suzy

```

Rather than specifying `defineStudent()` as an IIFE, we just define it as a normal standalone function, which is commonly referred to in this context as a “module factory” function.

We then call the module factory, producing an instance of the module that we label `fullTime`. This module instance implies a new instance of the inner scope, and thus a new closure that `getName(..)` holds over `records`. `fullTime.getName(..)` now invokes the method on that specific instance.

Classic Module Definition So to clarify what makes something a classic module:

- There must be an outer scope, typically from a module factory function running at least once.
- The module’s inner scope must have at least one piece of hidden information that represents state for the module.
- The module must return on its public API a reference to at least one function that has closure over the hidden module state (so that this state is actually preserved).

You’ll likely run across other variations on this classic module approach, which we’ll look at in more detail in Appendix A.

Node CommonJS Modules

In Chapter 4, we introduced the CommonJS module format used by Node. Unlike the classic module format described earlier, where you could bundle the module factory or IIFE alongside any other code including other modules, CommonJS modules are file-based; one module per file.

Let’s tweak our module example to adhere to that format:

```

module.exports.getName = getName;

// *****

var records = [
    { id: 14, name: "Kyle", grade: 86 },

```

```

        { id: 73, name: "Suzy", grade: 87 },
        { id: 112, name: "Frank", grade: 75 },
        { id: 6, name: "Sarah", grade: 91 }
    ];

    function getName(studentID) {
        var student = records.find(
            student => student.id == studentID
        );
        return student.name;
    }
}

```

The `records` and `getName` identifiers are in the top-level scope of this module, but that's not the global scope (as explained in Chapter 4). As such, everything here is *by default* private to the module.

To expose something on the public API of a CommonJS module, you add a property to the empty object provided as `module.exports`. In some older legacy code, you may run across references to just a bare `exports`, but for code clarity you should always fully qualify that reference with the `module.` prefix.

For style purposes, I like to put my “exports” at the top and my module implementation at the bottom. But these exports can be placed anywhere. I strongly recommend collecting them all together, either at the top or bottom of your file.

Some developers have the habit of replacing the default exports object, like this:

```

// defining a new object for the API
module.exports = {
    // ..exports..
};

```

There are some quirks with this approach, including unexpected behavior if multiple such modules circularly depend on each other. As such, I recommend against replacing the object. If you want to assign multiple exports at once, using object literal style definition, you can do this instead:

```

Object.assign(module.exports, {
    // .. exports ..
});

```

What's happening here is defining the `{ .. }` object literal with your module's public API specified, and then `Object.assign(..)` is performing a shallow copy of all those properties onto the existing `module.exports` object, instead of replacing it. This is a nice balance of convenience and safer module behavior.

To include another module instance into your module/program, use Node's `require(..)` method. Assuming this module is located at “/path/to/student.js”, this is how we can access it:

```
var Student = require("/path/to/student.js");

Student.getName(73);
// Suzy
```

Student now references the public API of our example module.

CommonJS modules behave as singleton instances, similar to the IIFE module definition style presented before. No matter how many times you `require(..)` the same module, you just get additional references to the single shared module instance.

`require(..)` is an all-or-nothing mechanism; it includes a reference of the entire exposed public API of the module. To effectively access only part of the API, the typical approach looks like this:

```
var getName = require("/path/to/student.js").getName;

// or alternately:
```

```
var { getName } = require("/path/to/student.js");
```

Similar to the classic module format, the publicly exported methods of a CommonJS module's API hold closures over the internal module details. That's how the module singleton state is maintained across the lifetime of your program.

NOTE:

In Node `require("student")` statements, non-absolute paths ("student") assume a ".js" file extension and search "node_modules".

Modern ES Modules (ESM)

The ESM format shares several similarities with the CommonJS format. ESM is file-based, and module instances are singletons, with everything private *by default*. One notable difference is that ESM files are assumed to be strict-mode, without needing a "use strict" pragma at the top. There's no way to define an ESM as non-strict-mode.

Instead of `module.exports` in CommonJS, ESM uses an `export` keyword to expose something on the public API of the module. The `import` keyword replaces the `require(..)` statement. Let's adjust "students.js" to use the ESM format:

```
export { getName };

// *****

var records = [
  { id: 14, name: "Kyle", grade: 86 },
```

```

        { id: 73, name: "Suzy", grade: 87 },
        { id: 112, name: "Frank", grade: 75 },
        { id: 6, name: "Sarah", grade: 91 }
    ];
}

function getName(studentID) {
    var student = records.find(
        student => student.id == studentID
    );
    return student.name;
}

```

The only change here is the `export { getName }` statement. As before, `export` statements can appear anywhere throughout the file, though `export` must be at the top-level scope; it cannot be inside any other block or function.

ESM offers a fair bit of variation on how the `export` statements can be specified. For example:

```

export function getName(studentID) {
    // ...
}

```

Even though `export` appears before the `function` keyword here, this form is still a `function` declaration that also happens to be exported. That is, the `getName` identifier is *function hoisted* (see Chapter 5), so it's available throughout the whole scope of the module.

Another allowed variation:

```

export default function getName(studentID) {
    // ...
}

```

This is a so-called “default export,” which has different semantics from other exports. In essence, a “default export” is a shorthand for consumers of the module when they `import`, giving them a terser syntax when they only need this single default API member.

Non-default exports are referred to as “named exports.”

The `import` keyword—like `export`, it must be used only at the top level of an ESM outside of any blocks or functions—also has a number of variations in syntax. The first is referred to as “named import”:

```

import { getName } from "/path/to/students.js";

getName(73); // Suzy

```

As you can see, this form imports only the specifically named public API members from a module (skipping anything not named explicitly), and it adds those

identifiers to the top-level scope of the current module. This type of import is a familiar style to those used to package imports in languages like Java.

Multiple API members can be listed inside the `{ .. }` set, separated with commas. A named import can also be *renamed* with the `as` keyword:

```
import { getName as getStudentName }  
  from "/path/to/students.js";
```

```
getStudentName(73);  
// Suzy
```

If `getName` is a “default export” of the module, we can import it like this:

```
import getName from "/path/to/students.js";  
  
getName(73); // Suzy
```

The only difference here is dropping the `{ .. }` around the import binding. If you want to mix a default import with other named imports:

```
import { default as getName, /* .. others .. */ }  
  from "/path/to/students.js";  
  
getName(73); // Suzy
```

By contrast, the other major variation on `import` is called “namespace import”:

```
import * as Student from "/path/to/students.js";  
  
Student.getName(73); // Suzy
```

As is likely obvious, the `*` imports everything exported to the API, default and named, and stores it all under the single namespace identifier as specified. This approach most closely matches the form of classic modules for most of JS’s history.

NOTE:

As of the time of this writing, modern browsers have supported ESM for a few years now, but Node’s stable-ish support for ESM is fairly recent, and has been evolving for quite a while. The evolution is likely to continue for another year or more; the introduction of ESM to JS back in ES6 created a number of challenging compatibility concerns for Node’s interop with CommonJS modules. Consult Node’s ESM documentation for all the latest details:
<https://nodejs.org/api/esm.html>

Exit Scope

Whether you use the classic module format (browser or Node), CommonJS format (in Node), or ESM format (browser or Node), modules are one of the most effective ways to structure and organize your program's functionality and data.

The module pattern is the conclusion of our journey in this book of learning how we can use the rules of lexical scope to place variables and functions in proper locations. POLE is the defensive *private by default* posture we always take, making sure we avoid over-exposure and interact only with the minimal public API surface area necessary.

And underneath modules, the *magic* of how all our module state is maintained is closures leveraging the lexical scope system.

That's it for the main text. Congratulations on quite a journey so far! As I've said numerous times throughout, it's a really good idea to pause, reflect, and practice what we've just discussed.

When you're comfortable and ready, check out the appendices, which dig deeper into some of the corners of these topics, and also challenge you with some practice exercises to solidify what you've learned.

You Don't Know JS Yet: Scope & Closures - 2nd Edition

Appendix A: Exploring Further

We will now explore a number of nuances and edges around many of the topics covered in the main text of this book. This appendix is optional, supporting material.

Some people find diving too deeply into the nuanced corner cases and varying opinions creates nothing but noise and distraction—supposedly, developers are better served by sticking to the commonly-tread paths. My approach has been criticized as being impractical and counterproductive. I understand and appreciate that perspective, even if I don't necessarily share it.

I believe it's better to be empowered by knowledge of how things work than to just gloss over details with assumptions and lack of curiosity. Ultimately, you will encounter situations where something bubbles up from a piece you hadn't explored. In other words, you won't get to spend all your time riding on the smooth *happy path*. Wouldn't you rather be prepared for the inevitable bumps of off-roading?

These discussions will also be more heavily influenced by my opinions than the main text was, so keep that in mind as you consume and consider what is presented. This appendix is a bit like a collection of mini-blog posts that elaborate on various book topics. It's long and deep in the weeds, so take your time and don't rush through everything here.

Implied Scopes

Scopes are sometimes created in non-obvious places. In practice, these implied scopes don't often impact your program behavior, but it's still useful to know they're happening. Keep an eye out for the following surprising scopes:

- Parameter scope
- Function name scope

Parameter Scope

The conversation metaphor in Chapter 2 implies that function parameters are basically the same as locally declared variables in the function scope. But that's not always true.

Consider:

```
// outer/global scope: RED(1)

function getStudentName(studentID) {
```

```
// function scope: BLUE(2)

// ...
}
```

Here, `studentID` is considered a “simple” parameter, so it does behave as a member of the `BLUE(2)` function scope. But if we change it to be a non-simple parameter, that’s no longer technically the case. Parameter forms considered non-simple include parameters with default values, rest parameters (using `...`), and destructured parameters.

Consider:

```
// outer/global scope: RED(1)

function getStudentName(/*BLUE(2)*/ studentID = 0) {
    // function scope: GREEN(3)

    // ...
}
```

Here, the parameter list essentially becomes its own scope, and the function’s scope is then nested inside *that* scope.

Why? What difference does it make? The non-simple parameter forms introduce various corner cases, so the parameter list becomes its own scope to more effectively deal with them.

Consider:

```
function getStudentName(studentID = maxID, maxID) {
    // ...
}
```

Assuming left-to-right operations, the `default = maxID` for the `studentID` parameter requires a `maxID` to already exist (and to have been initialized). This code produces a TDZ error (Chapter 5). The reason is that `maxID` is declared in the parameter scope, but it’s not yet been initialized because of the order of parameters. If the parameter order is flipped, no TDZ error occurs:

```
function getStudentName(maxID,studentID = maxID) {
    // ...
}
```

The complication gets even more *in the weeds* if we introduce a function expression into the default parameter position, which then can create its own closure (Chapter 7) over parameters in this implied parameter scope:

```
function whatsTheDealHere(id,defaultID = () => id) {
    id = 5;
    console.log( defaultID() );
```

```

}

whatsTheDealHere(3);
// 5

```

That snippet probably makes sense, because the `defaultID()` arrow function closes over the `id` parameter/variable, which we then re-assign to 5. But now let's introduce a shadowing definition of `id` in the function scope:

```

function whatsTheDealHere(id,defaultID = () => id) {
  var id = 5;
  console.log( defaultID() );
}

whatsTheDealHere(3);
// 3

```

Uh oh! The `var id = 5` is shadowing the `id` parameter, but the closure of the `defaultID()` function is over the parameter, not the shadowing variable in the function body. This proves there's a scope bubble around the parameter list.

But it gets even crazier than that!

```

function whatsTheDealHere(id,defaultID = () => id) {
  var id;

  console.log(`local variable 'id': ${ id }`);
  console.log(
    `parameter 'id' (closure): ${ defaultID() }`
  );

  console.log("reassigning 'id' to 5");
  id = 5;

  console.log(`local variable 'id': ${ id }`);
  console.log(
    `parameter 'id' (closure): ${ defaultID() }`
  );
}

whatsTheDealHere(3);
// local variable 'id': 3  <-- Huh!? Weird!
// parameter 'id' (closure): 3
// reassigning 'id' to 5
// local variable 'id': 5
// parameter 'id' (closure): 3

```

The strange bit here is the first console message. At that moment, the shadowing `id` local variable has just been `var id` declared, which Chapter 5 asserts is

typically auto-initialized to `undefined` at the top of its scope. Why doesn't it print `undefined`?

In this specific corner case (for legacy compat reasons), JS doesn't auto-initialize `id` to `undefined`, but rather to the value of the `id` parameter (3)!

Though the two `ids` look at that moment like they're one variable, they're actually still separate (and in separate scopes). The `id = 5` assignment makes the divergence observable, where the `id` parameter stays 3 and the local variable becomes 5.

My advice to avoid getting bitten by these weird nuances:

- Never shadow parameters with local variables
- Avoid using a default parameter function that closes over any of the parameters

At least now you're aware and can be careful about the fact that the parameter list is its own scope if any of the parameters are non-simple.

Function Name Scope

In the “Function Name Scope” section in Chapter 3, I asserted that the name of a function expression is added to the function's own scope. Recall:

```
var askQuestion = function ofTheTeacher(){  
  // ..  
};
```

It's true that `ofTheTeacher` is not added to the enclosing scope (where `askQuestion` is declared), but it's also not *just* added to the scope of the function, the way you're likely assuming. It's another strange corner case of implied scope.

The name identifier of a function expression is in its own implied scope, nested between the outer enclosing scope and the main inner function scope.

If `ofTheTeacher` was in the function's scope, we'd expect an error here:

```
var askQuestion = function ofTheTeacher(){  
  // why is this not a duplicate declaration error?  
  let ofTheTeacher = "Confused, yet?";  
};
```

The `let` declaration form does not allow re-declaration (see Chapter 5). But this is perfectly legal shadowing, not re-declaration, because the two `ofTheTeacher` identifiers are in separate scopes.

You'll rarely run into any case where the scope of a function's name identifier matters. But again, it's good to know how these mechanisms actually work. To avoid being bitten, never shadow function name identifiers.

Anonymous vs. Named Functions

As discussed in Chapter 3, functions can be expressed either in named or anonymous form. It's vastly more common to use the anonymous form, but is that a good idea?

As you contemplate naming your functions, consider:

- Name inference is incomplete
- Lexical names allow self-reference
- Names are useful descriptions
- Arrow functions have no lexical names
- IIFEs also need names

Explicit or Inferred Names?

Every function in your program has a purpose. If it doesn't have a purpose, take it out, because you're just wasting space. If it *does* have a purpose, there *is* a name for that purpose.

So far many readers likely agree with me. But does that mean we should always put that name into the code? Here's where I'll raise more than a few eyebrows. I say, unequivocally, yes!

First of all, “anonymous” showing up in stack traces is just not all that helpful to debugging:

```
btn.addEventListener("click",function(){
  setTimeout(function(){
    ["a",42].map(function(v){
      console.log(v.toUpperCase());
    });
  },100);
});  
// Uncaught TypeError: v.toUpperCase is not a function
//     at myProgram.js:4
//     at Array.map (<anonymous>)
//     at myProgram.js:3
```

Ugh. Compare to what is reported if I give the functions names:

```
btn.addEventListener("click",function onClick(){
  setTimeout(function waitAMoment(){
    ["a",42].map(function allUpper(v){
      console.log(v.toUpperCase());
    });
  },100);
});  
// Uncaught TypeError: v.toUpperCase is not a function
//     at allUpper (myProgram.js:4)
```

```
//      at Array.map (<anonymous>)
//      at waitAMoment (myProgram.js:3)
```

See how `waitAMoment` and `allUpper` names appear and give the stack trace more useful information/context for debugging? The program is more debuggable if we use reasonable names for all our functions.

NOTE:

The unfortunate “`<anonymous>`” that still shows up refers to the fact that the implementation of `Array.map(..)` isn’t present in our program, but is built into the JS engine. It’s not from any confusion our program introduces with readability shortcuts.

By the way, let’s make sure we’re on the same page about what a named function is:

```
function thisIsNamed() {
  // ..
}

ajax("some.url",function thisIsAlsoNamed(){
  // ..
});

var notNamed = function(){
  // ..
};

makeRequest({
  data: 42,
  cb /* also not a name */: function(){
    // ..
  }
});

var stillNotNamed = function butThisIs(){
  // ..
};

“But wait!”, you say. Some of those are named, right!?

var notNamed = function(){
  // ..
};

var config = {
```

```

cb: function(){
  // ..
}
};

notNamed.name;
// notNamed

config.cb.name;
// cb

```

These are referred to as *inferred* names. Inferred names are fine, but they don't really address the full concern I'm discussing.

Missing Names?

Yes, these inferred names might show up in stack traces, which is definitely better than "anonymous" showing up. But...

```

function ajax(url,cb) {
  console.log(cb.name);
}

ajax("some.url",function(){
  // ..
});
// ""

```

Oops. Anonymous `function` expressions passed as callbacks are incapable of receiving an inferred name, so `cb.name` holds just the empty string `""`. The vast majority of all `function` expressions, especially anonymous ones, are used as callback arguments; none of these get a name. So relying on name inference is incomplete, at best.

And it's not just callbacks that fall short with inference:

```

var config = {};

config.cb = function(){
  // ..
};

config.cb.name;
// ""

var [ noName ] = [ function(){ } ];
noName.name
// ""

```

Any assignment of a `function` expression that's not a *simple assignment* will also fail name inferencing. So, in other words, unless you're careful and intentional about it, essentially almost all anonymous `function` expressions in your program will in fact have no name at all.

Name inference is just... not enough.

And even if a `function` expression *does* get an inferred name, that still doesn't count as being a full named function.

Who am I?

Without a lexical name identifier, the function has no internal way to refer to itself. Self-reference is important for things like recursion and event handling:

```
// broken
runOperation(function(num){
  if (num <= 1) return 1;
  return num * oopsNoNameToCall(num - 1);
});

// also broken
btn.addEventListener("click",function(){
  console.log("should only respond to one click!");
  btn.removeEventListener("click",oopsNoNameHere);
});
```

Leaving off the lexical name from your callback makes it harder to reliably self-reference the function. You *could* declare a variable in an enclosing scope that references the function, but this variable is *controlled* by that enclosing scope—it could be re-assigned, etc.—so it's not as reliable as the function having its own internal self-reference.

Names are Descriptors

Lastly, and I think most importantly of all, leaving off a name from a function makes it harder for the reader to tell what the function's purpose is, at a quick glance. They have to read more of the code, including the code inside the function, and the surrounding code outside the function, to figure it out.

Consider:

```
[ 1, 2, 3, 4, 5 ].filter(function(v){
  return v % 2 == 1;
});
// [ 1, 3, 5 ]

[ 1, 2, 3, 4, 5 ].filter(function keepOnlyOdds(v){
  return v % 2 == 1;
```

```
});  
// [ 1, 3, 5 ]
```

There's just no reasonable argument to be made that **omitting** the name `keepOnlyOdds` from the first callback more effectively communicates to the reader the purpose of this callback. You saved 13 characters, but lost important readability information. The name `keepOnlyOdds` very clearly tells the reader, at a quick first glance, what's happening.

The JS engine doesn't care about the name. But human readers of your code absolutely do.

Can the reader look at `v % 2 == 1` and figure out what it's doing? Sure. But they have to infer the purpose (and name) by mentally executing the code. Even a brief pause to do so slows down reading of the code. A good descriptive name makes this process almost effortless and instant.

Think of it this way: how many times does the author of this code need to figure out the purpose of a function before adding the name to the code? About once. Maybe two or three times if they need to adjust the name. But how many times will readers of this code have to figure out the name/purpose? Every single time this line is ever read. Hundreds of times? Thousands? More?

No matter the length or complexity of the function, my assertion is, the author should figure out a good descriptive name and add it to the code. Even the one-liner functions in `map(..)` and `then(..)` statements should be named:

```
lookupTheRecords(someData)  
.then(function extractSalesRecords(resp){  
    return resp.allSales;  
})  
.then(storeRecords);
```

The name `extractSalesRecords` tells the reader the purpose of this `then(..)` handler *better* than just inferring that purpose from mentally executing `return resp.allSales`.

The only excuse for not including a name on a function is either laziness (don't want to type a few extra characters) or uncreativity (can't come up with a good name). If you can't figure out a good name, you likely don't understand the function and its purpose yet. The function is perhaps poorly designed, or it does too many things, and should be re-worked. Once you have a well-designed, single-purpose function, its proper name should become evident.

Here's a trick I use: while first writing a function, if I don't fully understand its purpose and can't think of a good name to use, I just use `TODO` as the name. That way, later when reviewing my code, I'm likely to find those name placeholders, and I'm more inclined (and more prepared!) to go back and figure out a better name, rather than just leave it as `TODO`.

All functions need names. Every single one. No exceptions. Any name you omit is making the program harder to read, harder to debug, harder to extend and maintain later.

Arrow Functions

Arrow functions are **always** anonymous, even if (rarely) they're used in a way that gives them an inferred name. I just spent several pages explaining why anonymous functions are a bad idea, so you can probably guess what I think about arrow functions.

Don't use them as a general replacement for regular functions. They're more concise, yes, but that brevity comes at the cost of omitting key visual delimiters that help our brains quickly parse out what we're reading. And, to the point of this discussion, they're anonymous, which makes them worse for readability from that angle as well.

Arrow functions have a purpose, but that purpose is not to save keystrokes. Arrow functions have *lexical this* behavior, which is somewhat beyond the bounds of our discussion in this book.

Briefly: arrow functions don't define a `this` identifier keyword at all. If you use a `this` inside an arrow function, it behaves exactly as any other variable reference, which is that the scope chain is consulted to find a function scope (non-arrow function) where it *is* defined, and to use that one.

In other words, arrow functions treat `this` like any other lexical variable.

If you're used to hacks like `var self = this`, or if you prefer to call `.bind(this)` on inner `function` expressions, just to force them to inherit a `this` from an outer function like it was a lexical variable, then => arrow functions are absolutely the better option. They're designed specifically to fix that problem.

So, in the rare cases you need *lexical this*, use an arrow function. It's the best tool for that job. But just be aware that in doing so, you're accepting the downsides of an anonymous function. You should expend additional effort to mitigate the readability *cost*, such as more descriptive variable names and code comments.

IIFE Variations

All functions should have names. I said that a few times, right!? That includes IIFEs.

```
(function(){  
  // don't do this!  
})();  
  
(function doThisInstead(){
```

```
// ...
})();
```

How do we come up with a name for an IIFE? Identify what the IIFE is there for. Why do you need a scope in that spot? Are you hiding a cache variable for student records?

```
var getStudents = (function StoreStudentRecords(){
  var studentRecords = [];

  return function getStudents() {
    // ...
  }
})();
```

I named the IIFE `StoreStudentRecords` because that's what it's doing: storing student records. Every IIFE should have a name. No exceptions.

IIFEs are typically defined by placing (..) around the `function` expression, as shown in those previous snippets. But that's not the only way to define an IIFE. Technically, the only reason we're using that first surrounding set of (..) is just so the `function` keyword isn't in a position to qualify as a `function` declaration to the JS parser. But there are other syntactic ways to avoid being parsed as a declaration:

```
!function thisIsAnIIFE(){
  // ...
}();

+function soIsThisOne(){
  // ...
}();

~function andThisOneToo(){
  // ...
}();
```

The `!`, `+`, `~`, and several other unary operators (operators with one operand) can all be placed in front of `function` to turn it into an expression. Then the final `()` call is valid, which makes it an IIFE.

I actually kind of like using the `void` unary operator when defining a standalone IIFE:

```
void function yepItsAnIIFE() {
  // ...
}();
```

The benefit of `void` is, it clearly communicates at the beginning of the function that this IIFE won't be returning any value.

However you define your IIFEs, show them some love by giving them names.

Hoisting: Functions and Variables

Chapter 5 articulated both *function hoisting* and *variable hoisting*. Since hoisting is often cited as mistake in the design of JS, I wanted to briefly explore why both these forms of hoisting *can* be beneficial and should still be considered.

Give hoisting a deeper level of consideration by considering the merits of:

- Executable code first, function declarations last
- Semantic placement of variable declarations

Function Hoisting

To review, this program works because of *function hoisting*:

```
getStudents();  
  
// ..  
  
function getStudents() {  
    // ..  
}
```

The `function` declaration is hoisted during compilation, which means that `getStudents` is an identifier declared for the entire scope. Additionally, the `getStudents` identifier is auto-initialized with the function reference, again at the beginning of the scope.

Why is this useful? The reason I prefer to take advantage of *function hoisting* is that it puts the *executable* code in any scope at the top, and any further declarations (functions) below. This means it's easier to find the code that will run in any given area, rather than having to scroll and scroll, hoping to find a trailing `}` marking the end of a scope/function somewhere.

I take advantage of this inverse positioning in all levels of scope:

```
getStudents();  
  
// *****  
  
function getStudents() {  
    var whatever = doSomething();  
  
    // other stuff  
  
    return whatever;
```

```
// ****
function doSomething() {
  // ..
}
}
```

When I first open a file like that, the very first line is executable code that kicks off its behavior. That's very easy to spot! Then, if I ever need to go find and inspect `getStudents()`, I like that its first line is also executable code. Only if I need to see the details of `doSomething()` do I go and find its definition down below.

In other words, I think *function hoisting* makes code more readable through a flowing, progressive reading order, from top to bottom.

Variable Hoisting

What about *variable hoisting*?

Even though `let` and `const` hoist, you cannot use those variables in their TDZ (see Chapter 5). So, the following discussion only applies to `var` declarations. Before I continue, I'll admit: in almost all cases, I completely agree that *variable hoisting* is a bad idea:

```
pleaseDontDoThis = "bad idea";

// much later
var pleaseDontDoThis;
```

While that kind of inverted ordering was helpful for *function hoisting*, here I think it usually makes code harder to reason about.

But there's one exception that I've found, somewhat rarely, in my own coding. It has to do with where I place my `var` declarations inside a CommonJS module definition.

Here's how I typically structure my module definitions in Node:

```
// dependencies
var aModuleINeed = require("very-helpful");
var anotherModule = require("kinda-helpful");

// public API
var publicAPI = Object.assign(module.exports, {
  getStudents,
  addStudents,
  // ..
});
```

```

// ****
// private implementation

var cache = {};
var otherData = [ ];

function getStudents() {
    // ...
}

function addStudents() {
    // ...
}

```

Notice how the `cache` and `otherData` variables are in the “private” section of the module layout? That’s because I don’t plan to expose them publicly. So I organize the module so they’re located alongside the other hidden implementation details of the module.

But I’ve had a few rare cases where I needed the assignments of those values to happen *above*, before I declare the exported public API of the module. For instance:

```

// public API
var publicAPI = Object.assign(module.exports, {
    getStudents,
    addStudents,
    refreshData: reloadData.bind(null, cache)
});

```

I need the `cache` variable to have already been assigned a value, because that value is used in the initialization of the public API (the `.bind(..)` partial-application).

Should I just move the `var cache = {}` up to the top, above this public API initialization? Well, perhaps. But now it’s less obvious that `var cache` is a *private* implementation detail. Here’s the compromise I’ve (somewhat rarely) used:

```

cache = {};
    // used here, but declared below

// public API
var publicAPI = Object.assign(module.exports, {
    getStudents,
    addStudents,
    refreshData: reloadData.bind(null, cache)
});

// ****

```

```
// private implementation
```

```
var cache /* = {} */;
```

See the *variable hoisting*? I've declared the `cache` down where it belongs, logically, but in this rare case I've used it earlier up above, in the area where its initialization is needed. I even left a hint at the value that's assigned to `cache` in a code comment.

That's literally the only case I've ever found for leveraging *variable hoisting* to assign a variable earlier in a scope than its declaration. But I think it's a reasonable exception to employ with caution.

The Case for `var`

Speaking of *variable hoisting*, let's have some real talk for a bit about `var`, a favorite villain devs love to blame for many of the woes of JS development. In Chapter 5, we explored `let`/`const` and promised we'd revisit where `var` falls in the whole mix.

As I lay out the case, don't miss:

- `var` was never broken
- `let` is your friend
- `const` has limited utility
- The best of both worlds: `var` and `let`

Don't Throw Out `var`

`var` is fine, and works just fine. It's been around for 25 years, and it'll be around and useful and functional for another 25 years or more. Claims that `var` is broken, deprecated, outdated, dangerous, or ill-designed are bogus bandwagoning.

Does that mean `var` is the right declarator for every single declaration in your program? Certainly not. But it still has its place in your programs. Refusing to use it because someone on the team chose an aggressive linter opinion that chokes on `var` is cutting off your nose to spite your face.

OK, now that I've got you really riled up, let me try to explain my position.

For the record, I'm a fan of `let`, for block-scoped declarations. I really dislike TDZ and I think that was a mistake. But `let` itself is great. I use it often. In fact, I probably use it as much or more than I use `var`.

`const`-antly Confused

`const` on the other hand, I don't use as often. I'm not going to dig into all the reasons why, but it comes down to `const` not *carrying its own weight*. That is, while there's a tiny bit of benefit of `const` in some cases, that benefit is

outweighed by the long history of troubles around `const` confusion in a variety of languages, long before it ever showed up in JS.

`const` pretends to create values that can't be mutated—a misconception that's extremely common in developer communities across many languages—whereas what it really does is prevent re-assignment.

```
const studentIDs = [ 14, 73, 112 ];  
  
// later  
  
studentIDs.push(6); // whoa, wait... what!?
```

Using a `const` with a mutable value (like an array or object) is asking for a future developer (or reader of your code) to fall into the trap you set, which was that they either didn't know, or sorta forgot, that *value immutability* isn't at all the same thing as *assignment immutability*.

I just don't think we should set those traps. The only time I ever use `const` is when I'm assigning an already-immutable value (like 42 or "Hello, friends!"), and when it's clearly a "constant" in the sense of being a named placeholder for a literal value, for semantic purposes. That's what `const` is best used for. That's pretty rare in my code, though.

If variable re-assignment were a big deal, then `const` would be more useful. But variable re-assignment just isn't that big of a deal in terms of causing bugs. There's a long list of things that lead to bugs in programs, but "accidental re-assignment" is way, way down that list.

Combine that with the fact that `const` (and `let`) are supposed to be used in blocks, and blocks are supposed to be short, and you have a really small area of your code where a `const` declaration is even applicable. A `const` on line 1 of your ten-line block only tells you something about the next nine lines. And the thing it tells you is already obvious by glancing down at those nine lines: the variable is never on the left-hand side of an `=`; it's not re-assigned.

That's it, that's all `const` really does. Other than that, it's not very useful. Stacked up against the significant confusion of value vs. assignment immutability, `const` loses a lot of its luster.

A `let` (or `var!`) that's never re-assigned is already behaviorally a "constant", even though it doesn't have the compiler guarantee. That's good enough in most cases.

`var` and `let`

In my mind, `const` is pretty rarely useful, so this is only two-horse race between `let` and `var`. But it's not really a race either, because there doesn't have to be just one winner. They can both win... different races.

The fact is, you should be using both `var` and `let` in your programs. They are not interchangeable: you shouldn't use `var` where a `let` is called for, but you also shouldn't use `let` where a `var` is most appropriate.

So where should we still use `var`? Under what circumstances is it a better choice than `let`?

For one, I always use `var` in the top-level scope of any function, regardless of whether that's at the beginning, middle, or end of the function. I also use `var` in the global scope, though I try to minimize usage of the global scope.

Why use `var` for function scoping? Because that's exactly what `var` does. There literally is no better tool for the job of function scoping a declaration than a declarator that has, for 25 years, done exactly that.

You *could* use `let` in this top-level scope, but it's not the best tool for that job. I also find that if you use `let` everywhere, then it's less obvious which declarations are designed to be localized and which ones are intended to be used throughout the function.

By contrast, I rarely use a `var` inside a block. That's what `let` is for. Use the best tool for the job. If you see a `let`, it tells you that you're dealing with a localized declaration. If you see `var`, it tells you that you're dealing with a function-wide declaration. Simple as that.

```
function getStudents(data) {
  var studentRecords = [];

  for (let record of data.records) {
    let id = `student-${record.id}`;
    studentRecords.push({
      id,
      record.name
    });
  }

  return studentRecords;
}
```

The `studentRecords` variable is intended for use across the whole function. `var` is the best declarator to tell the reader that. By contrast, `record` and `id` are intended for use only in the narrower scope of the loop iteration, so `let` is the best tool for that job.

In addition to this *best tool* semantic argument, `var` has a few other characteristics that, in certain limited circumstances, make it more powerful.

One example is when a loop is exclusively using a variable, but its conditional clause cannot see block-scoped declarations inside the iteration:

```

function commitAction() {
  do {
    let result = commit();
    var done = result && result.code == 1;
  } while (!done);
}

```

Here, `result` is clearly only used inside the block, so we use `let`. But `done` is a bit different. It's only useful for the loop, but the `while` clause cannot see `let` declarations that appear inside the loop. So we compromise and use `var`, so that `done` is hoisted to the outer scope where it can be seen.

The alternative—declaring `done` outside the loop—separates it from where it's first used, and either necessitates picking a default value to assign, or worse, leaving it unassigned and thus looking ambiguous to the reader. I think `var` inside the loop is preferable here.

Another helpful characteristic of `var` is seen with declarations inside unintended blocks. Unintended blocks are blocks that are created because the syntax requires a block, but where the intent of the developer is not really to create a localized scope. The best illustration of unintended scope is the `try..catch` statement:

```

function getStudents() {
  try {
    // not really a block scope
    var records = fromCache("students");
  }
  catch (err) {
    // oops, fall back to a default
    var records = [];
  }
  // ...
}

```

There are other ways to structure this code, yes. But I think this is the *best* way, given various trade-offs.

I don't want to declare `records` (with `var` or `let`) outside of the `try` block, and then assign to it in one or both blocks. I prefer initial declarations to always be as close as possible (ideally, same line) to the first usage of the variable. In this simple example, that would only be a couple of lines distance, but in real code it can grow to many more lines. The bigger the gap, the harder it is to figure out what variable from what scope you're assigning to. `var` used at the actual assignment makes it less ambiguous.

Also notice I used `var` in both the `try` and `catch` blocks. That's because I want to signal to the reader that no matter which path is taken, `records` always gets declared. Technically, that works because `var` is hoisted once to the function

scope. But it's still a nice semantic signal to remind the reader what either `var` ensures. If `var` were only used in one of the blocks, and you were only reading the other block, you wouldn't as easily discover where `records` was coming from.

This is, in my opinion, a little superpower of `var`. Not only can it escape the unintentional `try..catch` blocks, but it's allowed to appear multiple times in a function's scope. You can't do that with `let`. It's not bad, it's actually a little helpful feature. Think of `var` more like a declarative annotation that's reminding you, each usage, where the variable comes from. "Ah ha, right, it belongs to the whole function."

This repeated-annotation superpower is useful in other cases:

```
function getStudents() {
  var data = [];

  // do something with data
  // ... 50 more lines of code ...

  // purely an annotation to remind us
  var data;

  // use data again
  //
}
```

The second `var data` is not re-declaring `data`, it's just annotating for the readers' benefit that `data` is a function-wide declaration. That way, the reader doesn't need to scroll up 50+ lines of code to find the initial declaration.

I'm perfectly fine with re-using variables for multiple purposes throughout a function scope. I'm also perfectly fine with having two usages of a variable be separated by quite a few lines of code. In both cases, the ability to safely "re-declare" (annotate) with `var` helps make sure I can tell where my `data` is coming from, no matter where I am in the function.

Again, sadly, `let` cannot do this.

There are other nuances and scenarios when `var` turns out to offer some assistance, but I'm not going to belabor the point any further. The takeaway is that `var` can be useful in our programs alongside `let` (and the occasional `const`). Are you willing to creatively use the tools the JS language provides to tell a richer story to your readers?

Don't just throw away a useful tool like `var` because someone shamed you into thinking it wasn't cool anymore. Don't avoid `var` because you got confused once years ago. Learn these tools and use them each for what they're best at.

What's the Deal with TDZ?

The TDZ (temporal dead zone) was explained in Chapter 5. We illustrated how it occurs, but we skimmed over any explanation of *why* it was necessary to introduce in the first place. Let's look briefly at the motivations of TDZ.

Some breadcrumbs in the TDZ origin story:

- `consts` should never change
- It's all about time
- Should `let` behave more like `const` or `var`?

Where It All Started

TDZ comes from `const`, actually.

During early ES6 development work, TC39 had to decide whether `const` (and `let`) were going to hoist to the top of their blocks. They decided these declarations would hoist, similar to how `var` does. Had that not been the case, I think some of the fear was confusion with mid-scope shadowing, such as:

```
let greeting = "Hi!";  
  
{  
  // what should print here?  
  console.log(greeting);  
  
  // ... a bunch of lines of code ...  
  
  // now shadowing the `greeting` variable  
  let greeting = "Hello, friends!";  
  
  // ...  
}
```

What should we do with that `console.log(..)` statement? Would it make any sense to JS devs for it to print “Hi!”? Seems like that could be a gotcha, to have shadowing kick in only for the second half of the block, but not the first half. That's not very intuitive, JS-like behavior. So `let` and `const` have to hoist to the top of the block, visible throughout.

But if `let` and `const` hoist to the top of the block (like `var` hoists to the top of a function), why don't `let` and `const` auto-initialize (to `undefined`) the way `var` does? Here was the main concern:

```
{  
  // what should print here?  
  console.log(studentName);  
  
  // later
```

```
const studentName = "Frank";  
  
// ..  
}
```

Let's imagine that `studentName` not only hoisted to the top of this block, but was also auto-initialized to `undefined`. For the first half of the block, `studentName` could be observed to have the `undefined` value, such as with our `console.log(..)` statement. Once the `const studentName = ..` statement is reached, now `studentName` is assigned "`Frank`". From that point forward, `studentName` can't ever be re-assigned.

But, is it strange or surprising that a constant observably has two different values, first `undefined`, then "`Frank`"? That does seem to go against what we think a `constant` means; it should only ever be observable with one value.

So... now we have a problem. We can't auto-initialize `studentName` to `undefined` (or any other value for that matter). But the variable has to exist throughout the whole scope. What do we do with the period of time from when it first exists (beginning of scope) and when it's assigned its value?

We call this period of time the "dead zone," as in the "temporal dead zone" (TDZ). To prevent confusion, it was determined that any sort of access of a variable while in its TDZ is illegal and must result in the TDZ error.

OK, that line of reasoning does make some sense, I must admit.

Who `let` the TDZ Out?

But that's just `const`. What about `let`?

Well, TC39 made the decision: since we need a TDZ for `const`, we might as well have a TDZ for `let` as well. *In fact, if we make let have a TDZ, then we discourage all that ugly variable hoisting people do.* So there was a consistency perspective and, perhaps, a bit of social engineering to shift developers' behavior.

My counter-argument would be: if you're favoring consistency, be consistent with `var` instead of `const`; `let` is definitely more like `var` than `const`. That's especially true since they had already chosen consistency with `var` for the whole hoisting-to-the-top-of-the-scope thing. Let `const` be its own unique deal with a TDZ, and let the answer to TDZ purely be: just avoid the TDZ by always declaring your constants at the top of the scope. I think this would have been more reasonable.

But alas, that's not how it landed. `let` has a TDZ because `const` needs a TDZ, because `let` and `const` mimic `var` in their hoisting to the top of the (block) scope. There ya go. Too circular? Read it again a few times.

Are Synchronous Callbacks Still Closures?

Chapter 7 presented two different models for tackling closure:

- Closure is a function instance remembering its outer variables even as that function is passed around and **invoked in** other scopes.
- Closure is a function instance and its scope environment being preserved in-place while any references to it are passed around and **invoked from** other scopes.

These models are not wildly divergent, but they do approach from a different perspective. And that different perspective changes what we identify as a closure.

Don't get lost following this rabbit trail through closures and callbacks:

- Calling back to what (or where)?
- Maybe "synchronous callback" isn't the best label
- **IIF** functions don't move around, why would they need closure?
- Deferring over time is key to closure

What is a Callback?

Before we revisit closure, let me spend a brief moment addressing the word "callback." It's a generally accepted norm that saying "callback" is synonymous with both *asynchronous callbacks* and *synchronous callbacks*. I don't think I agree that this is a good idea, so I want to explain why and propose we move away from that to another term.

Let's first consider an *asynchronous callback*, a function reference that will be invoked at some future *later* point. What does "callback" mean, in this case?

It means that the current code has finished or paused, suspended itself, and that when the function in question is invoked later, execution is entering back into the suspended program, resuming it. Specifically, the point of re-entry is the code that was wrapped in the function reference:

```
setTimeout(function waitForASecond(){
  // this is where JS should call back into
  // the program when the timer has elapsed
}, 1000);

// this is where the current program finishes
// or suspends
```

In this context, "calling back" makes a lot of sense. The JS engine is resuming our suspended program by *calling back in* at a specific location. OK, so a callback is asynchronous.

Synchronous Callback?

But what about *synchronous callbacks*? Consider:

```
function getLabels(studentIDs) {
  return studentIDs.map(
    function formatIDLabel(id){
      return `Student ID: ${String(id).padStart(6)}
    }`;
  );
}

getLabels([ 14, 73, 112, 6 ]);
// [
//   "Student ID: 000014",
//   "Student ID: 000073",
//   "Student ID: 000112",
//   "Student ID: 000006"
// ]
```

Should we refer to `formatIDLabel(..)` as a callback? Is the `map(..)` utility really *calling back* into our program by invoking the function we provided?

There's nothing to *call back into* per se, because the program hasn't paused or exited. We're passing a function (reference) from one part of the program to another part of the program, and then it's immediately invoked.

There's other established terms that might match what we're doing—passing in a function (reference) so that another part of the program can invoke it on our behalf. You might think of this as *Dependency Injection* (DI) or *Inversion of Control* (IoC).

DI can be summarized as passing in necessary part(s) of functionality to another part of the program so that it can invoke them to complete its work. That's a decent description for the `map(..)` call above, isn't it? The `map(..)` utility knows to iterate over the list's values, but it doesn't know what to *do* with those values. That's why we pass it the `formatIDLabel(..)` function. We pass in the dependency.

IoC is a pretty similar, related concept. Inversion of control means that instead of the current area of your program controlling what's happening, you hand control off to another part of the program. We wrapped the logic for computing a label string in the function `formatIDLabel(..)`, then handed invocation control to the `map(..)` utility.

Notably, Martin Fowler cites IoC as the difference between a framework and a library: with a library, you call its functions; with a framework, it calls your

functions.¹

In the context of our discussion, either DI or IoC could work as an alternative label for a *synchronous callback*.

But I have a different suggestion. Let's refer to (the functions formerly known as) *synchronous callbacks*, as *inter-invoked functions* (IIFs). Yes, exactly, I'm playing off IIFEs. These kinds of functions are *inter-invoked*, meaning: another entity invokes them, as opposed to IIFEs, which invoke themselves immediately.

What's the relationship between an *asynchronous callback* and an IIF? An *asynchronous callback* is an IIF that's invoked asynchronously instead of synchronously.

Synchronous Closure?

Now that we've re-labeled *synchronous callbacks* as IIFs, we can return to our main question: are IIFs an example of closure? Obviously, the IIF would have to reference variable(s) from an outer scope for it to have any chance of being a closure. The `formatIDLabel(..)` IIF from earlier does not reference any variables outside its own scope, so it's definitely not a closure.

What about an IIF that does have external references, is that closure?

```
function printLabels(labels) {
  var list = document.getElementById("labelsList");

  labels.forEach(
    function renderLabel(label){
      var li = document.createElement("li");
      li.innerText = label;
      list.appendChild(li);
    }
  );
}
```

The inner `renderLabel(..)` IIF references `list` from the enclosing scope, so it's an IIF that *could* have closure. But here's where the definition/model we choose for closure matters:

- If `renderLabel(..)` is a **function that gets passed somewhere else**, and that function is then invoked, then yes, `renderLabel(..)` is exercising a closure, because closure is what preserved its access to its original scope chain.
- But if, as in the alternative conceptual model from Chapter 7, `renderLabel(..)` stays in place, and only a reference to it is passed to `forEach(..)`, is there any need for closure to preserve the scope chain of

¹*Inversion of Control*, Martin Fowler, <https://martinfowler.com/bliki/InversionOfControl.html>, 26 June 2005.

`renderLabel(..)`, while it executes synchronously right inside its own scope?

No. That's just normal lexical scope.

To understand why, consider this alternative form of `printLabels(..)`:

```
function printLabels(labels) {
  var list = document.getElementById("labelsList");

  for (let label of labels) {
    // just a normal function call in its own
    // scope, right? That's not really closure!
    renderLabel(label);
  }

  // *****

  function renderLabel(label) {
    var li = document.createElement("li");
    li.innerText = label;
    list.appendChild(li);
  }
}
```

These two versions of `printLabels(..)` are essentially the same.

The latter one is definitely not an example of closure, at least not in any useful or observable sense. It's just lexical scope. The former version, with `forEach(..)` calling our function reference, is essentially the same thing. It's also not closure, but rather just a plain ol' lexical scope function call.

Defer to Closure

By the way, Chapter 7 briefly mentioned partial application and currying (which *do* rely on closure!). This is an interesting scenario where manual currying can be used:

```
function printLabels(labels) {
  var list = document.getElementById("labelsList");
  var renderLabel = renderTo(list);

  // definitely closure this time!
  labels.forEach( renderLabel );

  // *****

  function renderTo(list) {
    return function createLabel(label){
```

```

        var li = document.createElement("li");
        li.innerText = label;
        list.appendChild(li);
    };
}
}

```

The inner function `createLabel(..)`, which we assign to `renderLabel`, is closed over `list`, so closure is definitely being utilized.

Closure allows us to remember `list` for later, while we defer execution of the actual label-creation logic from the `renderTo(..)` call to the subsequent `forEach(..)` invocations of the `createLabel(..)` IIF. That may only be a brief moment here, but any amount of time could pass, as closure bridges from call to call.

Classic Module Variations

Chapter 8 explained the classic module pattern, which can look like this:

```

var StudentList = (function defineModule(Student){
    var elems = [];

    var publicAPI = {
        renderList() {
            // ..
        }
    };

    return publicAPI;
})(Student);

```

Notice that we're passing `Student` (another module instance) in as a dependency. But there's lots of useful variations on this module form you may encounter. Some hints for recognizing these variations:

- Does the module know about its own API?
- Even if we use a fancy module loader, it's just a classic module
- Some modules need to work universally

Where's My API?

First, most classic modules don't define and use a `publicAPI` the way I have shown in this code. Instead, they typically look like:

```

var StudentList = (function defineModule(Student){
    var elems = [];

```

```

    return {
      renderList() {
        // ...
      }
    };
  })(Student);

```

The only difference here is directly returning the object that serves as the public API for the module, as opposed to first saving it to an inner `publicAPI` variable. This is by far how most classic modules are defined.

But I strongly prefer, and always use myself, the former `publicAPI` form. Two reasons:

- `publicAPI` is a semantic descriptor that aids readability by making it more obvious what the purpose of the object is.
- Storing an inner `publicAPI` variable that references the same external public API object returned, can be useful if you need to access or modify the API during the lifetime of the module.

For example, you may want to call one of the publicly exposed functions, from inside the module. Or, you may want to add or remove methods depending on certain conditions, or update the value of an exposed property.

Whatever the case may be, it just seems rather silly to me that we *wouldn't* maintain a reference to access our own API. Right?

Asynchronous Module Definition (AMD)

Another variation on the classic module form is AMD-style modules (popular several years back), such as those supported by the RequireJS utility:

```

define([ "./Student" ],function StudentList(Student){
  var elems = [];

  return {
    renderList() {
      // ...
    }
  };
});

```

If you look closely at `StudentList(..)`, it's a classic module factory function. Inside the machinery of `define(..)` (provided by RequireJS), the `StudentList(..)` function is executed, passing to it any other module instances declared as dependencies. The return value is an object representing the public API for the module.

This is based on exactly the same principles (including how the closure works!) as we explored with classic modules.

Universal Modules (UMD)

The final variation we'll look at is UMD, which is less a specific, exact format and more a collection of very similar formats. It was designed to create better interop (without any build-tool conversion) for modules that may be loaded in browsers, by AMD-style loaders, or in Node. I personally still publish many of my utility libraries using a form of UMD.

Here's the typical structure of a UMD:

```
(function UMD(name, context, definition){
    // loaded by an AMD-style loader?
    if (
        typeof define === "function" &&
        define.amd
    ) {
        define(definition);
    }
    // in Node?
    else if (
        typeof module !== "undefined" &&
        module.exports
    ) {
        module.exports = definition(name, context);
    }
    // assume standalone browser script
    else {
        context[name] = definition(name, context);
    }
})( "StudentList", this, function DEF(name, context){

    var elems = [];

    return {
        renderList() {
            // ...
        }
    };
});
```

Though it may look a bit unusual, UMD is really just an IIFE.

What's different is that the main `function` expression part (at the top) of the IIFE contains a series of `if..else if` statements to detect which of the three

supported environments the module is being loaded in.

The final () that normally invokes an IIFE is being passed three arguments: "StudentsList", `this`, and another `function` expression. If you match those arguments to their parameters, you'll see they are: `name`, `context`, and `definition`, respectively. "StudentList" (`name`) is the name label for the module, primarily in case it's defined as a global variable. `this` (`context`) is generally the `window` (aka, global object; see Chapter 4) for defining the module by its name.

`definition(..)` is invoked to actually retrieve the definition of the module, and you'll notice that, sure enough, that's just a classic module form!

There's no question that as of the time of this writing, ESM (ES Modules) are becoming popular and widespread rapidly. But with millions and millions of modules written over the last 20 years, all using some pre-ESM variation of classic modules, they're still very important to be able to read and understand when you come across them.

You Don't Know JS Yet: Scope & Closures - 2nd Edition

Appendix B: Practice

This appendix aims to give you some challenging and interesting exercises to test and solidify your understanding of the main topics from this book. It's a good idea to try out the exercises yourself—in an actual code editor!—instead of skipping straight to the solutions at the end. No cheating!

These exercises don't have a specific right answer that you have to get exactly. Your approach may differ some (or a lot!) from the solutions presented, and that's OK.

There's no judging you on how you write your code. My hope is that you come away from this book feeling confident that you can tackle these sorts of coding tasks built on a strong foundation of knowledge. That's the only objective, here. If you're happy with your code, I am, too!

Buckets of Marbles

Remember Figure 2 from back in Chapter 2?

Fig. 2 (Ch. 2): Colored Scope Bubbles

This exercise asks you to write a program—any program!—that contains nested functions and block scopes, which satisfies these constraints:

- If you color all the scopes (including the global scope!) different colors, you need at least six colors. Make sure to add a code comment labeling each scope with its color.

BONUS: identify any implied scopes your code may have.

- Each scope has at least one identifier.
- Contains at least two function scopes and at least two block scopes.
- At least one variable from an outer scope must be shadowed by a nested scope variable (see Chapter 3).
- At least one variable reference must resolve to a variable declaration at least two levels higher in the scope chain.

TIP:

You *can* just write junk foo/bar/baz-type code for this exercise, but I suggest you try to come up with some sort of non-trivial real'ish code that at least does something kind of reasonable.

Try the exercise for yourself, then check out the suggested solution at the end of this appendix.

Closure (PART 1)

Let's first practice closure with some common computer-math operations: determining if a value is prime (has no divisors other than 1 and itself), and generating a list of prime factors (divisors) for a given number.

For example:

```
isPrime(11);          // true
isPrime(12);          // false

factorize(11);        // [ 11 ]
factorize(12);        // [ 3, 2, 2 ] --> 3*2*2=12
```

Here's an implementation of `isPrime(..)`, adapted from the `Math.js` library:¹

```
function isPrime(v) {
  if (v <= 3) {
    return v > 1;
  }
  if (v % 2 == 0 || v % 3 == 0) {
    return false;
  }
  var vSqrt = Math.sqrt(v);
  for (let i = 5; i <= vSqrt; i += 6) {
    if (v % i == 0 || v % (i + 2) == 0) {
      return false;
    }
  }
  return true;
}
```

And here's a somewhat basic implementation of `factorize(..)` (not to be confused with `factorial(..)` from Chapter 6):

```
function factorize(v) {
  if (!isPrime(v)) {
    let i = Math.floor(Math.sqrt(v));
    while (v % i != 0) {
      i--;
    }
    return [
      ...factorize(i),
```

¹ *Math.js: isPrime(..)*, <https://github.com/josdejong/mathjs/blob/develop/src/function/utils/isPrime.js>, 3 March 2020.

```

        ...factorize(v / i)
    ];
}
return [v];
}

```

NOTE:

I call this basic because it's not optimized for performance. It's binary-recursive (which isn't tail-call optimizable), and it creates a lot of intermediate array copies. It also doesn't order the discovered factors in any way. There are many, many other algorithms for this task, but I wanted to use something short and roughly understandable for our exercise.

If you were to call `isPrime(4327)` multiple times in a program, you can see that it would go through all its dozens of comparison/computation steps every time. If you consider `factorize(..)`, it's calling `isPrime(..)` many times as it computes the list of factors. And there's a good chance most of those calls are repeats. That's a lot of wasted work!

The first part of this exercise is to use closure to implement a cache to remember the results of `isPrime(..)`, so that the primality (`true` or `false`) of a given number is only ever computed once. Hint: we already showed this sort of caching in Chapter 6 with `factorial(..)`.

If you look at `factorize(..)`, it's implemented with recursion, meaning it calls itself repeatedly. That again means we may likely see a lot of wasted calls to compute prime factors for the same number. So the second part of the exercise is to use the same closure cache technique for `factorize(..)`.

Use separate closures for caching of `isPrime(..)` and `factorize(..)`, rather than putting them inside a single scope.

Try the exercise for yourself, then check out the suggested solution at the end of this appendix.

A Word About Memory

I want to share a little quick note about this closure cache technique and the impacts it has on your application's performance.

We can see that in saving the repeated calls, we improve computation speed (in some cases, by a dramatic amount). But this usage of closure is making an explicit trade-off that you should be very aware of.

The trade-off is memory. We're essentially growing our cache (in memory) unboundedly. If the functions in question were called many millions of times with mostly unique inputs, we'd be chewing up a lot of memory. This can definitely

be worth the expense, but only if we think it's likely we see repetition of common inputs so that we're taking advantage of the cache.

If most every call will have a unique input, and the cache is essentially never *used* to any benefit, this is an inappropriate technique to employ.

It also might be a good idea to have a more sophisticated caching approach, such as an LRU (least recently used) cache, that limits its size; as it runs up to the limit, an LRU evicts the values that are... well, least recently used!

The downside here is that LRU is quite non-trivial in its own right. You'll want to use a highly optimized implementation of LRU, and be keenly aware of all the trade-offs at play.

Closure (PART 2)

In this exercise, we're going to again practice closure by defining a `toggle(..)` utility that gives us a value toggler.

You will pass one or more values (as arguments) into `toggle(..)`, and get back a function. That returned function will alternate/rotate between all the passed-in values in order, one at a time, as it's called repeatedly.

```
function toggle(/* .. */) {
  // ..
}

var hello = toggle("hello");
var onOff = toggle("on", "off");
var speed = toggle("slow", "medium", "fast");

hello();      // "hello"
hello();      // "hello"

onOff();      // "on"
onOff();      // "off"
onOff();      // "on"

speed();      // "slow"
speed();      // "medium"
speed();      // "fast"
speed();      // "slow"
```

The corner case of passing in no values to `toggle(..)` is not very important; such a toggler instance could just always return `undefined`.

Try the exercise for yourself, then check out the suggested solution at the end of this appendix.

Closure (PART 3)

In this third and final exercise on closure, we're going to implement a basic calculator. The `calculator()` function will produce an instance of a calculator that maintains its own state, in the form of a function (`calc(..)`, below):

```
function calculator() {  
  // ..  
}  
  
var calc = calculator();
```

Each time `calc(..)` is called, you'll pass in a single character that represents a keypress of a calculator button. To keep things more straightforward, we'll restrict our calculator to supporting entering only digits (0-9), arithmetic operations (+, -, *, /), and “=” to compute the operation. Operations are processed strictly in the order entered; there's no “()” grouping or operator precedence.

We don't support entering decimals, but the divide operation can result in them. We don't support entering negative numbers, but the “-” operation can result in them. So, you should be able to produce any negative or decimal number by first entering an operation to compute it. You can then keep computing with that value.

The return of `calc(..)` calls should mimic what would be shown on a real calculator, like reflecting what was just pressed, or computing the total when pressing “=”.

For example:

```
calc("4");    // 4  
calc("+");   // +  
calc("7");   // 7  
calc("3");   // 3  
calc("-");   // -  
calc("2");   // 2  
calc("=");   // 75  
calc("*");   // *  
calc("4");   // 4  
calc("=");   // 300  
calc("5");   // 5  
calc("-");   // -  
calc("5");   // 5  
calc("=");   // 0
```

Since this usage is a bit clumsy, here's a `useCalc(..)` helper, that runs the calculator with characters one at a time from a string, and computes the display each time:

```
function useCalc(calc,keys) {
```

```

    return [...keys].reduce(
      function showDisplay(display, key){
        var ret = String( calc(key) );
        return (
          display +
          (
            (ret != "" && key == "=") ?
            "=" :
            ""
          ) +
          ret
        );
      },
      ""
    );
  }

useCalc(calc, "4+3=");           // 4+3=7
useCalc(calc, "+9=");           // +9=16
useCalc(calc, "*8=");           // *5=128
useCalc(calc, "7*2*3=");         // 7*2*3=42
useCalc(calc, "1/0=");           // 1/0=ERR
useCalc(calc, "+3=");           // +3=ERR
useCalc(calc, "51=");           // 51

```

The most sensible usage of this `useCalc(..)` helper is to always have “=” be the last character entered.

Some of the formatting of the totals displayed by the calculator require special handling. I’m providing this `formatTotal(..)` function, which your calculator should use whenever it’s going to return a current computed total (after an “=” is entered):

```

function formatTotal(display) {
  if (Number.isFinite(display)) {
    // constrain display to max 11 chars
    let maxDigits = 11;
    // reserve space for "e+" notation?
    if (Math.abs(display) > 9999999999) {
      maxDigits = 6;
    }
    // reserve space for "-"?
    if (display < 0) {
      maxDigits--;
    }
    // whole number?
  }
}

```

```

    if (Number.isInteger(display)) {
        display = display
            .toPrecision(maxDigits)
            .replace(/\.\0+$/,"");
    }
    // decimal
    else {
        // reserve space for "."
        maxDigits--;
        // reserve space for leading "0"?
        if (
            Math.abs(display) >= 0 &&
            Math.abs(display) < 1
        ) {
            maxDigits--;
        }
        display = display
            .toPrecision(maxDigits)
            .replace(/\0+$/,"");
    }
    else {
        display = "ERR";
    }
    return display;
}

```

Don't worry too much about how `formatTotal(..)` works. Most of its logic is a bunch of handling to limit the calculator display to 11 characters max, even if negatives, repeating decimals, or even “e+” exponential notation is required.

Again, don't get too mired in the mud around calculator-specific behavior. Focus on the *memory* of closure.

Try the exercise for yourself, then check out the suggested solution at the end of this appendix.

Modules

This exercise is to convert the calculator from Closure (PART 3) into a module.

We're not adding any additional functionality to the calculator, only changing its interface. Instead of calling a single function `calc(..)`, we'll be calling specific methods on the public API for each “keypress” of our calculator. The outputs stay the same.

This module should be expressed as a classic module factory function called `calculator()`, instead of a singleton IIFE, so that multiple calculators can be

created if desired.

The public API should include the following methods:

- `number(..)` (input: the character/number “pressed”)
- `plus()`
- `minus()`
- `mult()`
- `div()`
- `eq()`

Usage would look like:

```
var calc = calculator();

calc.number("4");      // 4
calc.plus();          // +
calc.number("7");      // 7
calc.number("3");      // 3
calc.minus();          // -
calc.number("2");      // 2
calc.eq();              // 75
```

`formatTotal(..)` remains the same from that previous exercise. But the `useCalc(..)` helper needs to be adjusted to work with the module API:

```
function useCalc(calc,keys) {
  var keyMappings = {
    "+": "plus",
    "-": "minus",
    "*": "mult",
    "/": "div",
    "=": "eq"
  };

  return [...keys].reduce(
    function showDisplay(display,key){
      var fn = keyMappings[key] || "number";
      var ret = String( calc[fn](key) );
      return (
        display +
        (
          (ret != "" && key == "=") ?
          "=" :
          ""
        ) +
        ret
      );
    },
    ""
  );
}
```

```

    ";
}

useCalc(calc, "4+3=");           // 4+3=7
useCalc(calc, "+9=");           // +9=16
useCalc(calc, "*8=");           // *5=128
useCalc(calc, "7*2*3=");         // 7*2*3=42
useCalc(calc, "1/0=");           // 1/0=ERR
useCalc(calc, "+3=");           // +3=ERR
useCalc(calc, "51=");           // 51

```

Try the exercise for yourself, then check out the suggested solution at the end of this appendix.

As you work on this exercise, also spend some time considering the pros/cons of representing the calculator as a module as opposed to the closure-function approach from the previous exercise.

BONUS: write out a few sentences explaining your thoughts.

BONUS #2: try converting your module to other module formats, including: UMD, CommonJS, and ESM (ES Modules).

Suggested Solutions

Hopefully you've tried out the exercises before you're reading this far. No cheating!

Remember, each suggested solution is just one of a bunch of different ways to approach the problems. They're not "the right answer," but they do illustrate a reasonable way to approach each exercise.

The most important benefit you can get from reading these suggested solutions is to compare them to your code and analyze why we each made similar or different choices. Don't get into too much bikeshedding; try to stay focused on the main topic rather than the small details.

Suggested: Buckets of Marbles

The *Buckets of Marbles Exercise* can be solved like this:

```

// RED(1)
const howMany = 100;

// Sieve of Eratosthenes
function findPrimes(howMany) {
    // BLUE(2)
    var sieve = Array(howMany).fill(true);

```

```

var max = Math.sqrt(howMany);

for (let i = 2; i < max; i++) {
  // GREEN(3)
  if (sieve[i]) {
    // ORANGE(4)
    let j = Math.pow(i, 2);
    for (let k = j; k < howMany; k += i) {
      // PURPLE(5)
      sieve[k] = false;
    }
  }
}

return sieve
  .map(function getPrime(flag, prime){
    // PINK(6)
    if (flag) return prime;
    return flag;
  })
  .filter(function onlyPrimes(v){
    // YELLOW(7)
    return !!v;
  })
  .slice(1);
}

findPrimes(howMany);
// [
//   2, 3, 5, 7, 11, 13, 17,
//   19, 23, 29, 31, 37, 41,
//   43, 47, 53, 59, 61, 67,
//   71, 73, 79, 83, 89, 97
// ]

```

Suggested: Closure (PART 1)

The *Closure Exercise (PART 1)* for `isPrime(..)` and `factorize(..)`, can be solved like this:

```

var isPrime = (function isPrime(v){
  var primes = {};

  return function isPrime(v) {
    if (v in primes) {
      return primes[v];
    }
  }
});

```

```

    }
    if (v <= 3) {
        return (primes[v] = v > 1);
    }
    if (v % 2 == 0 || v % 3 == 0) {
        return (primes[v] = false);
    }
    let vSqrt = Math.sqrt(v);
    for (let i = 5; i <= vSqrt; i += 6) {
        if (v % i == 0 || v % (i + 2) == 0) {
            return (primes[v] = false);
        }
    }
    return (primes[v] = true);
};

})();

var factorize = (function factorize(v){
    var factors = {};

    return function findFactors(v) {
        if (v in factors) {
            return factors[v];
        }
        if (!isPrime(v)) {
            let i = Math.floor(Math.sqrt(v));
            while (v % i != 0) {
                i--;
            }
            return (factors[v] = [
                ...findFactors(i),
                ...findFactors(v / i)
            ]);
        }
        return (factors[v] = [v]);
    };
})();

```

The general steps I used for each utility:

1. Wrap an IIFE to define the scope for the cache variable to reside.
2. In the underlying call, first check the cache, and if a result is already known, return.
3. At each place where a **return** was happening originally, assign to the cache and just return the results of that assignment operation—this is a space savings trick mostly just for brevity in the book.

I also renamed the inner function from `factorize(..)` to `findFactors(..)`. That's not technically necessary, but it helps it make clearer which function the recursive calls invoke.

Suggested: Closure (PART 2)

The *Closure Exercise (PART 2)* `toggle(..)` can be solved like this:

```
function toggle(...vals) {
  var unset = {};
  var cur = unset;

  return function next(){
    // save previous value back at
    // the end of the list
    if (cur != unset) {
      vals.push(cur);
    }
    cur = vals.shift();
    return cur;
  };
}

var hello = toggle("hello");
var onOff = toggle("on", "off");
var speed = toggle("slow", "medium", "fast");

hello();      // "hello"
hello();      // "hello"

onOff();      // "on"
onOff();      // "off"
onOff();      // "on"

speed();      // "slow"
speed();      // "medium"
speed();      // "fast"
speed();      // "slow"
```

Suggested: Closure (PART 3)

The *Closure Exercise (PART 3)* `calculator()` can be solved like this:

```
// from earlier:
//
// function useCalc(..) { .. }
// function formatTotal(..) { .. }
```

```

function calculator() {
  var currentTotal = 0;
  var currentVal = "";
  var currentOper = "=";

  return pressKey;

  // ****

  function pressKey(key){
    // number key?
    if (/^\d/.test(key)) {
      currentVal += key;
      return key;
    }
    // operator key?
    else if (/[+*/-]/.test(key)) {
      // multiple operations in a series?
      if (
        currentOper != "=" &&
        currentVal != ""
      ) {
        // implied '=' keypress
        pressKey("=");
      }
      else if (currentVal != "") {
        currentTotal = Number(currentVal);
      }
      currentOper = key;
      currentVal = "";
      return key;
    }
    // = key?
    else if (
      key == "=" &&
      currentOper != "="
    ) {
      currentTotal = op(
        currentTotal,
        currentOper,
        Number(currentVal)
      );
      currentOper = "=";
      currentVal = "";
      return formatTotal(currentTotal);
    }
  }
}

```

```

        }
        return "";
    };

    function op(val1,oper,val2) {
        var ops = {
            // NOTE: using arrow functions
            // only for brevity in the book
            "+": (v1,v2) => v1 + v2,
            "-": (v1,v2) => v1 - v2,
            "*": (v1,v2) => v1 * v2,
            "/": (v1,v2) => v1 / v2
        };
        return ops[oper](val1,val2);
    }
}

var calc = calculator();

useCalc(calc,"4+3=");           // 4+3=7
useCalc(calc,"+9=");           // +9=16
useCalc(calc,"*8=");           // *5=128
useCalc(calc,"7*2*3=");        // 7*2*3=42
useCalc(calc,"1/0=");          // 1/0=ERR
useCalc(calc,"+3=");           // +3=ERR
useCalc(calc,"51=");           // 51

```

NOTE:

Remember: this exercise is about closure. Don't focus too much on the actual mechanics of a calculator, but rather on whether you are properly *remembering* the calculator state across function calls.

Suggested: Modules

The *Modules Exercise* `calculator()` can be solved like this:

```

// from earlier:
//
// function useCalc(..) { .. }
// function formatTotal(..) { .. }

function calculator() {
    var currentTotal = 0;
    var currentVal = "";
    var currentOper = "=";

```

```

var publicAPI = {
  number,
  eq,
  plus() { return operator("+"); },
  minus() { return operator("-"); },
  mult() { return operator("*"); },
  div() { return operator("/"); }
};

return publicAPI;

// *****

function number(key) {
  // number key?
  if (/^\d/.test(key)) {
    currentVal += key;
    return key;
  }
}

function eq() {
  // = key?
  if (currentOper != "=") {
    currentTotal = op(
      currentTotal,
      currentOper,
      Number(currentVal)
    );
    currentOper = "=";
    currentVal = "";
    return formatTotal(currentTotal);
  }
  return "";
}

function operator(key) {
  // multiple operations in a series?
  if (
    currentOper != "=" &&
    currentVal != ""
  ) {
    // implied '=' keypress
    eq();
  }
}

```

```

        else if (currentVal != "") {
            currentTotal = Number(currentVal);
        }
        currentOper = key;
        currentVal = "";
        return key;
    }

    function op(val1,oper,val2) {
        var ops = {
            // NOTE: using arrow functions
            // only for brevity in the book
            "+": (v1,v2) => v1 + v2,
            "-": (v1,v2) => v1 - v2,
            "*": (v1,v2) => v1 * v2,
            "/": (v1,v2) => v1 / v2
        };
        return ops[oper](val1,val2);
    }
}

var calc = calculator();

useCalc(calc,"4+3=");           // 4+3=7
useCalc(calc,"+9=");           // +9=16
useCalc(calc,"*8=");           // *5=128
useCalc(calc,"7*2*3=");         // 7*2*3=42
useCalc(calc,"1/0=");           // 1/0=ERR
useCalc(calc,"+3=");           // +3=ERR
useCalc(calc,"51=");           // 51

```

That's it for this book, congratulations on your achievement! When you're ready, move on to Book 3, *Objects & Classes*.

You Don't Know JS Yet: Objects & Classes - 2nd Edition

Foreword

Kyle Simpson has an unmatched brilliance in the art of explanation.

In April of 2015 I was honored to author the foreword for Kyle's book You Don't Know JS: ES6 & Beyond, an exciting and detailed deep-dive into new JavaScript language features that hadn't even yet been ratified by Ecma (that wouldn't happen until June 2015), but had already made their way out into the world. At the time, there was gap in meaningful documentation and educational resources, and I declared that no better person had stepped up to fill it. I stand by everything I wrote in 2015, and I'm here in 2022 to double down for You Don't Know JS Yet: Objects & Classes.

We are all better off for the time, effort and resources that Kyle pours into forming a better understanding of the JavaScript programming language for himself, and for the betterment of his peers: learning, honing and distilling complex semantics into easily digestable concepts that can be built upon in infinite, myriad ways. And that's exactly what we witness here: Kyle becomes an expert in programming subject matter by consuming it at every level. By probing the semantics of new language features, luring them out from the shadows, releasing them from arcane origins and freeing them for our consumption and growth.

Even as a successful professional software engineer, I keep Kyle's works close at hand. At times these tomes are helpful in explaining deeply complex concepts to teammates and peers, while other times they serve as refreshers for myself, because there's always some interesting take-away or new perspective to gain.

Rick Waldron (@rwaldron) Lead Software Engineer Lightning Web Security
Salesforce

You Don't Know JS Yet - 2nd Edition

Preface

Welcome to the 2nd edition of the widely acclaimed *You Don't Know JS* (**YDKJS**) book series: *You Don't Know JS Yet* (**YDKJSY**).

If you've read any of the 1st edition books, you can expect a refreshed approach in these new ones, with plenty of updated coverage of what's changed in JS over the last five years. But what I hope and believe you'll still *get* is the same commitment to respecting JS and digging into what really makes it tick.

If this is your first time reading these books, I'm glad you're here. Prepare for a deep and extensive journey into all the corners of JavaScript.

If you are new to programming or JS, be aware that these books are not intended as a gentle "intro to JavaScript." This material is, at times, complex and challenging, and goes much deeper than is typical for a first-time learner. You're welcome here no matter what your background is, but these books are written assuming you're already comfortable with JS and have at least 6–9 months experience with it.

The Parts

These books approach JavaScript intentionally opposite of how *The Good Parts* treats the language. No, that doesn't mean we're looking at *the bad parts*, but rather, exploring **all the parts**.

You may have been told, or felt yourself, that JS is a deeply flawed language that was poorly designed and inconsistently implemented. Many have asserted that it's the worst most popular language in the world; that nobody writes JS because they want to, only because they have to given its place at the center of the web. That's a ridiculous, unhealthy, and wholly condescending claim.

Millions of developers write JavaScript every day, and many of them appreciate and respect the language.

Like any great language, it has its brilliant parts as well as its scars. Even the creator of JavaScript himself, Brendan Eich, laments some of those parts as mistakes. But he's wrong: they weren't mistakes at all. JS is what it is today—the world's most ubiquitous and thus most influential programming language—precisely because of *all those parts*.

Don't buy the lie that you should only learn and use a small collection of *good parts* while avoiding all the bad stuff. Don't buy the "X is the new Y" snake oil, that some new feature of the language instantly relegates all usage of a previous feature as obsolete and ignorant. Don't listen when someone says your code isn't "modern" because it isn't yet using a stage-0 feature that was only proposed a few weeks ago!

Every part of JS is useful. Some parts are more useful than others. Some parts require you to be more careful and intentional.

I find it absurd to try to be a truly effective JavaScript developer while only using a small sliver of what the language has to offer. Can you imagine a construction worker with a toolbox full of tools, who only uses their hammer and scoffs at the screwdriver or tape measure as inferior? That's just silly.

My unreserved claim is that you should go about learning all parts of JavaScript, and where appropriate, use them! And if I may be so bold as to suggest: it's time to discard any JS books that tell you otherwise.

The Title?

So what's the title of the series all about?

I'm not trying to insult you with criticism about your current lack of knowledge or understanding of JavaScript. I'm not suggesting you can't or won't be able to learn JavaScript. I'm not boasting about secret advanced insider wisdom that I and only a select few possess.

Seriously, all those were real reactions to the original series title before folks even read the books. And they're baseless.

The primary point of the title “You Don’t Know JS Yet” is to point out that most JS developers don’t take the time to really understand how the code that they write works. They know *that* it works—that it produces a desired outcome. But they either don’t understand exactly *how*, or worse, they have an inaccurate mental model for the *how* that falters on closer scrutiny.

I'm presenting a gentle but earnest challenge to you the reader, to set aside the assumptions you have about JS, and approach it with fresh eyes and an invigorated curiosity that leads you to ask *why* for every line of code you write. Why does it do what it does? Why is one way better or more appropriate than the other half-dozen ways you could have accomplished it? Why do all the “popular kids” say to do X with your code, but it turns out that Y might be a better choice?

I added “Yet” to the title, not only because it's the second edition, but because ultimately I want these books to challenge you in a hopeful rather than discouraging way.

But let me be clear: I don't think it's possible to ever fully *know* JS. That's not an achievement to be obtained, but a goal to strive after. You don't finish knowing everything about JS, you just keep learning more and more as you spend more time with the language. And the deeper you go, the more you revisit what you *knew* before, and you re-learn it from that more experienced perspective.

I encourage you to adopt a mindset around JavaScript, and indeed all of software

development, that you will never fully have mastered it, but that you can and should keep working to get closer to that end, a journey that will stretch for the entirety of your software development career, and beyond.

You can always know JS better than you currently do. That's what I hope these YDKJSY books represent.

The Mission

The case doesn't really need to be made for why developers should take JS seriously—I think it's already more than proven worthy of first-class status among the world's programming languages.

But a different, more important case still needs to be made, and these books rise to that challenge.

I've taught more than 5,000 developers from teams and companies all over the world, in more than 25 countries on six continents. And what I've seen is that far too often, what *counts* is generally just the result of the program, not how the program is written or how/why it works.

My experience not only as a developer but in teaching many other developers tells me: you will always be more effective in your development work if you more completely understand how your code works than you are solely *just* getting it to produce a desired outcome.

In other words, *good enough to work* is not, and should not be, *good enough*.

All developers regularly struggle with some piece of code not working correctly, and they can't figure out why. But far too often, JS developers will blame this on the language rather than admitting it's their own understanding that is falling short. These books serve as both the question and answer: why did it do *this*, and here's how to get it to do *that* instead.

My mission with YDKJSY is to empower every single JS developer to fully own the code they write, to understand it and to write with intention and clarity.

The Path

Some of you have started reading this book with the goal of completing all six books, back to back.

I would like to caution you to consider changing that plan.

It is not my intention that YDKJSY be read straight through. The material in these books is dense, because JavaScript is powerful, sophisticated, and in parts rather complex. Nobody can really hope to *download* all this information to their brains in a single pass and retain any significant amount of it. That's unreasonable, and it's foolish to try.

My suggestion is you take your time going through YDKJSY. Take one chapter, read it completely through start to finish, and then go back and re-read it section by section. Stop in between each section, and practice the code or ideas from that section. For larger concepts, it probably is a good idea to expect to spend several days digesting, re-reading, practicing, then digesting some more.

You could spend a week or two on each chapter, and a month or two on each book, and a year or more on the whole series, and you would still not be squeezing every ounce of YDKJSY out.

Don't binge these books; be patient and spread out your reading. Interleave reading with lots of practice on real code in your job or on projects you participate in. Wrestle with the opinions I've presented along the way, debate with others, and most of all, disagree with me! Run a study group or book club. Teach mini-workshops at your office. Write blog posts on what you've learned. Speak about these topics at local JS meetups.

It's never my goal to convince you to agree with my opinion, but to encourage you to own and be able to defend your opinions. You can't get *there* with an expedient read-through of these books. That's something that takes a long while to emerge, little by little, as you study and ponder and re-visit.

These books are meant to be a field-guide on your wanderings through JavaScript, from wherever you currently are with the language, to a place of deeper understanding. And the deeper you understand JS, the more questions you will ask and the more you will have to explore! That's what I find so exciting!

I'm so glad you're embarking on this journey, and I am so honored you would consider and consult these books along the way. It's time to start *getting to know JS*.

You Don't Know JS Yet: Objects & Classes - 2nd Edition

Chapter 1: Object Foundations

NOTE:

Work in progress

Everything in JS is an object.

This is one of the most pervasive, but most incorrect, “facts” that perpetually circulates about JS. Let the myth busting commence.

JS definitely has objects, but that doesn’t mean that all values are objects. Nevertheless, objects are arguably the most important (and varied!) value type in the language, so mastering them is critical to your JS journey.

The object mechanism is certainly the most flexible and powerful container type – something you put other values into; every JS program you write will use them in one way or another. But that’s not why objects deserve top billing for this book. Objects are the foundation for the second of JS’s three pillars: the prototype.

Why are prototypes (along with the `this` keyword, covered later in the book) so core to JS as to be one of its three pillars? Among other things, prototypes are how JS’s object system can express the class design pattern, one of the most widely relied on design patterns in all of programming.

So our journey here will start with objects, build up a complete understanding of prototypes, de-mystify the `this` keyword, and explore the `class` system.

About This Book

Welcome to book 3 in the *You Don't Know JS Yet* series! If you already finished *Get Started* (the first book) and *Scope & Closures* (the second book), you’re in the right spot! If not, before you proceed I encourage you to read those two as foundations before diving into this book.

The first edition of this book is titled, “`this` & Object Prototypes”. In that book, our focus started with the `this` keyword, as it’s arguably one of the most confused topics in all of JS. The book then spent the majority of its time focused on expositing the prototype system and advocating for embrace of the lesser-known “delegation” pattern instead of class designs. At the time of that book’s writing (2014), ES6 would still be almost 2 years to its completion, so I felt the early sketches of the `class` keyword only merited a brief addendum of coverage.

It's quite an understatement to say a lot has changed in the JS landscape in the almost 8 years since that book. ES6 is old news now; at the time of *this* book's writing, JS has seen 7 yearly updates **after ES6** (ES2016 through ES2022).

Now, we still need to talk about how `this` works, and how that relates to methods invoked against various objects. And `class` actually operates (mostly!) via the prototype chain deep under the covers. But JS developers in 2022 are almost never writing code to explicitly wire up prototypal inheritance anymore. And as much as I personally wish differently, class design patterns – not “behavior delegation” – are how the majority of data and behavior organization (data structures) in JS are expressed.

This book reflects JS's current reality: thus the new sub-title, new organization and focus of topics, and complete re-write of the previous edition's text.

Objects As Containers

One common way of gathering up multiple values in a single container is with an object. Objects are collections of key/value pairs. There are also sub-types of object in JS with specialized behaviors, such as arrays (numerically indexed) and even functions (callable); more on these sub-types later.

NOTE:

Keys are often referred to as “property names”, with the pairing of a property name and a value often called a “property”. This book will use those terms distinctly in that manner.

Regular JS objects are typically declared with literal syntax, like this:

```
myObj = {  
  // ...  
};
```

Note: There's an alternate way to create an object (using `myObj = new Object()`), but this is not common or preferred, and is almost never the appropriate way to go about it. Stick with object literal syntax.

It's easy to get confused what pairs of `{ .. }` mean, since JS overloads the curly brackets to mean any of the following, depending on the context used:

- delimit values, like object literals
- define object destructuring patterns (more on this later)
- delimit interpolated string expressions, like ``some ${ getNumber() } thing``
- define blocks, like on `if` and `for` loops
- define function bodies

Though it can sometimes be challenging as you read code, look for whether a { .. } curly brace pair is used in the program where a value/expression is valid to appear; if so, it's an object literal, otherwise it's one of the other overloaded uses.

Defining Properties

Inside the object literal curly braces, you define properties (name and value) with `propertyName: PropertyValue` pairs, like this:

```
myObj = {  
    favoriteNumber: 42,  
    isDeveloper: true,  
    firstName: "Kyle"  
};
```

The values you assign to the properties can be literals, as shown, or can be computed by expression:

```
function twenty() { return 20; }  
  
myObj = {  
    favoriteNumber: (twenty() + 1) * 2,  
};
```

The expression `(twenty() + 1) * 2` is evaluated immediately, with the result (42) assigned as the property value.

Developers sometimes wonder if there's a way to define an expression for a property value where the expression is "lazy", meaning it's not computed at the time of assignment, but defined later. JS does not have lazy expressions, so the only way to do so is for the expression to be wrapped in a function:

```
function twenty() { return 20; }  
function myNumber() { return (twenty() + 1) * 2; }  
  
myObj = {  
    favoriteNumber: myNumber // notice, NOT `myNumber()` as a function call  
};
```

In this case, `favoriteNumber` is not holding a numeric value, but rather a function reference. To compute the result, that function reference must be explicitly executed.

Looks Like JSON?

You may notice that this object-literal syntax we've seen thus far resembles a related syntax, "JSON" (JavaScript Object Notation):

```
{
  "favoriteNumber": 42,
  "isDeveloper": true,
  "firstName": "Kyle"
}
```

The biggest differences between JS's object literals and JSON are, for objects defined as JSON:

1. property names must be quoted with " double-quote characters
2. property values must be literals (either primitives, objects, or arrays), not arbitrary JS expressions

In JS programs, an object literal does not require quoted property names – you *can* quote them (' or " allowed), but it's usually optional. There are however characters that are valid in a property name, but which cannot be included without surrounding quotes; for example, leading numbers or whitespace:

```
myObj = {
  favoriteNumber: 42,
  isDeveloper: true,
  firstName: "Kyle",
  "2 nicknames": [ "getify", "ydkjs" ]
};
```

One other minor difference is, JSON syntax – that is, text that will be *parsed* as JSON, such as from a `.json` file – is stricter than general JS. For example, JS allows comments (`// ..` and `/* .. */`), and trailing `,` commas in object and array expressions; JSON does not allow any of these. Thankfully, JSON does still allow arbitrary whitespace.

Property Names

Property names in object literals are almost always treated/converted as string values. One exception to this is for integer (or “integer looking”) property “names”:

```
anotherObj = {
  42:      "<-- this property name will be treated as an integer",
  "41":    "<-- ...and so will this one",

  true:    "<-- this property name will be treated as a string",
  [myObj]: "<-- ...and so will this one"
};
```

The `42` property name will be treated as an integer property name (aka, index); the `"41"` string value will also be treated as such since it *looks like* an integer. By contrast, the `true` value will become the string property name `"true"`, and

the `myObj` identifier reference, *computed* via the surrounding `[...]`, will coerce the object's value to a string (generally the default `"[object Object]"`).

WARNING:

If you need to actually use an object as a key/property name, never rely on this computed string coercion; its behavior is surprising and almost certainly not what's expected, so program bugs are likely to occur. Instead, use a more specialized data structure, called a `Map` (added in ES6), where objects used as property "names" are left as-is instead of being coerced to a string value.

As with `[myObj]` above, you can *compute* any **property name** (distinct from computing the property value) at the time of object literal definition:

```
anotherObj = {  
  ["x" + (21 * 2)]: true  
};
```

The expression `"x" + (21 * 2)`, which must appear inside of `[...]` brackets, is computed immediately, and the result (`"x42"`) is used as the property name.

Symbols As Property Names

ES6 added a new primitive value type of `Symbol`, which is often used as a special property name for storing and retrieving property values. They're created via the `Symbol(..)` function call (**without** the `new` keyword), which accepts an optional description string used only for friendlier debugging purposes; if specified, the description is inaccessible to the JS program and thus not used for any other purpose than debug output.

```
myPropSymbol = Symbol("optional, developer-friendly description");
```

NOTE:

Symbols are sort of like numbers or strings, except that their value is *opaque* to, and globally unique within, the JS program. In other words, you can create and use symbols, but JS doesn't let you know anything about, or do anything with, the underlying value; that's kept as a hidden implementation detail by the JS engine.

Computed property names, as previously described, are how to define a symbol property name on an object literal:

```
myPropSymbol = Symbol("optional, developer-friendly description");  
  
anotherObj = {  
  [myPropSymbol]: "Hello, symbol!"  
};
```

The computed property name used to define the property on `anotherObj` will be the actual primitive symbol value (whatever it is), not the optional description string ("optional, developer-friendly description").

Because symbols are globally unique in your program, there's **no** chance of accidental collision where one part of the program might accidentally define a property name the same as another part of the program tried defined/assigned.

Symbols are also useful to hook into special default behaviors of objects, and we'll cover that in more detail in "Extending the MOP" in the next chapter.

Concise Properties

When defining an object literal, it's common to use a property name that's the same as an existing in-scope identifier that holds the value you want to assign.

```
coolFact = "the first person convicted of speeding was going 8 mph";  
  
anotherObj = {  
  coolFact: coolFact  
};
```

NOTE:

That would have been the same thing as the quoted property name definition `"coolFact": coolFact`, but JS developers rarely quote property names unless strictly necessary. Indeed, it's idiomatic to avoid the quotes unless required, so it's discouraged to include them unnecessarily.

In this situation, where the property name and value expression identifier are identical, you can omit the property-name portion of the property definition, as a so-called "concise property" definition:

```
coolFact = "the first person convicted of speeding was going 8 mph";  
  
anotherObj = {  
  coolFact  // <-- concise property short-hand  
};
```

The property name is `"coolFact"` (string), and the value assigned to the property is what's in the `coolFact` variable at that moment: `"the first person convicted of speeding was going 8 mph"`.

At first, this shorthand convenience may seem confusing. But as you get more familiar with seeing this very common and popular feature being used, you'll likely favor it for typing (and reading!) less.

Concise Methods

Another similar shorthand is defining functions/methods in an object literal using a more concise form:

```
anotherObj = {
  // standard function property
  greet: function() { console.log("Hello!"); },

  // concise function/method property
  greet2() { console.log("Hello, friend!"); }
};
```

While we're on the topic of concise method properties, we can also define generator functions (another ES6 feature):

```
anotherObj = {
  // instead of:
  // greet3: function*() { yield "Hello, everyone!"; }

  // concise generator method
  *greet3() { yield "Hello, everyone!"; }
};
```

And though it's not particularly common, concise methods/generators can even have quoted or computed names:

```
anotherObj = {
  "greet-4"() { console.log("Hello, audience!"); },

  // concise computed name
  [ "gr" + "eet 5" ]() { console.log("Hello, audience!"); },

  // concise computed generator name
  *[ "ok, greet 6".toUpperCase() ]() { yield "Hello, audience!"; }
};
```

Object Spread

Another way to define properties at object literal creation time is with a form of the `...` syntax – it's not technically an operator, but it certainly seems like one – often referred to as “object spread”.

The `...` when used inside an object literal will “spread” out the contents (properties, aka key/value pairs) of another object value into the object being defined:

```
anotherObj = {
  favoriteNumber: 12,
```

```

...myObj,    // object spread, shallow copies `myObj`

greeting: "Hello!"
}

```

The spreading of `myObj`'s properties is shallow, in that it only copies the top-level properties from `myObj`; any values those properties hold are simply assigned over. If any of those values are references to other objects, the references themselves are assigned (by copy), but the underlying object values are *not* duplicated – so you end up with multiple shared references to the same object(s).

You can think of object spreading like a `for` loop that runs through the properties one at a time and does an `=` style assignment from the source object (`myObj`) to the target object (`anotherObj`).

Also, consider these property definition operations to happen “in order”, from top to bottom of the object literal. In the above snippet, since `myObj` has a `favoriteNumber` property, the object spread will end up overwriting the `favoriteNumber: 12` property assignment from the previous line. Moreover, if `myObj` had contained a `greeting` property that was copied over, the next line (`greeting: "Hello!"`) would override that property definition.

NOTE:

Object spread also only copies *owned* properties (those directly on the object) that are *enumerable* (allowed to be enumerated/listed). It does not duplicate the property – as in, actually mimic the property’s exact characteristics – but rather do a simple assignment style copy. We’ll cover more such details in the “Property Descriptors” section of the next chapter.

A common way `... object` spread is used is for performing *shallow* object duplication:

```
myObjShallowCopy = { ...myObj };
```

Keep in mind you cannot `... spread` into an existing object value; the `... object` spread syntax can only appear inside the `{ ... }` object literal, which is creating a new object value. To perform a similar shallow object copy but with APIs instead of syntax, see the “Object Entries” section later in this chapter (with coverage of `Object.entries(..)` and `Object.fromEntries(..)`).

But if you instead want to copy object properties (shallowly) into an *existing* object, see the “Assigning Properties” section later in this chapter (with coverage of `Object.assign(..)`).

Deep Object Copy

Also, since `...` doesn’t do full, deep object duplication, the object spread is generally only suitable for duplicating objects that hold simple, primitive values

only, not references to other objects.

Deep object duplication is an incredibly complex and nuanced operation. Duplicating a value like 42 is obvious and straightforward, but what does it mean to copy a function (which is a special kind of object, also held by reference), or to copy an external (not entirely in JS) object reference, such as a DOM element? And what happens if an object has circular references (like where a nested descendant object holds a reference back up to an outer ancestor object)? There's a variety of opinions in the wild about how all these corner cases should be handled, and thus no single standard exists for deep object duplication.

For deep object duplication, the standard approaches have been:

1. Use a library utility that declares a specific opinion on how the duplication behaviors/nuances should be handled.
2. Use the `JSON.parse(JSON.stringify(..))` round-trip trick – this only “works” correctly if there are no circular references, and if there are no values in the object that cannot be properly serialized with JSON (such as functions).

Recently, though, a third option has landed. This is not a JS feature, but rather a companion API provided to JS by environments like the web platform. Objects can be deep copied now using `structuredClone(..)`¹.

```
myObjCopy = structuredClone(myObj);
```

The underlying algorithm behind this built-in utility supports duplicating circular references, as well as **many more** types of values than the JSON round-trip trick. However, this algorithm still has its limits, including no support for cloning functions or DOM elements.

Accessing Properties

Property access of an existing object is preferably done with the `.` operator:

```
myObj.favoriteNumber; // 42
myObj.isDeveloper; // true
```

If it's possible to access a property this way, it's strongly suggested to do so.

If the property name contains characters that cannot appear in identifiers, such as leading numbers or whitespace, `[..]` brackets can be used instead of the `..`:

```
myObj["2 nicknames"]; // [ "getify", "ydkjs" ]
anotherObj[42]; // <-- this property name will...
anotherObj["41"]; // <-- this property name will...
```

¹“Structured Clone Algorithm”, HTML Specification; <https://html.spec.whatwg.org/multipage/structured-data.html#structured-cloning> ; Accessed July 2022

Even though numeric property “names” remain as numbers, property access via the [...] brackets will coerce a string representation to a number (e.g., "42" as the 42 numeric equivalent), and then access the associated numeric property accordingly.

Similar to the object literal, the property name to access can be computed via the [...] brackets. The expression can be a simple identifier:

```
propName = "41";
anotherObj[propName];
```

Actually, what you put between the [...] brackets can be any arbitrary JS expression, not just identifiers or literal values like 42 or "isDeveloper". JS will first evaluate the expression, and the resulting value will then be used as the property name to look up on the object:

```
function howMany(x) {
  return x + 1;
}

myObj[`${howMany(1)} nicknames`]; // [ "getify", "ydkjs" ]
```

In this snippet, the expression is a back-tick delimited `template string literal` with an interpolated expression of the function call `howMany(1)`. The overall result of that expression is the string value "2 nicknames", which is then used as the property name to access.

Object Entries

You can get a listing of the properties in an object, as an array of tuples (two-element sub-arrays) holding the property name and value:

```
myObj = {
  favoriteNumber: 42,
  isDeveloper: true,
  firstName: "Kyle"
};

Object.entries(myObj);
// [ ["favoriteNumber", 42], ["isDeveloper", true], ["firstName", "Kyle"] ]
```

Added in ES6, `Object.entries(..)` retrieves this list of entries – containing only owned enumerable properties; see the “Property Descriptors” section in the next chapter – from a source object.

Such a list can be looped/iterated over, potentially assigning properties to another existing object. However, it’s also possible to create a new object from a list of entries, using `Object.fromEntries(..)` (added in ES2019):

```
myObjShallowCopy = Object.fromEntries( Object.entries(myObj) );
```

```
// alternate approach to the earlier discussed:  
// myObjShallowCopy = { ...myObj };
```

Destructuring

Another approach to accessing properties is through object destructuring (added in ES6). Think of destructuring as defining a “pattern” that describes what an object value is supposed to “look like” (structurally), and then asking JS to follow that “pattern” to systematically access the contents of an object value.

The end result of object destructuring is not another object, but rather one or more assignments to other targets (variables, etc) of the values from the source object.

Imagine this sort of pre-ES6 code:

```
myObj = {  
    favoriteNumber: 42,  
    isDeveloper: true,  
    firstName: "Kyle"  
};  
  
const favoriteNumber = (  
    myObj.favoriteNumber !== undefined ? myObj.favoriteNumber : 42  
);  
const isDev = myObj.isDeveloper;  
const firstName = myObj.firstName;  
const lname = (  
    myObj.lastName !== undefined ? myObj.lastName : "--missing--"  
);
```

Those accesses of the property values, and assignments to other identifiers, is generally called “manual destructuring”. To use the declarative object destructuring syntax, it might look like this:

```
myObj = {  
    favoriteNumber: 42,  
    isDeveloper: true,  
    firstName: "Kyle"  
};  
  
const { favoriteNumber = 12 } = myObj;  
const {  
    isDeveloper: isDev,  
    firstName: firstName,  
    lastName: lname = "--missing--"  
} = myObj;
```

```

favoriteNumber;    // 42
isDev;            // true
firstName;        // "Kyle"
lname;            // "--missing--"

```

As shown, the `{ ... }` object destructuring resembles an object literal value definition, but it appears on the left-hand side of the `=` operator rather than on the right-hand side where an object value expression would appear. That makes the `{ ... }` on the left-hand side a destructuring pattern rather than another object definition.

The `{ favoriteNumber } = myObj` destructuring tells JS to find a property named `favoriteNumber` on the object, and to assign its value to an identifier of the same name. The single instance of the `favoriteNumber` identifier in the pattern is similar to “concise properties” as discussed earlier in this chapter: if the source (property name) and target (identifier) are the same, you can omit one of them and only list it once.

The `= 12` part tells JS to provide 12 as a default value for the assignment to `favoriteNumber`, if the source object either doesn’t have a `favoriteNumber` property, or if the property holds an `undefined` value.

In the second destructuring pattern, the `isDeveloper: isDev` pattern is instructing JS to find a property named `isDeveloper` on the source object, and assign its value to an identifier named `isDev`. It’s sort of a “renaming” of the source to the target. By contrast, `firstName: firstName` is providing the source and target for an assignment, but is redundant since they’re identical; a single `firstName` would have sufficed here, and is generally more preferred.

The `lastName: lname = "--missing--"` combines both source-target renaming and a default value (if the `lastName` source property is missing or `undefined`).

The above snippet combines object destructuring with variable declarations – in this example, `const` is used, but `var` and `let` work as well – but it’s not inherently a declaration mechanism. Destructuring is about access and assignment (source to target), so it can operate against existing targets rather than declaring new ones:

```

let fave;

// surrounding ( ) are required syntax here,
// when a declarator is not used
({ favoriteNumber: fave } = myObj);

fave; // 42

```

Object destructuring syntax is generally preferred for its declarative and more readable style, over the heavily imperative pre-ES6 equivalents. But don’t go

overboard with destructuring. Sometimes just doing `x = someObj.x` is perfectly fine!

Conditional Property Access

Recently (in ES2020), a feature known as “optional chaining” was added to JS, which augments property access capabilities (especially nested property access). The primary form is the two-character compound operator `?.`, like `A?.B`.

This operator will check the left-hand side reference (`A`) to see if it’s null’ish (`null` or `undefined`). If so, the rest of the property access expression is short-circuited (skipped), and `undefined` is returned as the result (even if it was `null` that was actually encountered!). Otherwise, `?.` will access the property just as a normal `.` operator would.

For example:

```
myObj?.favoriteNumber
```

Here, the null’ish check is performed against the `myObj`, meaning that the `favoriteNumber` property access is only performed if the value in `myObj` is non-null’ish. Note that it doesn’t verify that `myObj` is actually holding a real object, only that it’s non-nullish. However, all non-nullish values can “safely” (no JS exception) be “accessed” via the `.` operator, even if there’s no matching property to retrieve.

It’s easy to get confused into thinking that the null’ish check is against the `favoriteNumber` property. But one way to keep it straight is to remember that the `?` is on the side where the safety check is performed, while the `.` is on the side that is only conditionally evaluated if the non-null’ish check passes.

Typically, the `?.` operator is used in nested property accesses that may be 3 or more levels deep, such as:

```
myObj?.address?.city
```

The equivalent operation with the `?.` operator would look like this:

```
(myObj != null && myObj.address != null) ? myObj.address.city : undefined
```

Again, remember that no check has been performed against the right-most property (`city`) here.

Also, the `?.` should not universally be used in place of every single `.` operator in your programs. You should endeavor to know if a `.` property access will succeed or not before making the access, whenever possible. Use `?.` only when the nature of the values being accessed is subject to conditions that cannot be predicted/controlled.

For example, in the previous snippet, the `myObj?.` usage is probably mis-guided, because it really shouldn’t be the case that you start a chain of property access

against a variable that might not even hold a top-level object (aside from its contents potentially missing certain properties in certain conditions).

Instead, I would recommend usage more like this:

```
myObj.address?.city
```

And that expression should only be used in part of your program where you're sure that `myObj` is at least holding a valid object (whether or not it has an `address` property with a sub-object in it).

Another form of the “optional chaining” operator is `?.`, which is used when the property access you want to make conditional/safe requires a `[...]` bracket.

```
myObj["2 nicknames"]?.[0]; // "getify"
```

Everything asserted about how `?.` behaves goes the same for `?.`.

WARNING:

There's a third form of this feature, named “optional call”, which uses `?.` as the operator. It's used for performing a non-null’ish check on a property before executing the function value in the property. For example, instead of `myObj.someFunc(42)`, you can do `myObj.someFunc?.(42)`. The `?.` checks to make sure `myObj.someFunc` is non-null’ish before invoking it (with the `(42)` part). While that may sound like a useful feature, I think this is dangerous enough to warrant complete avoidance of this form/construct. My concern is that `?.` makes it seem as if we're ensuring that the function is “callable” before calling it, when in fact we're only checking if it's non-null’ish. Unlike `?.` which can allow a “safe” `. access` against a non-null’ish value that's also not an object, the `?.` non-null’ish check isn't similarly “safe”. If the property in question has any non-null’ish, non-function value in it, like `true` or `"Hello"`, the `(42)` call part will be invoked and yet throw a JS exception. So in other words, this form is unfortunately masquerading as more “safe” than it actually is, and should thus be avoided in essentially all circumstances. If a property value can ever *not be* a function, do a more fullsome check for its function’ness before trying to invoke it. Don't pretend that `?.` is doing that for you, or future readers/maintainers of your code (including your future self!) will likely regret it.

Accessing Properties On Non-Objects

This may sound counter-intuitive, but you can generally access properties/methods from values that aren't themselves objects:

```
fave = 42;  
  
fave;           // 42  
fave.toString(); // "42"
```

Here, `fave` holds a primitive `42` number value. So how can we do `.toString` to access a property from it, and then `()` to invoke the function held in that property?

This is a tremendously more indepth topic than we'll get into in this book; see book 4, "Types & Grammar", of this series for more. However, as a quick glimpse: if you perform a property access (`.` or `[...]`) against a non-object, non-null'ish value, JS will by default (temporarily!) coerce the value into an object-wrapped representation, allowing the property access against that implicitly instantiated object.

This process is typically called "boxing", as in putting a value inside a "box" (object container).

So in the above snippet, just for the moment that `.toString` is being accessed on the `42` value, JS will box this value into a `Number` object, and then perform the property access.

Note that `null` and `undefined` can be object-ified, by calling `Object(null)` / `Object(undefined)`. However, JS does not automatically box these null'ish values, so property access against them will fail (as discussed earlier in the "Conditional Property Access" section).

NOTE:

Boxing has a counterpart: unboxing. For example, the JS engine will take an object wrapper – like a `Number` object wrapped around `42` – created with `Number(42)` or `Object(42)` – and unwrap it to retrieve the underlying primitive `42`, whenever a mathematical operation (like `*` or `-`) encounters such an object. Unboxing behavior is way out of scope for our discussion, but is covered fully in the aforementioned "Types & Grammar" title.

Assigning Properties

Whether a property is defined at the time of object literal definition, or added later, the assignment of a property value is done with the `=` operator, as any other normal assignment would be:

```
myObj.favoriteNumber = 123;
```

If the `favoriteNumber` property doesn't already exist, that statement will create a new property of that name and assign its value. But if it already exists, that statement will re-assign its value.

WARNING:

An `=` assignment to a property may fail (silently or throwing an exception), or it may not directly assign the value but instead invoke a *setter* function that performs some operation(s). More details on these behaviors in the next chapter.

It's also possible to assign one or more properties at once – assuming the source properties (name and value pairs) are in another object – using the `Object.assign(..)` (added in ES6) method:

```
// shallow copy all (owned and enumerable) properties
// from `myObj` into `anotherObj`
Object.assign(anotherObj, myObj);

Object.assign(
  /*target=*/
  anotherObj,
  /*source1=*/
  {
    someProp: "some value",
    anotherProp: 1001,
  },
  /*source2=*/
  {
    yetAnotherProp: false
  }
);
```

`Object.assign(..)` takes the first object as target, and the second (and optionally subsequent) object(s) as source(s). Copying is done in the same manner as described earlier in the “Object Spread” section.

Deleting Properties

Once a property is defined on an object, the only way to remove it is with the `delete` operator:

```
anotherObj = {
  counter: 123
};

anotherObj.counter; // 123

delete anotherObj.counter;

anotherObj.counter; // undefined
```

Contrary to common misconception, the JS `delete` operator does **not** directly do any deallocation/freeing up of memory, through garbage collection (GC).

The only thing it does is remove a property from an object. If the value in the property was a reference (to another object/etc), and there are no other surviving references to that value once the property is removed, that value would likely then be eligible for removal in a future sweep of the GC.

Calling `delete` on anything other than an object property is a misuse of the `delete` operator, and will either fail silently (in non-strict mode) or throw an exception (in strict mode).

Deleting a property from an object is distinct from assigning it a value like `undefined` or `null`. A property assigned `undefined`, either initially or later, is still present on the object, and might still be revealed when enumerating the contents

Determining Container Contents

You can determine an object's contents in a variety of ways. To ask an object if it has a specific property:

```
myObj = {  
    favoriteNumber: 42,  
    coolFact: "the first person convicted of speeding was going 8 mph",  
    beardLength: undefined,  
    nicknames: [ "getify", "ydkjs" ]  
};  
  
"favoriteNumber" in myObj;           // true  
  
myObj.hasOwnProperty("coolFact");    // true  
myObj.hasOwnProperty("beardLength"); // true  
  
myObj.nicknames = undefined;  
myObj.hasOwnProperty("nicknames");   // true  
  
delete myObj.nicknames;  
myObj.hasOwnProperty("nicknames");   // false
```

There *is* an important difference between how the `in` operator and the `hasOwnProperty(..)` method behave. The `in` operator will check not only the target object specified, but if not found there, it will also consult the object's `[[Prototype]]` chain (covered in the next chapter). By contrast, `hasOwnProperty(..)` only consults the target object.

If you're paying close attention, you may have noticed that `myObj` appears to have a method property called `hasOwnProperty(..)` on it, even though we didn't define such. That's because `hasOwnProperty(..)` is defined as a built-in on `Object.prototype`, which by default is "inherited by" all normal objects. There is risk inherent to accessing such an "inherited" method, though. Again,

more on prototypes in the next chapter.

Better Existence Check

ES2022 (almost official at time of writing) has already settled on a new feature, `Object.hasOwn(..)`. It does essentially the same thing as `hasOwnProperty(..)`, but it's invoked as a static helper external to the object value instead of via the object's `[[Prototype]]`, making it safer and more consistent in usage:

```
// instead of:  
myObj.hasOwnProperty("favoriteNumber")  
  
// we should now prefer:  
Object.hasOwn(myObj, "favoriteNumber")
```

Even though (at time of writing) this feature is just now emerging in JS, there are polyfills that make this API available in your programs even when running in a previous JS environment that doesn't yet have the feature defined. For example, a quick stand-in polyfill sketch:

```
// simple polyfill sketch for `Object.hasOwn(..)`  
if (!Object.hasOwn) {  
    Object.hasOwn = function hasOwn(obj, propName) {  
        return Object.prototype.hasOwnProperty.call(obj, propName);  
    };  
}
```

Including a polyfill patch such as that in your program means you can safely start using `Object.hasOwn(..)` for property existence checks no matter whether a JS environment has `Object.hasOwn(..)` built in yet or not.

Listing All Container Contents

We already discussed the `Object.entries(..)` API earlier, which tells us what properties an object has (as long as they're enumerable – more in the next chapter).

There's a variety of other mechanisms available, as well. `Object.keys(..)` gives us list of the enumerable property names (aka, keys) in an object – names only, no values; `Object.values(..)` instead gives us list of all values held in enumerable properties.

But what if we wanted to get *all* the keys in an object (enumerable or not)? `Object.getOwnPropertyNames(..)` seems to do what we want, in that it's like `Object.keys(..)` but also returns non-enumerable property names. However, this list **will not** include any Symbol property names, as those are treated as special locations on the object. `Object.getOwnPropertySymbols(..)` returns

all of an object's Symbol properties. So if you concatenate both of those lists together, you'd have all the direct (*owned*) contents of an object.

Yet as we've implied several times already, and will cover in full detail in the next chapter, an object can also "inherit" contents from its `[[Prototype]]` chain. These are not considered *owned* contents, so they won't show up in any of these lists.

Recall that the `in` operator will potentially traverse the entire chain looking for the existence of a property. Similarly, a `for..in` loop will traverse the chain and list any enumerable (owned or inherited) properties. But there's no built-in API that will traverse the whole chain and return a list of the combined set of both *owned* and *inherited* contents.

Temporary Containers

Using a container to hold multiple values is sometimes just a temporary transport mechanism, such as when you want to pass multiple values to a function via a single argument, or when you want a function to return multiple values:

```
function formatValues({ one, two, three }) {
  // the actual object passed in as an
  // argument is not accessible, since
  // we destructured it into three
  // separate variables

  one = one.toUpperCase();
  two = `--${two}--`;
  three = three.substring(0,5);

  // this object is only to transport
  // all three values in a single
  // return statement
  return { one, two, three };
}

// destructuring the return value from
// the function, because that returned
// object is just a temporary container
// to transport us multiple values
const { one, two, three } =

  // this object argument is a temporary
  // transport for multiple input values
formatValues({
  one: "Kyle",
  two: "Simpson",
```

```

        three: "getify"
    });

one;    // "KYLE"
two;    // "--Simpson--"
three;  // "getif"

```

The object literal passed into `formatValues(..)` is immediately parameter destructured, so inside the function we only deal with three separate variables (`one`, `two`, and `three`). The object literal `returned` from the function is also immediately destructured, so again we only deal with three separate variables (`one`, `two`, `three`).

This snippet illustrates the idiom/pattern that an object is sometimes just a temporary transport container rather than a meaningful value in and of itself.

Containers Are Collections Of Properties

The most common usage of objects is as containers for multiple values. We create and manage property container objects by:

- defining properties (named locations), either at object creation time or later
- assigning values, either at object creation time or later
- accessing values later, using the location names (property names)
- deleting properties via `delete`
- determining container contents with `in`, `hasOwnProperty(..)` / `hasOwn(..)`, `Object.entries(..)` / `Object.keys(..)`, etc

But there's a lot more to objects than just static collections of property names and values. In the next chapter, we'll dive under the hood to look at how they actually work.

You Don't Know JS Yet: Objects & Classes - 2nd Edition

Chapter 2: How Objects Work

NOTE:

Work in progress

Objects are not just containers for multiple values, though clearly that's the context for most interactions with objects.

To fully understand the object mechanism in JS, and get the most out of using objects in our programs, we need to look more closely at a number of characteristics of objects (and their properties) which can affect their behavior when interacting with them.

These characteristics that define the underlying behavior of objects are collectively referred to in formal terms as the “metaobject protocol” (MOP)¹. The MOP is useful not only for understanding how objects will behave, but also for overriding the default behaviors of objects to bend the language to fit our program's needs more fully.

Property Descriptors

Each property on an object is internally described by what's known as a “property descriptor”. This is, itself, an object (aka, “metaobject”) with several properties (aka “attributes”) on it, dictating how the target property behaves.

We can retrieve a property descriptor for any existing property using `Object.getOwnPropertyDescriptor(..)` (ES5):

```
myObj = {  
    favoriteNumber: 42,  
    isDeveloper: true,  
    firstName: "Kyle"  
};  
  
Object.getOwnPropertyDescriptor(myObj, "favoriteNumber");  
// {  
//     value: 42,  
//     enumerable: true,  
//     writable: true,  
//     configurable: true  
// }
```

¹“Metaobject”, Wikipedia; <https://en.wikipedia.org/wiki/Metaobject> ; Accessed July 2022.

We can even use such a descriptor to define a new property on an object, using `Object.defineProperty(..)` (ES5):

```
anotherObj = {};  
  
Object.defineProperty(anotherObj, "fave", {  
    value: 42,  
    enumerable: true,      // default if omitted  
    writable: true,       // default if omitted  
    configurable: true    // default if omitted  
});  
  
anotherObj.fave;           // 42
```

If an existing property has not already been marked as non-configurable (with `configurable: false` in its descriptor), it can always be re-defined/overwritten using `Object.defineProperty(..)`.

WARNING:

A number of earlier sections in this chapter refer to “copying” or “duplicating” properties. One might assume such copying/duplication would be at the property descriptor level. However, none of those operations actually work that way; they all do simple = style access and assignment, which has the effect of ignoring any nuances in how the underlying descriptor for a property is defined.

Though it seems far less common out in the wild, we can even define multiple properties at once, each with their own descriptor:

```
anotherObj = {};  
  
Object.defineProperties(anotherObj, {  
    "fave": {  
        // a property descriptor  
    },  
    "superFave": {  
        // another property descriptor  
    }  
});
```

It’s not very common to see this usage, because it’s rarer that you need to specifically control the definition of multiple properties. But it may be useful in some cases.

Accessor Properties

A property descriptor usually defines a `value` property, as shown above. However, a special kind of property, known as an “accessor property” (aka, a getter/setter), can be defined. For these a property like this, its descriptor does not define a fixed `value` property, but would instead look something like this:

```
{  
  get() { .. },    // function to invoke when retrieving the value  
  set(v) { .. },  // function to invoke when assigning the value  
  // .. enumerable, etc  
}
```

A getter looks like a property access (`obj.prop`), but under the covers it invokes the `get()` method as defined; it's sort of like if you had called `obj.prop()`. A setter looks like a property assignment (`obj.prop = value`), but it invokes the `set(..)` method as defined; it's sort of like if you had called `obj.prop(value)`.

Let's illustrate a getter/setter accessor property:

```
anotherObj = {};  
  
Object.defineProperty(anotherObj, "fave", {  
  get() { console.log("Getting 'fave' value!"); return 123; },  
  set(v) { console.log(`Ignoring ${v} assignment.`); }  
});  
  
anotherObj.fave;  
// Getting 'fave' value!  
// 123  
  
anotherObj.fave = 42;  
// Ignoring 42 assignment.  
  
anotherObj.fave;  
// Getting 'fave' value!  
// 123
```

Enumerable, Writable, Configurable

Besides `value` or `get()` / `set(..)`, the other 3 attributes of a property descriptor are (as shown above):

- `enumerable`
- `writable`
- `configurable`

The `enumerable` attribute controls whether the property will appear in various enumerations of object properties, such as `Object.keys(..)`,

`Object.entries(..)`, `for..in` loops, and the copying that occurs with the `... object` spread and `Object.assign(..)`. Most properties should be left enumerable, but you can mark certain special properties on an object as non-enumerable if they shouldn't be iterated/copied.

The `writable` attribute controls whether a `value` assignment (via `=`) is allowed. To make a property “read only”, define it with `writable: false`. However, as long as the property is still configurable, `Object.defineProperty(..)` can still change the value by setting `value` differently.

The `configurable` attribute controls whether a property’s `descriptor` can be re-defined/overwritten. A property that’s `configurable: false` is locked to its definition, and any further attempts to change it with `Object.defineProperty(..)` will fail. A non-configurable property can still be assigned new values (via `=`), as long as `writable: true` is still set on the property’s descriptor.

Object Sub-Types

There are a variety of specialized sub-types of objects in JS. But by far, the two most common ones you’ll interact with are `arrays` and `functions`.

NOTE:

By “sub-type”, we mean the notion of a derived type that has inherited the behaviors from a parent type but then specialized or extended those behaviors. In other words, values of these sub-types are fully objects, but are also *more than just* objects.

Arrays

Arrays are objects that are specifically intended to be **numerically indexed**, rather than using string named property locations. They are still objects, so a named property like `favoriteNumber` is legal. But it’s greatly frowned upon to mix named properties into numerically indexed arrays.

Arrays are preferably defined with literal syntax (similar to objects), but with the `[..]` square brackets rather than `{ .. }` curly brackets:

```
myList = [ 23, 42, 109 ];
```

JS allows any mixture of value types in arrays, including objects, other arrays, functions, etc. As you’re likely already aware, arrays are “zero-indexed”, meaning the first element in the array is at the index 0, not 1:

```
myList = [ 23, 42, 109 ];
```

```
myList[0];      // 23
myList[1];      // 42
```

Recall that any string property name on an object that “looks like” an integer – is able to be validly coerced to a numeric integer – will actually be treated like an integer property (aka, integer index). The same goes for arrays. You should always use 42 as an integer index (aka, property name), but if you use the string "42", JS will assume you meant that as an integer and do that for you.

```
// "2" works as an integer index here, but it's not advised
myList["2"];    // 109
```

One exception to the “no named properties on arrays” *rule* is that all arrays automatically expose a `length` property, which is automatically kept updated with the “length” of the array.

```
myList = [ 23, 42, 109 ];

myList.length;    // 3

// "push" another value onto the end of the list
myList.push("Hello");

myList.length;    // 4
```

WARNING:

Many JS developers incorrectly believe that array `length` is basically a *getter* (see “Accessor Properties” earlier in this chapter), but it’s not. The offshoot is that these developers feel like it’s “expensive” to access this property – as if JS has to on-the-fly recompute the length – and will thus do things like capture/store the length of an array before doing a non-mutating loop over it. This used to be “best practice” from a performance perspective. But for at least 10 years now, that’s actually been an anti-pattern, because the JS engine is more efficient at managing the `length` property than our JS code is at trying to “outsmart” the engine to avoid invoking something we think is a *getter*. It’s more efficient to let the JS engine do its job, and just access the property whenever and however often it’s needed.

Empty Slots JS arrays also have a really unfortunate “flaw” in their design, referred to as “empty slots”. If you assign an index of an array more than one position beyond the current end of the array, JS will leave the in between slots “empty” rather than auto-assigning them to `undefined` as you might expect:

```
myList = [ 23, 42, 109 ];
myList.length;    // 3

myList[14] = "Hello";
myList.length;    // 15
```

```

myList;           // [ 23, 42, 109, empty x 11, "Hello" ]

// looks like a real slot with a
// real `undefined` value in it,
// but beware, it's a trick!
myList[9];       // undefined

```

You might wonder why empty slots are so bad? One reason: there are APIs in JS, like array's `map(..)`, where empty slots are surprisingly skipped over! Never, ever intentionally create empty slots in your arrays. This is undebateably one of JS's "bad parts".

Functions

I don't have much specifically to say about functions here, other than to point out that they are also sub-object-types. This means that in addition to being executable, they can also have named properties added to or accessed from them.

Functions have two pre-defined properties you may find yourself interacting with, specifically for meta-programming purposes:

```

function help(opt1,opt2,...remainingOpts) {
  // ...
}

help.name;        // "help"
help.length;      // 2

```

The `length` of a function is the count of its explicitly defined parameters, up to but not including a parameter that either has a default value defined (e.g., `param = 42`) or a "rest parameter" (e.g., `...remainingOpts`).

Avoid Setting Function-Object Properties You should avoid assigning properties on function objects. If you're looking to store extra information associated with a function, use a separate `Map(..)` (or `WeakMap(..)`) with the function object as the key, and the extra information as the value.

```

extraInfo = new Map();

extraInfo.set(help, "this is some important information");

// later:
extraInfo.get(help); // "this is some important information"

```

Object Characteristics

In addition to defining behaviors for specific properties, certain behaviors are configurable across the whole object:

- extensible
- sealed
- frozen

Extensible

Extensibility refers to whether an object can have new properties defined/added to it. By default, all objects are extensible, but you can change shut off extensibility for an object:

```
myObj = {  
    favoriteNumber: 42  
};  
  
myObj.firstName = "Kyle"; // works fine  
  
Object.preventExtensions(myObj);  
  
myObj.nicknames = [ "getify", "ydkjs" ]; // fails  
myObj.favoriteNumber = 123; // works fine
```

In non-strict-mode, an assignment that creates a new property will silently fail, whereas in strict mode an exception will be thrown.

Sealed

```
// TODO
```

Frozen

```
// TODO
```

Extending The MOP

As mentioned at the start of this chapter, objects in JS behave according to a set of rules referred to as the Metaobject Protocol (MOP)². Now that we understand more fully how objects work by default, we want to turn our attention to how we can hook into some of these default behaviors and override/customize them.

```
// TODO
```

²“Metaobject”, Wikipedia; <https://en.wikipedia.org/wiki/Metaobject> ; Accessed July 2022.

[[Prototype]] Chain

One of the most important, but least obvious, characteristics of an object (part of the MOP) is referred to as its “prototype chain”; the official JS specification notation is `[[Prototype]]`. Make sure not to confuse this `[[Prototype]]` with a public property named `prototype`. Despite the naming, these are distinct concepts.

The `[[Prototype]]` is an internal linkage that an object gets by default when its created, pointing to another object. This linkage is a hidden, often subtle characteristic of an object, but it has profound impacts on how interactions with the object will play out. It’s referred to as a “chain” because one object links to another, which in turn links to another, ... and so on. There is an *end* or *top* to this chain, where the linkage stops and there’s no further to go. More on that shortly.

We already saw several implications of `[[Prototype]]` linkage in Chapter 1. For example, by default, all objects are `[[Prototype]]`-linked to the built-in object named `Object.prototype`.

WARNING:

That `Object.prototype` name itself can be confusing, since it uses a property called `prototype`. How are `[[Prototype]]` and `prototype` related!? Put such questions/confusion on pause for a bit, as we’ll come back and explain the differences between `[[Prototype]]` and `prototype` later in this chapter. For the moment, just assume the presence of this important but weirdly named built-in object, `Object.prototype`.

Let’s consider some code:

```
myObj = {  
    favoriteNumber: 42  
};
```

That should look familiar from Chapter 1. But what you *don’t see* in this code is that the object there was automatically linked (via its internal `[[Prototype]]`) to that automatically built-in, but weirdly named, `Object.prototype` object.

When we do things like:

```
myObj.toString(); // "[object Object]"  
myObj.hasOwnProperty("favoriteNumber"); // true
```

We’re taking advantage of this internal `[[Prototype]]` linkage, without really realizing it. Since `myObj` does not have `toString` or `hasOwnProperty` properties defined on it, those property accesses actually end up **DELEGATING** the access to continue its lookup along the `[[Prototype]]` chain.

Since `myObj` is `[[Prototype]]`-linked to the object named `Object.prototype`, the lookup for `toString` and `hasOwnProperty` properties continues on that object; and indeed, these methods are found there!

The ability for `myObj.toString` to access the `toString` property even though it doesn't actually have it, is commonly referred to as "inheritance", or more specifically, "prototypal inheritance". The `toString` and `hasOwnProperty` properties, along with many others, are said to be "inherited properties" on `myObj`.

NOTE:

I have a lot of frustrations with the usage of the word "inheritance" here – it should be called "delegation"! – but that's what most people refer to it as, so we'll begrudgingly comply and use that same terminology for now (albeit under protest, with " quotes). I'll save my objections for an appendix of this book.

`Object.prototype` has several built-in properties and methods, all of which are "inherited" by any object that is `[[Prototype]]`-linked, either directly or indirectly through another object's linkage, to `Object.prototype`.

Some common "inherited" properties from `Object.prototype` include:

- `constructor`
- `__proto__`
- `toString()`
- `valueOf()`
- `hasOwnProperty(..)`
- `isPrototypeOf(..)`

Recall `hasOwnProperty(..)`, which we saw earlier gives us a boolean check for whether a certain property (by string name) is owned by an object:

```
myObj = {  
    favoriteNumber: 42  
};  
  
myObj.hasOwnProperty("favoriteNumber"); // true
```

It's always been considered somewhat unfortunate (semantic organization, naming conflicts, etc) that such an important utility as `hasOwnProperty(..)` was included on the `Object` `[[Prototype]]` chain as an instance method, instead of being defined as a static utility.

As of ES2022, JS has finally added the static version of this utility: `Object.hasOwn(..)`.

```
myObj = {  
    favoriteNumber: 42  
};
```

```
Object.hasOwnProperty(myObj, "favoriteNumber"); // true
```

This form is now considered the more preferable and robust option, and the instance method (`hasOwnProperty(..)`) form should now generally be avoided.

Somewhat unfortunately and inconsistently, there's not (yet, as of time of writing) corresponding static utilities, like `Object.isPrototypeOf(..)` (instead of the instance method `isPrototypeOfOf(..)`). But at least `Object.hasOwnProperty(..)` exists, so that's progress.

Creating An Object With A Different [[Prototype]]

By default, any object you create in your programs will be `[[Prototype]]`-linked to that `Object.prototype` object. However, you can create an object with a different linkage like this:

```
myObj = Object.create(differentObj);
```

The `Object.create(..)` method takes its first argument as the value to set for the newly created object's `[[Prototype]]`.

One downside to this approach is that you aren't using the `{ .. }` literal syntax, so you don't initially define any contents for `myObj`. You typically then have to define properties one-by-one, using `=`.

NOTE:

The second, optional argument to `Object.create(..)` is – like the second argument to `Object.defineProperties(..)` as discussed earlier – an object with properties that hold descriptors to initially define the new object with. In practice out in the wild, this form is rarely used, likely because it's more awkward to specify full descriptors instead of just name/value pairs. But it may come in handy in some limited cases.

Alternately, but less preferably, you can use the `{ .. }` literal syntax along with a special (and strange looking!) property:

```
myObj = {  
  __proto__: differentObj,  
  // .. the rest of the object definition  
};
```

WARNING:

The strange looking `__proto__` property has been in some JS engines for more than 20 years, but was only standardized in JS as of ES6 (in 2015). Even still, it was added in Appendix B of the specification³, which lists features that TC39 begrudgingly includes because they exist popularly in various browser-based JS engines and therefore are a de-facto reality even if they didn't originate with TC39. This feature is thus “guaranteed” by the spec to exist in all conforming browser-based JS engines, but is not necessarily guaranteed to work in other independent JS engines. Node.js uses the JS engine (v8) from the Chrome browser, so Node.js gets `__proto__` by default/accident. Be careful when using `__proto__` to be aware of all the JS engine environments your code will run in.

Whether you use `Object.create(..)` or `__proto__`, the created object in question will usually be `[[Prototype]]`-linked to a different object than the default `Object.prototype`.

Empty `[[Prototype]]` Linkage We mentioned above that the `[[Prototype]]` chain has to stop somewhere, so as to have lookups not continue forever. `Object.prototype` is typically the top/end of every `[[Prototype]]` chain, as its own `[[Prototype]]` is `null`, and therefore there's nowhere else to continue looking.

However, you can also define objects with their own `null` value for `[[Prototype]]`, such as:

```
emptyObj = Object.create(null);
// or: emptyObj = { __proto__: null }

emptyObj.toString; // undefined
```

It can be quite useful to create an object with no `[[Prototype]]` linkage to `Object.prototype`. For example, as mentioned in Chapter 1, the `in` and `for..in` constructs will consult the `[[Prototype]]` chain for inherited properties. But this may be undesirable, as you may not want something like `"toString" in myObj` to resolve successfully.

Moreover, an object with an empty `[[Prototype]]` is safe from any accidental “inheritance” collision between its own property names and the ones it “inherits” from elsewhere. These types of (useful!) objects are sometimes referred to in popular parlance as “dictionary objects”.

³“Appendix B: Additional ECMAScript Features for Web Browsers”, ECMAScript 2022 Language Specification; <https://262.ecma-international.org/13.0/#sec-additional-ecmascript-features-for-web-browsers> ; Accessed July 2022

[[Prototype]] vs prototype

Notice that public property name `prototype` in the name/location of this special object, `Object.prototype`? What's that all about?

`Object` is the `Object(..)` function; by default, all functions (which are themselves objects!) have such a `prototype` property on them, pointing at an object.

Any here's where the name conflict between `[[Prototype]]` and `prototype` really bites us. The `prototype` property on a function doesn't define any linkage that the function itself experiences. Indeed, functions (as objects) have their own internal `[[Prototype]]` linkage somewhere else – more on that in a second.

Rather, the `prototype` property on a function refers to an object that should be *linked TO* by any other object that is created when calling that function with the `new` keyword:

```
myObj = {};  
  
// is basically the same as:  
myObj = new Object();
```

Since the `{ .. }` object literal syntax is essentially the same as a `new Object()` call, the built-in object named/located at `Object.prototype` is used as the internal `[[Prototype]]` value for the new object we create and name `myObj`.

Phew! Talk about a topic made significantly more confusing just because of the name overlap between `[[Prototype]]` and `prototype`!

But where do functions themselves (as objects!) link to, `[[Prototype]]` wise? They link to `Function.prototype`, yet another built-in object, located at the `prototype` property on the `Function(..)` function.

In other words, you can think of functions themselves as having been “created” by a `new Function(..)` call, and then `[[Prototype]]`-linked to the `Function.prototype` object. This object contains properties/methods all functions “inherit” by default, such as `toString()` (to string serialize the source code of a function) and `call(..) / apply(..) / bind(..)` (we’ll explain these later in this book).

Objects Behavior

Properties on objects are internally defined and controlled by a “descriptor” metaobject, which includes attributes such as `value` (the property’s present value) and `enumerable` (a boolean controlling whether the property is included in enumerable-only listings of properties/property names).

The way object and their properties work in JS is referred to as the “metaob-

ject protocol” (MOP)⁴. We can control the precise behavior of properties via `Object.defineProperty(..)`, as well as object-wide behaviors with `Object.freeze(..)`. But even more powerfully, we can hook into and override certain default behaviors on objects using special pre-defined Symbols.

Prototypes are internal linkages between objects that allow property or method access against one object – if the property/method requested is absent – to be handled by “delegating” that access lookup to another object. When the delegation involves a method, the context for the method to run in is shared from the initial object to the target object via the `this` keyword.

⁴“Metaobject”, Wikipedia; <https://en.wikipedia.org/wiki/Metaobject> ; Accessed July 2022.

You Don't Know JS Yet: Objects & Classes - 2nd Edition

Chapter 3: Classy Objects

NOTE:

Work in progress

The class-design pattern generally entails defining a *type of thing* (class), including data (members) and behaviors (methods), and then creating one or more concrete *instances* of this class definition as actual objects that can interact and perform tasks. Moreover, class-orientation allows declaring a relationship between two or more classes, through what's called "inheritance", to derive new and augmented "subclasses" that mix-n-match and even re-define behaviors.

Prior to ES6 (2015), JS developers mimicked aspects of class-oriented (aka "object-oriented") design using plain functions and objects, along with the `[[Prototype]]` mechanism (as explained in the previous chapter) – so called "prototypal classes".

But to many developers joy and relief, ES6 introduced dedicated syntax, including the `class` and `extends` keywords, to express class-oriented design more declaratively.

At the time of ES6's `class` being introduced, this new dedicated syntax was almost entirely *just syntactic sugar* to make class definitions more convenient and readable. However, in the many years since ES6, `class` has matured and grown into its own first class feature mechanism, accruing a significant amount of dedicated syntax and complex behaviors that far surpass the pre-ES6 "prototypal class" capabilities.

Even though `class` now bears almost no resemblance to older "prototypal class" code style, the JS engine is still *just* wiring up objects to each other through the existing `[[Prototype]]` mechanism. In other words, `class` is not its own separate pillar of the language (as `[[Prototype]]` is), but more like the fancy, decorative *Capital* that tops the pillar/column.

That said, since `class` style code has now replaced virtually all previous "prototypal class" coding, the main text here focuses only on `class` and its various particulars. For historical purposes, we'll briefly cover the old "prototypal class" style in an appendix.

When Should I Class-Orient My Code?

Class-orientation is a design pattern, which means it's a choice for how you organize the information and behavior in your program. It has pros and cons.

It's not a universal solution for all tasks.

So how do you know when you should use classes?

In a theoretical sense, class-orientation is a way of dividing up the business domain of a program into one or more pieces that can each be defined by an “is-a” classification: grouping a thing into the set (or sets) of characteristics that thing shares with other similar things. You would say “X is a Y”, meaning X has (at least) all the characteristics of a thing of kind Y.

For example, consider computers. We could say a computer is electrical, since it uses electrical current (voltage, amps, etc) as power. It's furthermore electronic, because it manipulates the electrical current beyond simply routing electrons around (electrical/magnetic fields), creating a meaningful circuit to manipulate the current into performing more complex tasks. By contrast, a basic desk lamp is electrical, but not really electronic.

We could thus define a class `Electrical` to describe what electrical devices need and can do. We could then define a further class `Electronic`, and define that in addition to being electrical, `Electronic` things manipulate electricity to create more specialized outcomes.

Here's where class-orientation starts to shine. Rather than re-define all the `Electrical` characteristics in the `Electronic` class, we can define `Electronic` in such a way that it “shares” or “inherits” those characteristics from `Electrical`, and then augments/redefines the unique behaviors that make a device electronic. This relationship between the two classes – called “inheritance” – is a key aspect of class-orientation.

So class-orientation is a way of thinking about the entities our program needs, and classifying them into groupings based on their characteristics (what information they hold, what operations can be performed on that data), and defining the relationships between the different grouping of characteristics.

But moving from the theoretical into a bit more pragmatic perspective: if your program needs to hold and use multiple collections (instances) of alike data/behavior at once, you *may* benefit from class-orientation.

Time For An Example

Here's a short illustration.

A couple of decades ago, right after I had gone through nearly all of a Computer Science degree in college, I found myself sitting in my first professional software developer job. I was tasked with building, all by myself, a timesheet and payroll tracking system. I built the backend in PHP (using MySQL for the DB) and used JS for the interface (early as it was in its maturity way back around the turn of the century).

Since my CS degree had emphasized class-orientation heavily throughout my

courses, I was eager to put all that theory to work. For my program's design, I defined the concept of a "timesheet" entity as a collection of 2-3 "week" entities, and each "week" as a collection of 5-7 "day" entities, and each "day" as a collection of "task" entities.

If I wanted to know how many hours were logged into a timesheet instance, I could call a `totalTime()` operation on that instance. The timesheet defined this operation by looping over its collection of weeks, calling `totalTime()` on each of them and summing the values. Each week did the same for all its days, and each day did the same for all its tasks.

The notion being illustrated here, one of the fundamentals of design patterns like class-orientation, is called *encapsulation*. Each entity level encapsulated (e.g., controlled, hid, abstracted) internal details (data and behavior) while presenting a useful external interface.

But encapsulation alone isn't a sufficient justification for class-orientation. Other design patterns offer sufficient encapsulation.

How did my class design take advantage of inheritance? I had a base class that defined a set of operations like `totalTime()`, and each of my entity class types extended/subclassed this base class. That meant that each of them inherited this summation-of-total-time capability, but where each of them applied their own extensions and definitions for the internal details of *how* to do that work.

There's yet another aspect of the design pattern at play, which is *composition*: each entity was defined as a collection of other entities.

Single vs Multiple

I mentioned above that a pragmatic way of deciding if you need class-orientation is if your program is going to have multiple instances of a single kind/type of behavior (aka, "class"). In the timesheet example, we had 4 classes: Timesheet, Week, Day, and Task. But for each class, we had multiple instances of each at once.

Had we instead only needed a single instance of a class, like just one `Computer` thing that was an instance of the `Electronic` class, which was a subclass of the `Electrical` class, then class-orientation may not offer quite as much benefit. In particular, if the program doesn't need to create an instance of the `Electrical` class, then there's no particular benefit to separating `Electrical` from `Electronic`, so we aren't really getting any help from the inheritance aspect of class-orientation.

So, if you find yourself designing a program by dividing up a business problem domain into different "classes" of entities, but in the actual code of the program you are only ever need one concrete *thing* of one kind/definition of behavior (aka, "class"), you might very well not actually need class-orientation. There are other design patterns which may be a more efficient match to your effort.

But if you find yourself wanting to define classes, and subclasses which inherit from them, and if you're going to be instantiating one or more of those classes multiple times, then class-orientation is a good candidate. And to do class-orientation in JS, you're going to need the `class` keyword.

Keep It `classy`

`class` defines either a declaration or expression for a class. As a declaration, a class definition appears in a statement position and looks like this:

```
class Point2d {  
  // ...  
}
```

As an expression, a class definition appears in a value position and can either have a name or be anonymous:

```
// named class expression  
const pointClass = class Point2d {  
  // ...  
};  
  
// anonymous class expression  
const anotherClass = class {  
  // ...  
};
```

The contents of a `class` body typically include one or more method definitions:

```
class Point2d {  
  setX(x) {  
    // ...  
  }  
  setY(y) {  
    // ...  
  }  
}
```

Inside a `class` body, methods are defined without the `function` keyword, and there's no `,` or `;` separators between the method definitions.

NOTE:

Inside a `class` block, all code runs in strict-mode even without the `"use strict"` pragma present in the file or its functions. In particular, this impacts the `this` behavior for function calls, as explained in Chapter 4.

The Constructor

One special method that all classes have is called a “constructor”. If omitted, there’s a default empty constructor assumed in the definition.

The constructor is invoked any time a `new` instance of the class is created:

```
class Point2d {
  constructor() {
    console.log("Here's your new instance!");
  }
}

var point = new Point2d();
// Here's your new instance!
```

Even though the syntax implies a function actually named `constructor` exists, JS defines a function as specified, but with the name of the class (`Point2d` above):

```
typeof Point2d; // "function"
```

It’s not *just* a regular function, though; this special kind of function behaves a bit differently:

```
Point2d.toString();
// class Point2d {
//   ..
// }

Point2d();
// TypeError: Class constructor Point2d cannot
// be invoked without 'new'

Point2d.call({});
// TypeError: Class constructor Point2d cannot
// be invoked without 'new'
```

You can construct as many different instances of a class as you need:

```
var one = new Point2d();
var two = new Point2d();
var three = new Point2d();
```

Each of `one`, `two`, and `three` here are objects that are independent instances of the `Point2d` class.

NOTE:

Each of the `one`, `two`, and `three` objects have a `[[Prototype]]` linkage to the `Point2d.prototype` object (see Chapter 2). In this code, `Point2d` is both a `class` definition and the constructor function of the same name.

If you add a property to the object `one`:

```
one.value = 42;
```

That property now exists only on `one`, and does not exist in any way that the independent `two` or `three` objects can access:

```
two.value;      // undefined
three.value;    // undefined
```

Class Methods

As shown above, a class definition can include one or more method definitions:

```
class Point2d {
  constructor() {
    console.log("Here's your new instance!");
  }
  setX(x) {
    console.log(`Setting x to: ${x}`);
    // ...
  }
}

var point = new Point2d();

point.setX(3);
// Setting x to: 3
```

The `setX` property (method) *looks like* it exists on (is owned by) the `point` object here. But that's a mirage. Each class method is added to the `prototype` object, a property of the constructor function.

So, `setX(..)` only exists as `Point2d.prototype.setX`. Since `point` is `[[Prototype]]` linked to `Point2d.prototype` (see Chapter 2) via the `new` keyword instantiation, the `point.setX(..)` reference traverses the `[[Prototype]]` chain and finds the method to execute.

Class methods should only be invoked via an instance; `Point2d.setX(..)` doesn't work because there *is no* such property. You *could* invoke `Point2d.prototype.setX(..)`, but that's not generally proper/advised in standard class-oriented coding. Always access class methods via the instances.

Class Instance `this`

We will cover the `this` keyword in much more detail in a subsequent chapter. But as it relates to class-oriented code, the `this` keyword generally refers to the current instance that is the context of any method invocation.

In the constructor, as well as any methods, you can use `this.` to either add or access properties on the current instance:

```
class Point2d {
  constructor(x,y) {
    // add properties to the current instance
    this.x = x;
    this.y = y;
  }
  toString() {
    // access the properties from the current instance
    console.log(`(${this.x},${this.y})`);
  }
}

var point = new Point2d(3,4);

point.x;           // 3
point.y;           // 4

point.toString();  // (3,4)
```

Any properties not holding function values, which are added to a class instance (usually via the constructor), are referred to as *members*, as opposed to the term *methods* for executable functions.

While the `point.toString()` method is running, its `this` reference is pointing at the same object that `point` references. That's why both `point.x` and `this.x` reveal the same 3 value that the constructor set with its `this.x = x` operation.

Public Fields

Instead of defining a class instance member imperatively via `this.` in the constructor or a method, classes can declaratively define *fields* in the `class` body, which correspond directly to members that will be created on each instance:

```
class Point2d {
  // these are public fields
  x = 0
  y = 0

  constructor(x,y) {
    // set properties (fields) on the current instance
```

```

        this.x = x;
        this.y = y;
    }
    toString() {
        // access the properties from the current instance
        console.log(`(${this.x},${this.y})`);
    }
}

```

Public fields can have a value initialization, as shown above, but that's not required. If you don't initialize a field in the class definition, you almost always should initialize it in the constructor.

Fields can also reference each other, via natural `this.` access syntax:

```

class Point3d {
    // these are public fields
    x
    y = 4
    z = this.y * 5

    // ...
}

```

TIP:

You can mostly think of public field declarations as if they appear at the top of the `constructor(..)`, each prefixed with an implied `this.` that you get to omit in the declarative `class` body form. But, there's a catch! See “That's Super!” later for more information about it.

Just like computed property names (see Chapter 1), field names can be computed:

```

var coordName = "x";

class Point2d {
    // computed public field
    [coordName.toUpperCase()] = 42

    // ...
}

var point = new Point2d(3,4);

point.x;          // 3
point.y;          // 4

```

```
point.x;           // 42
```

Avoid This One pattern that has emerged and grown quite popular, but which I firmly believe is an anti-pattern for `class`, looks like the following:

```
class Point2d {
  x = null
  y = null
  getDoubleX = () => this.x * 2

  constructor(x,y) {
    this.x = x;
    this.y = y;
  }
  toString() { /* .. */ }
}

var point = new Point2d(3,4);

point.getDoubleX();    // 6
```

See the field holding an `=>` arrow function? I say this is a no-no. But why? Let's unwind what's going on.

First, why do this? Because JS developers seem to be perpetually frustrated by the dynamic `this` binding rules (see Chapter 4), so they force a `this` binding via the `=>` arrow function. That way, no matter how `getDoubleX()` is invoked, it's always `this`-bound to the particular instance. That's an understandable convenience to desire, but... it betrays the very nature of the `this` / `[[Prototype]]` pillar of the language. How?

Let's consider the equivalent code to the previous snippet:

```
class Point2d {
  constructor(x,y) {
    this.x = null;
    this.y = null;
    this.getDoubleX = () => this.x * 2;

    this.x = x;
    this.y = y;
  }
  toString() { /* .. */ }
}

var point = new Point2d(3,4);
```

```
point.getDoubleX();    // 6
```

Can you spot the problem? Look closely. I'll wait.

...

We've made it clear repeatedly so far that `class` definitions put their methods on the class constructor's `prototype` object – that's where they belong! – such that there's just one of each function and it's inherited (shared) by all instances. That's what will happen with `toString()` in the above snippet.

But what about `getDoubleX()`? That's essentially a class method, but it won't be handled by JS quite the same as `toString()` will. Consider:

```
Object.hasOwnProperty(point, "x");           // true -- good
Object.hasOwnProperty(point, "toString");      // false -- good
Object.hasOwnProperty(point, "getDoubleX");    // true -- oops :(
```

You see now? By defining a function value and attaching it as a field/member property, we're losing the shared prototypal method'ness of the function, and it becomes just like any per-instance property. That means we're creating a new function property **for each instance**, rather than it being created just once on the class constructor's `prototype`.

That's wasteful in performance and memory, even if by a tiny bit. That alone should be enough to avoid it.

But I would argue that way more importantly, what you've done with this pattern is invalidate the very reason why using `class` and `this`-aware methods is even remotely useful/powerful!

If you go to all the trouble to define class methods with `this.` references throughout them, but then you lock/bind most or all of those methods to a specific object instance, you've basically travelled all the way around the world just to go next door.

If all you want are function(s) that are statically fixed to a particular “context”, and don't need any dynamicism or sharing, what you want is... **closure**. And you're in luck: I wrote a whole book in this series (“Scope & Closures”) on how to use closure so functions remember/access their statically defined scope (aka “context”). That's a way more appropriate, and simpler to code, approach to get what you're after.

Don't abuse/misuse `class` and turn it into a over-hyped, glorified collection of closure.

To be clear, I'm *not* saying: never use `=>` arrow functions inside classes.

I *am* saying: never attach an `=>` arrow function as an instance property in place of a dynamic prototypal class method, either out of mindless habit, or laziness in typing fewer characters, or misguided `this`-binding convenience.

In a subsequent chapter, we'll dive deep into how to understand and properly leverage the full power of the dynamic `this` mechanism.

Class Extension

The way to unlock the power of class inheritance is through the `extends` keyword, which defines a relationship between two classes:

```
class Point2d {
  x = 3
  y = 4

  getX() {
    return this.x;
  }
}

class Point3d extends Point2d {
  x = 21
  y = 10
  z = 5

  printDoubleX() {
    console.log(`double x: ${this.getX() * 2}`);
  }
}

var point = new Point2d();

point.getX(); // 3

var anotherPoint = new Point3d();

anotherPoint.getX(); // 21
anotherPoint.printDoubleX(); // double x: 42
```

Take a few moments to re-read that code snippet and make sure you fully understand what's happening.

The base class `Point2d` defines fields (members) called `x` and `y`, and gives them the initial values 3 and 4, respectively. It also defines a `getX()` method that accesses this `x` instance member and returns it. We see that behavior illustrated in the `point.getX()` method call.

But the `Point3d` class extends `Point2d`, making `Point3d` a derived-class, child-class, or (most commonly) subclass. In `Point3d`, the same `x` property that's inherited from `Point2d` is re-initialized with a different 21 value, as is the `y` overridden to value from 4, to 10.

It also adds a new `z` field/member method, as well as a `printDoubleX()` method, which itself calls `this.getX()`.

When `anotherPoint.printDoubleX()` is invoked, the inherited `this.getX()` is thus invoked, and that method makes reference to `this.x`. Since `this` is pointing at the class instance (aka, `anotherPoint`), the value it finds is now 21 (instead of 3 from the `point` object's `x` member).

Extending Expressions

```
// TODO: cover class Foo extends ... where ... is an expression, not a class-name
```

Overriding Methods

In addition to overriding a field/member in a subclass, you can also override (redefine) a method:

```
class Point2d {
    x = 3
    y = 4

    getX() {
        return this.x;
    }
}

class Point3d extends Point2d {
    x = 21
    y = 10
    z = 5

    getX() {
        return this.x * 2;
    }
    printX() {
        console.log(`double x: ${this.getX()}`);
    }
}

var point = new Point3d();

point.printX();      // double x: 42
```

The `Point3d` subclass overrides the inherited `getX()` method to give it different behavior. However, you can still instantiate the base `Point2d` class, which would then give an object that uses the original (`return this.x;`) definition for `getX()`.

If you want to access an inherited method from a subclass even if it's been overridden, you can use `super` instead of `this`:

```
class Point2d {
  x = 3
  y = 4

  getX() {
    return this.x;
  }
}

class Point3d extends Point2d {
  x = 21
  y = 10
  z = 5

  getX() {
    return this.x * 2;
  }
  printX() {
    console.log(`x: ${super.getX()}`);
  }
}

var point = new Point3d();

point.printX();      // x: 21
```

The ability for methods of the same name, at different levels of the inheritance hierarchy, to exhibit different behavior when either accessed directly, or relatively with `super`, is called *method polymorphism*. It's a very powerful part of class-orientation, when used appropriately.

That's Super!

In addition to a subclass method accessing an inherited method definition (even if overridden on the subclass) via `super.` reference, a subclass constructor must manually invoke the inherited base class constructor via `super(..)` function invocation:

```
class Point2d {
  x
  y
  constructor(x,y) {
    this.x = x;
    this.y = y;
```

```

        }
    }

    class Point3d extends Point2d {
        z
        constructor(x,y,z) {
            super(x,y);
            this.z = z;
        }
        toString() {
            console.log(`(${this.x},${this.y},${this.z})`);
        }
    }

    var point = new Point3d(3,4,5);

    point.toString();      // (3,4,5)

```

WARNING:

An explicitly defined subclass constructor *must* call `super(..)` to run the inherited class's initialization, and that must occur before the subclass constructor makes any references to `this` or finishes/returns. Otherwise, a runtime exception will be thrown when that subclass constructor is invoked (via `new`). If you omit the subclass constructor, the default constructor automatically – thankfully! – invokes `super()` for you.

One nuance to be aware of: if you define a field (public or private) inside a subclass, and explicitly define a `constructor(..)` for this subclass, the field initializations will be processed not at the top of the constructor, but *between* the `super(..)` call and any subsequent code in the constructor.

Pay close attention to the order of console messages here:

```

class Point2d {
    x
    y
    constructor(x,y) {
        console.log("Running Point2d(..) constructor");
        this.x = x;
        this.y = y;
    }
}

class Point3d extends Point2d {
    z = console.log("Initializing field 'z'")

```

```

constructor(x,y,z) {
    console.log("Running Point3d(..) constructor");
    super(x,y);

    console.log(`Setting instance property 'z' to ${z}`);
    this.z = z;
}
toString() {
    console.log(`${this.x},${this.y},${this.z}`);
}
}

var point = new Point3d(3,4,5);
// Running Point3d(..) constructor
// Running Point2d(..) constructor
// Initializing field 'z'
// Setting instance property 'z' to 5

```

As the console messages illustrate, the `z = ..` field initialization happens *immediately after* the `super(x,y)` call, *before* the `console.log(`Setting instance...`)` is executed. Perhaps think of it like the field initializations attached to the end of the `super(..)` call, so they run before anything else in the constructor does.

Which Class? You may need to determine in a constructor if that class is being instantiated directly, or being instantiated from a subclass with a `super()` call. We can use a special “pseudo property” `new.target`:

```

class Point2d {
    // ..

    constructor(x,y) {
        if (new.target === Point2d) {
            console.log("Constructing 'Point2d' instance");
        }
    }

    // ..
}

class Point3d extends Point2d {
    // ..

    constructor(x,y,z) {
        super(x,y);

```

```

        if (new.target === Point3d) {
            console.log("Constructing 'Point3d' instance");
        }
    }

    // ...
}

var point = new Point2d(3,4);
// Constructing 'Point2d' instance

var anotherPoint = new Point3d(3,4,5);
// Constructing 'Point3d' instance

```

But Which Kind Of Instance?

You may want to introspect a certain object instance to see if it's an instance of a specific class. We do this with the `instanceof` operator:

```

class Point2d { /* .. */ }
class Point3d extends Point2d { /* .. */ }

var point = new Point2d(3,4);

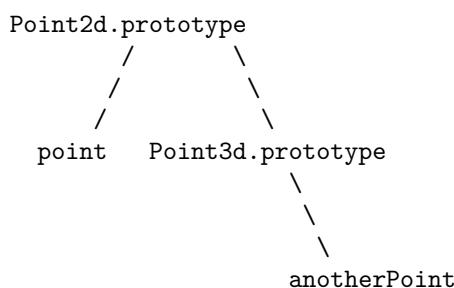
point instanceof Point2d;           // true
point instanceof Point3d;          // false

var anotherPoint = new Point3d(3,4,5);

anotherPoint instanceof Point2d;    // true
anotherPoint instanceof Point3d;   // true

```

It may seem strange to see `anotherPoint instanceof Point2d` result in `true`. To understand why better, perhaps it's useful to visualize both `[[Prototype]]` chains:



The `instanceof` operator doesn't just look at the current object, but rather

traverses the entire class inheritance hierarchy (the `[[Prototype]]` chain) until it finds a match. Thus, `anotherPoint` is an instance of both `Point3d` and `Point2d`.

To illustrate this fact a little more obviously, another (less ergonomic) way of going about the same kind of check as `instanceof` is with the (inherited from `Object.prototype`) utility, `isPrototypeOf(..)`:

```
Point2d.prototype.isPrototypeOf(point);           // true
Point3d.prototype.isPrototypeOf(point);           // false

Point2d.prototype.isPrototypeOf(anotherPoint);    // true
Point3d.prototype.isPrototypeOf(anotherPoint);    // true
```

This utility makes it a little clearer why both `Point2d.prototype.isPrototypeOf(anotherPoint)` and `anotherPoint instanceof Point2d` result in `true`: the object `Point2d.prototype` is in the `[[Prototype]]` chain of `anotherPoint`.

If you instead wanted to check if the object instance was *only and directly* created by a certain class, check the instance's `constructor` property.

```
point.constructor === Point2d;           // true
point.constructor === Point3d;           // false

anotherPoint.constructor === Point2d;    // false
anotherPoint.constructor === Point3d;    // true
```

NOTE:

The `constructor` property shown here is *not* actually present on (owned) the `point` or `anotherPoint` instance objects. So where does it come from!? It's on each object's `[[Prototype]]` linked prototype object:

`Point2d.prototype.constructor === Point2d` and
`Point3d.prototype.constructor === Point3d`.

“Inheritance” Is Sharing, Not Copying

It may seem as if `Point3d`, when it extends the `Point2d` class, is in essence getting a *copy* of all the behavior defined in `Point2d`. Moreover, it may seem as if the concrete object instance `anotherPoint` receives, *copied down* to it, all the methods from `Point3d` (and by extension, also from `Point2d`).

However, that's not the correct mental model to use for JS's implementation of class-orientation. Recall this base class and subclass definition, as well as instantiation of `anotherPoint`:

```
class Point2d {
  x
  y
```

```

    constructor(x,y) {
      this.x = x;
      this.y = y;
    }
}

class Point3d extends Point2d {
  z
  constructor(x,y,z) {
    super(x,y);
    this.z = z;
  }
  toString() {
    console.log(`(${this.x},${this.y},${this.z})`);
  }
}

var anotherPoint = new Point3d(3,4,5);

```

If you inspect the `anotherPoint` object, you'll see it only has the `x`, `y`, and `z` properties (instance members) on it, but not the `toString()` method:

```

Object.hasOwnProperty(anotherPoint,"x");           // true
Object.hasOwnProperty(anotherPoint,"y");           // true
Object.hasOwnProperty(anotherPoint,"z");           // true

Object.hasOwnProperty(anotherPoint,"toString");    // false

```

Where is that `toString()` method located? On the prototype object:

```
Object.hasOwnProperty(Point3d.prototype,"toString"); // true
```

And `anotherPoint` has access to that method via its `[[Prototype]]` linkage (see Chapter 2). In other words, the prototype objects **share access** to their method(s) with the subclass(es) and instance(s). The method(s) stay in place, and are not copied down the inheritance chain.

As nice as the `class` syntax is, don't forget what's really happening under the syntax: JS is *just* wiring up objects to each other along a `[[Prototype]]` chain.

Static Class Behavior

We've so far emphasized two different locations for data or behavior (methods) to reside: on the constructor's prototype, or on the instance. But there's a third option: on the constructor (function object) itself.

In a traditional class-oriented system, methods defined on a class are not concrete things you could ever invoke or interact with. You have to instantiate a class to have a concrete object to invoke those methods with. Prototypal

languages like JS blur this line a bit: all class-defined methods are “real” functions residing on the constructor’s prototype, and you could therefore invoke them. But as I asserted earlier, you really *should not* do so, as this is not how JS assumes you will write your `classes`, and there are some weird corner-case behaviors you may run into. Best to stay on the narrow path that `class` lays out for you.

Not all behavior that we define and want to associate/organize with a class *needs* to be aware of an instance. Moreover, sometimes a class needs to publicly define data (like constants) that developers using that class need to access, independent of any instance they may or may not have created.

So, how does a class system enable defining such data and behavior that should be available with a class but independent of (unaware of) instantiated objects? **Static properties and functions.**

NOTE:

I’ll use “static property” / “static function”, rather than “member” / “method”, just so it’s clearer that there’s a distinction between instance-bound members / instance-aware methods, and non-instance properties and instance-unaware functions.

We use the `static` keyword in our `class` bodies to distinguish these definitions:

```
class Point2d {
    // class statics
    static origin = new Point2d(0,0)
    static distance(point1,point2) {
        return Math.sqrt(
            ((point2.x - point1.x) ** 2) +
            ((point2.y - point1.y) ** 2)
        );
    }

    // instance members and methods
    x
    y
    constructor(x,y) {
        this.x = x;
        this.y = y;
    }
    toString() {
        return `(${this.x},${this.y})`;
    }
}
```

```

console.log(`Starting point: ${Point2d.origin}`);
// Starting point: (0,0)

var next = new Point2d(3,4);
console.log(`Next point: ${next}`);
// Next point: (3,4)

console.log(`Distance: ${
  Point2d.distance( Point2d.origin, next )
}`);
// Distance: 5

```

The `Point2d.origin` is a static property, which just so happens to hold a constructed instance of our class. And `Point2d.distance(..)` is a static function that computes the 2-dimensional cartesian distance between two points.

Of course, we could have put these two somewhere other than as `statics` on the class definition. But since they're directly related to the `Point2d` class, it makes *most sense* to organize them there.

NOTE:

Don't forget that when you use the `class` syntax, the name `Point2d` is actually the name of a constructor function that JS defines. So `Point2d.origin` is just a regular property access on that function object. That's what I meant at the top of this section when I referred to a third location for storing *things* related to classes; in JS, `statics` are stored as properties on the constructor function. Take care not to confuse those with properties stored on the constructor's `prototype` (methods) and properties stored on the instance (members).

Static Property Initializations

The value in a static initialization (`static whatever = ..`) can include `this` references, which refers to the class itself (actually, the constructor) rather than to an instance:

```

class Point2d {
  // class statics
  static originX = 0
  static originY = 0
  static origin = new this(this.originX, this.originY)

  // ..
}

```

WARNING:

I don't recommend actually doing the `new this(..)` trick I've illustrated here. That's just for illustration purposes. The code would read more cleanly with `new Point2d(this.originX, this.originY)`, so prefer that approach.

An important detail not to gloss over: unlike public field initializations, which only happen once an instantiation (with `new`) occurs, class static initializations always run *immediately* after the `class` has been defined. Moreover, the order of static initializations matters; you can think of the statements as if they're being evaluated one at a time.

Also like class members, static properties do not have to be initialized (default: `undefined`), but it's much more common to do so. There's not much utility in declaring a static property with no initialized value (`static whatever`); Accessing either `Point2d.whatever` or `Point2d.nonExistent` would both result in `undefined`.

Recently (in ES2022), the `static` keyword was extended so it can now define a block inside the `class` body for more sophisticated initialization of `statics`:

```
class Point2d {
  // class statics
  static origin = new Point2d(0,0)
  static distance(point1,point2) {
    return Math.sqrt(
      ((point2.x - point1.x) ** 2) +
      ((point2.y - point1.y) ** 2)
    );
  }

  // static initialization block (as of ES2022)
  static {
    let outerPoint = new Point2d(6,8);
    this.maxDistance = this.distance(
      this.origin,
      outerPoint
    );
  }

  // ...
}

Point2d.maxDistance; // 10
```

The `let outerPoint = ..` here is not a special `class` feature; it's exactly like a normal `let` declaration in any normal block of scope (see the "Scope &

Closures” book of this series). We’re merely declaring a localized instance of `Point2d` assigned to `outerPoint`, then using that value to derive the assignment to the `maxDistance` static property.

Static initialization blocks are also useful for things like `try..catch` statements around expression computations.

Static Inheritance

Class statics are inherited by subclasses (obviously, as statics!), can be overridden, and `super` can be used for base class references (and static function polymorphism), all in much the same way as inheritance works with instance members/methods:

```
class Point2d {
    static origin = /* .. */
    static distance(x,y) { /* .. */ }

    static {
        // ..
        this.maxDistance = /* .. */;
    }

    // ..
}

class Point3d extends Point2d {
    // class statics
    static origin = new Point3d(
        // here, `this.origin` references wouldn't
        // work (self-referential), so we use
        // `super.origin` references instead
        super.origin.x, super.origin.y, 0
    )
    static distance(point1,point2) {
        // here, super.distance(..) is Point2d.distance(..),
        // if we needed to invoke it

        return Math.sqrt(
            ((point2.x - point1.x) ** 2) +
            ((point2.y - point1.y) ** 2) +
            ((point2.z - point1.z) ** 2)
        );
    }

    // instance members/methods
    z
```

```

constructor(x,y,z) {
  super(x,y);      // <-- don't forget this line!
  this.z = z;
}
toString() {
  return `(${this.x},${this.y},${this.z})`;
}
}

Point2d.maxDistance;      // 10
Point3d.maxDistance;      // 10

```

As you can see, the static property `maxDistance` we defined on `Point2d` was inherited as a static property on `Point3d`.

TIP:

Remember: any time you define a subclass constructor, you'll need to call `super(..)` in it, usually as the first statement. I find that all too easy to forget.

Don't skip over the underlying JS behavior here. Just like method inheritance discussed earlier, the static "inheritance" is *not* a copying of these static properties/functions from base class to subclass; it's sharing via the `[[Prototype]]` chain. Specifically, the constructor function `Point3d()` has its `[[Prototype]]` linkage changed by JS (from the default of `Function.prototype`) to `Point2d`, which is what allows `Point3d.maxDistance` to delegate to `Point2d.maxDistance`.

It's also interesting, perhaps only historically now, to note that static inheritance – which was part of the original ES6 `class` mechanism feature set! – was one specific feature that went beyond "just syntax sugar". Static inheritance, as we see it illustrated here, was *not* possible to achieve/emulate in JS prior to ES6, in the old prototypal-class style of code. It's a special new behavior introduced only as of ES6.

Private Class Behavior

Everything we've discussed so far as part of a `class` definition is publicly visible/accessible, either as static properties/functions on the class, methods on the constructor's `prototype`, or member properties on the instance.

But how do you store information that cannot be seen from outside the class? This was one of the most asked for features, and biggest complaints with JS's `class`, up until it was finally addressed in ES2022.

`class` now supports new syntax for declaring private fields (instance members) and private methods. In addition, private static properties/functions are possible.

Motivation?

Before we illustrate how to do `class` privates, it bears contemplating why this is a helpful feature?

With closure-oriented design patterns (again, see the “Scope & Closures” book of this series), we automatically get “privacy” built-in. When you declare a variable inside a scope, it cannot be seen outside that scope. Period. Reducing the scope visibility of a declaration is helpful in preventing namespace collisions (identical variable names).

But it’s even more important to ensure proper “defensive” design of software, the so called “Principle of Least Privilege”¹. POLP states that we should only expose a piece of information or capability in our software to the smallest surface area necessary.

Over-exposure opens our software up to several issues that complicate software security/maintenance, including another piece of code acting maliciously to do something our code didn’t expect or intend. Moreover, there’s the less critical but still as problematic concern of other parts of our software relying on (using) parts of our code that we should have reserved as hidden implementation detail. Once other code relies on our code’s implementation details, we stop being able to refactor our code without potentially breaking other parts of the program.

So, in short, we *should* hide implementation details if they’re not necessary to be exposed. In this sense, JS’s `class` system feels a bit too permissive in that everything defaults to being public. Class-private features are a welcomed addition to more proper software design.

Too Private? All that said, I have to throw a bit of a damper on the class-private party.

I’ve suggested strongly that you should only use `class` if you’re going to really take advantage of most or all of what class-orientation gives you. Otherwise, you’d be better suited using other core pillar features of JS for organizing code, such as with the closure pattern.

One of the most important aspects of class-orientation is subclass inheritance, as we’ve seen illustrated numerous times so far in this chapter. Guess what happens to a private member/method in a base class, when it’s extended by a subclass?

¹“Principle of Least Privilege”, Wikipedia; https://en.wikipedia.org/wiki/Principle_of_least_privilege; Accessed July 2022

Private members/methods are private **only to the class they're defined in**, and are **not** inherited in any way by a subclass. Uh oh.

That might not seem like too big of a concern, until you start working with `class` and private members/methods in real software. You might quickly run up against a situation where you need to access a private method, or more often even, just a private member, from the subclass, so that the subclass can extend/augment the behavior of the base class as desired. And you might scream in frustration pretty quickly once you realize this is not possible.

What comes next is inevitably an awkward decision: do you just go back to making it public, so the subclass can access it? Ugh. Or, worse, do you try to re-design the base class to contort the design of its members/methods, such that the lack of access is partially worked around. That often involves exhausting over-parameterization (with privates as default parameter values) of methods, and other such tricks. Double ugh.

There's not a particularly great answer here, to be honest. If you have experience with class-orientation in more traditional class languages like Java or C++, you're probably dubious as to why we don't have *protected* visibility in between *public* and *private*. That's exactly what *protected* is for: keeping something private to a class AND any of its subclasses. Those languages also have *friend* features, but that's beyond the scope of our discussion here.

Sadly, not only does JS not have *protected* visibility, it seems (even as useful as it is!) to be unlikely as a JS feature. It's been discussed in great detail for over a decade (before ES6 was even a thing), and there've been multiple proposals for it.

I shouldn't say it will *never* happen, because that's not solid ground to stake on in any software. But it's very unlikely, because it actually betrays the very pillar that `class` is built on. If you are curious, or (more likely) certain that there's just *got to be a way*, I'll cover the incompatibility of *protected* visibility within JS's mechanisms in an appendix.

The point here is, as of now, JS has no *protected* visibility, and it won't any time soon. And *protected* visibility is actually, in practice, way more useful than *private* visibility.

So we return to the question: **Why should you care to make any class contents private?**

If I'm being honest: maybe you shouldn't. Or maybe you should. That's up to you. Just go into it aware of the stumbling blocks.

Private Members/Methods

You're excited to finally see the syntax for magical *private* visibility, right? Please don't shoot the messenger if you feel angered or sad at what you're about to see.

```

class Point2d {
  // statics
  static samePoint(point1,point2) {
    return point1.#ID === point2.#ID;
  }

  // privates
  #ID = null
  #assignID() {
    this.#ID = Math.round(Math.random() * 1e9);
  }

  // publics
  x
  y
  constructor(x,y) {
    this.#assignID();
    this.x = x;
    this.y = y;
  }
}

var one = new Point2d(3,4);
var two = new Point2d(3,4);

Point2d.samePoint(one,two);           // false
Point2d.samePoint(one,one);          // true

```

No, JS didn't do the sensible thing and introduce a `private` keyword like they did with `static`. Instead, they introduced the `#`. (insert lame joke about social-media millenials loving hashtags, or something)

TIP:

And yes, there's a million and one discussions about why not. I could spend chapters recounting the whole history, but honestly I just don't care to. I think this syntax is ugly, and many others do, too. And some love it! If you're in the latter camp, though I rarely do something like this, I'm just going to say: **just accept it**. It's too late for any more debate or pleading.

The `#whatever` syntax (including `this.#whatever` form) is only valid inside `class` bodies. It will throw syntax errors if used outside of a `class`.

Unlike public fields/instance members, private fields/instance members *must* be declared in the `class` body. You cannot add a private member to a class declaration dynamically while in the constructor method; `this.#whatever =`

.. type assignments only work if the `#whatever` private field is declared in the class body. Moreover, though private fields can be re-assigned, they cannot be `deleted` from an instance, the way a public field/class member can.

Subclassing + Privates I warned earlier that subclassing with classes that have private members/methods can be a limiting trap. But that doesn't mean they cannot be used together.

Because "inheritance" in JS is sharing (through the `[[Prototype]]` chain), if you invoke an inherited method in a subclass, and that inherited method in turn accesses/invokes privates in its host (base) class, this works fine:

```
class Point2d { /* .. */ }

class Point3d extends Point2d {
  z
  constructor(x,y,z) {
    super(x,y);
    this.z = z;
  }
}

var one = new Point3d(3,4,5);
```

The `super(x,y)` call in this constructor invokes the inherited base class constructor (`Point2d(..)`), which itself accesses `Point2d`'s private method `#assignID()` (see the earlier snippet). No exception is thrown, even though `Point3d` cannot directly see or access the `#ID` / `#assignID()` privates that are indeed stored on the instance (named `one` here).

In fact, even the inherited `static samePoint(..)` function will work from either `Point3d` or `Point2d`:

```
Point2d.samePoint(one,one);           // true
Point3d.samePoint(one,one);           // true
```

Actually, that shouldn't be that surprising, since:

```
Point2d.samePoint === Point3d.samePoint;
```

The inherited function reference is *the exact same function* as the base function reference; it's not some copy of the function. Because the function in question has no `this` reference in it, no matter from where it's invoked, it should produce the same outcome.

It's still a shame though that `Point3d` has no way to access/influence, or indeed even knowledge of, the `#ID` / `#assignID()` privates from `Point2d`:

```
class Point2d { /* .. */ }
```

```

class Point3d extends Point2d {
  z
  constructor(x,y,z) {
    super(x,y);
    this.z = z;

    console.log(this.#ID);      // will throw!
  }
}

```

WARNING:

Notice that this snippet throws an early static syntax error at the time of defining the `Point3d` class, before ever even getting a chance to create an instance of the class. The same exception would be thrown if the reference was `super.#ID` instead of `this.#ID`.

Existence Check Keep in mind that only the `class` itself knows about, and can therefore check for, such a private field/method.

You may want to check to see if a private field/method exists on an object instance. For example (as shown below), you may have a static function or method in a class, which receives an external object reference passed in. To check to see if the passed-in object reference is of this same class (and therefore has the same private members/methods in it), you basically need to do a “brand check” against the object.

Such a check could be rather convoluted, because if you access a private field that doesn’t already exist on the object, you get a JS exception thrown, requiring ugly `try..catch` logic.

But there’s a cleaner approach, so called an “ergonomic brand check”, using the `in` keyword:

```

class Point2d {
  // statics
  static samePoint(point1,point2) {
    // "ergonomic brand checks"
    if (#ID in point1 && #ID in point2) {
      return point1.#ID === point2.#ID;
    }
    return false;
  }

  // privates
  #ID = null
  #assignID() {

```

```

        this.#ID = Math.round(Math.random() * 1e9);
    }

    // publics
    x
    y
    constructor(x,y) {
        this.#assignID();
        this.x = x;
        this.y = y;
    }
}

var one = new Point2d(3,4);
var two = new Point2d(3,4);

Point2d.samePoint(one,two);           // false
Point2d.samePoint(one,one);          // true

```

The `#privateField` in `someObject` check will not throw an exception if the field isn't found, so it's safe to use without `try..catch` and use its simple boolean result.

Exfiltration Even though a member/method may be declared with *private* visibility, it can still be exfiltrated (extracted) from a class instance:

```

var id, func;

class Point2d {
    // privates
    #ID = null
    #assignID() {
        this.#ID = Math.round(Math.random() * 1e9);
    }

    // publics
    x
    y
    constructor(x,y) {
        this.#assignID();
        this.x = x;
        this.y = y;

        // exfiltration
        id = this.#ID;
        func = this.#assignID;
    }
}

```

```

        }
    }

var point = new Point2d(3,4);

id; // 7392851012 (...for example)

func; // function #assignID() { .. }
func.call(point,42);

func.call({},100);
// TypeError: Cannot write private member #ID to an
// object whose class did not declare it

```

The main concern here is to be careful when passing private methods as callbacks (or in any way exposing privates to other parts of the program). There's nothing stopping you from doing so, which can create a bit of an unintended privacy disclosure.

Private Statics

Static properties and functions can also use `#` to be marked as private:

```

class Point2d {
    static #errorMsg = "Out of bounds."
    static #printError() {
        console.log(`Error: ${this.#errorMsg}`);
    }

    // publics
    x
    y
    constructor(x,y) {
        if (x > 100 || y > 100) {
            Point2d.#printError();
        }
        this.x = x;
        this.y = y;
    }
}

var one = new Point2d(30,400);
// Error: Out of bounds.

```

The `#printError()` static private function here has a `this`, but that's referencing the `Point2d` class, not an instance. As such, the `#errorMsg` and `#printError()` are independent of instances and thus are best as statics. More-

over, there's no reason for them to be accessible outside the class, so they're marked private.

Remember: private statics are similarly not-inherited by subclasses just as private members/methods are not.

Gotcha: Subclassing With Static Privates and `this` Recall that inherited methods, invoked from a subclass, have no trouble accessing (via `this.#whatever` style references) any privates from their own base class:

```
class Point2d {
    // ...

    getID() {
        return this.#ID;
    }

    // ...
}

class Point3d extends Point2d {
    // ...

    printID() {
        console.log(`ID: ${this.getID()}`);
    }
}

var point = new Point3d(3,4,5);
point.printID();
// ID: ..
```

That works just fine.

Unfortunately, and (to me) quite unexpectedly/inconsistently, the same is not true of private statics accessed from inherited public static functions:

```
class Point2d {
    static #errorMsg = "Out of bounds."
    static printError() {
        console.log(`Error: ${this.#errorMsg}`);
    }

    // ...
}

class Point3d extends Point2d {
    // ..
```

```

}

Point2d.printError();
// Error: Out of bounds.

Point3d.printError === Point2d.printError;
// true

Point3d.printError();
// TypeError: Cannot read private member #errorMsg
// from an object whose class did not declare it

```

The `printError()` static is inherited (shared via `[[Prototype]]`) from `Point2d` to `Point3d` just fine, which is why the function references are identical. Like the non-static snippet just above, you might have expected the `Point3d.printError()` static invocation to resolve via the `[[Prototype]]` chain to its original base class (`Point2d`) location, thereby letting it access the base class's `#errorMsg` static private.

But it fails, as shown by the last statement in that snippet. The reason it fails here, but not with the previous snippet, is a convoluted brain twister. I'm not going to dig into the *why* explanation here, frankly because it boils my blood to do so.

There's a *fix*, though. In the static function, instead of `this.#errorMsg`, swap that for `Point2d.#errorMsg`, and now it works:

```

class Point2d {
  static #errorMsg = "Out of bounds."
  static printError() {
    // the fixed reference vvvvvv
    console.log(`Error: ${Point2d.#errorMsg}`);
  }
  // ...
}

class Point3d extends Point2d {
  // ...
}

Point2d.printError();
// Error: Out of bounds.

Point3d.printError();
// Error: Out of bounds. <-- phew, it works now!

```

If public static functions are being inherited, use the class name to access any

private statics instead of using `this.` references. Beware that gotcha!

Class Example

OK, we've laid out a bunch of disparate class features. I want to wrap up this chapter by trying to illustrate a sampling of these capabilities in a single example that's a little less basic/contrived.

```
class CalendarItem {
    static #UNSET = Symbol("unset")
    static #isUnset(v) {
        return v === this.#UNSET;
    }
    static #error(num) {
        return this[`ERROR_${num}`];
    }
    static {
        for (let [idx,msg] of [
            "ID is already set.",
            "ID is unset.",
            "Don't instantiate 'CalendarItem' directly.",
        ].entries()) {
            this[`ERROR_${(idx+1)*100}`] = msg;
        }
    }
    static isSameItem(item1,item2) {
        if (#ID in item1 && #ID in item2) {
            return item1.#ID === item2.#ID;
        }
        else {
            return false;
        }
    }
}

#ID = CalendarItem.#UNSET
#setID(id) {
    if (CalendarItem.#isUnset(this.#ID)) {
        this.#ID = id;
    }
    else {
        throw new Error(CalendarItem.#error(100));
    }
}

description = null
startDateTime = null
```

```

constructor() {
    if (new.target !== CalendarItem) {
        let id = Math.round(Math.random() * 1e9);
        this.#setID(id);
    }
    else {
        throw new Error(CalendarItem.#error(300));
    }
}
getID() {
    if (!CalendarItem.#isUnset(this.#ID)) {
        return this.#ID;
    }
    else {
        throw new Error(CalendarItem.#error(200));
    }
}
getDateTimeStr() {
    if (this.startDateTime instanceof Date) {
        return this.startDateTime.toUTCString();
    }
}
summary() {
    console.log(` ${
        this.getID()
    }) ${
        this.description
    } at ${
        this.getDateTimeStr()
    }`);
}
}

class Reminder extends CalendarItem {
    #complete = false; // <-- no ASI, semicolon needed

    [Symbol.toStringTag] = "Reminder"
    constructor(description,startDateTime) {
        super();

        this.description = description;
        this.startDateTime = startDateTime;
    }
    isComplete() {
        return !!this.#complete;
    }
}

```

```

    }
    markComplete() {
        this.#complete = true;
    }
    summary() {
        if (this.isComplete()) {
            console.log(`(${this.getID()}) Complete.`);
        }
        else {
            super.summary();
        }
    }
}

class Meeting extends CalendarItem {
    #getEndDateTime() {
        if (this.endDateTime instanceof Date) {
            return this.endDateTime.toUTCString();
        }
    }
}

endDateTime = null; // <-- no ASI, semicolon needed

[Symbol.toStringTag] = "Meeting"
constructor(description,startDateTime,endDateTime) {
    super();

    this.description = description;
    this.startDateTime = startDateTime;
    this.endDateTime = endDateTime;
}
getDateTimeStr() {
    return `${super.getDateTimeStr()
} - ${this.#getEndDateTime()
}`;
}
}

```

Take some time to read and digest those `class` definitions. Did you spot most of the `class` features we talked about in this chapter?

NOTE:

One question you may have: why didn't I move the repeated logic of `description` and `startDateTime` setting from both subclass constructors into the single base constructor? This is a nuanced point, but it's not my intention that `CalendarItem` ever be directly instantiated; it's what in class-oriented terms we refer to as an "abstract class". That's why I'm using `new.target` to throw an error if the `CalendarItem` class is ever directly instantiated! So I don't want to imply by signature that the `CalendarItem(..)` constructor should ever be directly used.

Let's now see these three classes in use:

```
var callMyParents = new Reminder(
  "Call my parents to say hi",
  new Date("July 7, 2022 11:00:00 UTC")
);
callMyParents.toString();
// [object Reminder]
callMyParents.summary();
// (586380912) Call my parents to say hi at
// Thu, 07 Jul 2022 11:00:00 GMT
callMyParents.markComplete();
callMyParents.summary();
// (586380912) Complete.
callMyParents instanceof Reminder;
// true
callMyParents instanceof CalendarItem;
// true
callMyParents instanceof Meeting;
// false

var interview = new Meeting(
  "Job Interview: ABC Tech",
  new Date("June 23, 2022 08:30:00 UTC"),
  new Date("June 23, 2022 09:15:00 UTC")
);
interview.toString();
// [object Meeting]
interview.summary();
// (994337604) Job Interview: ABC Tech at Thu,
// 23 Jun 2022 08:30:00 GMT - Thu, 23 Jun 2022
// 09:15:00 GMT
interview instanceof Meeting;
// true
```

```
interview instanceof CalendarItem;  
// true  
interview instanceof Reminder;  
// false  
  
Reminder.isSameItem(callMyParents,callMyParents);  
// true  
Meeting.isSameItem(callMyParents,interview);  
// false
```

Admittedly, some bits of this example are a little contrived. But honestly, I think pretty much all of this is plausible and reasonable usages of the various `class` features.

By the way, there's probably a million different ways to structure the above code logic. I'm by no means claiming this is the *right* or *best* way to do so. As an exercise for the reader, try your hand and writing it yourself, and take note of things you did differently than my approach.

You Don't Know JS Yet: Objects & Classes - 2nd Edition

Chapter 4: This Works

NOTE:

Work in progress

We've seen the `this` keyword used quite a bit so far, but haven't really dug in to understand exactly how it works in JS. It's time we do so.

But to properly understand `this` in JS, you need to set aside any preconceptions you may have, especially assumptions from how `this` works in other programming languages you may have experience in.

Here's the most important thing to understand about `this`: the determination of what value (usually, object) `this` points at is not made at author time, but rather determined at runtime. That means you cannot simply look at a `this`-aware function (even a method in a `class` definition) and know for sure what `this` will hold while that function runs.

Instead, you have to find each place the function is invoked, and look at *how* it's invoked (not even *where* matters). That's the only way to fully answer what `this` will point to.

In fact, a single `this`-aware function can be invoked at least four different ways, and any of those approaches will end up assigning a different `this` for that particular function invocation.

So the typical question we might ask when reading code – “What does `this` point to the function?” – is not actually a valid question. The question you really have to ask is, “When the function is invoked a certain way, what `this` will be assigned for that invocation?”

If your brain is already twisting around just reading this chapter intro... good! Settle in for a rewiring of how you think about `this` in JS.

This Aware

I used the phrase `this`-aware just a moment ago. But what exactly do I mean by that?

Any function that has a `this` keyword in it.

If a function does not have `this` in it anywhere, then the rules of how `this` behaves don't affect that function in any way. But if it *does* have even a single `this` in it, then you absolutely cannot determine how the function will behave

without figuring out, for each invocation of the function, what `this` will point to.

It's sort of like the `this` keyword is a placeholder in a template. That placeholder's value-replacement doesn't get determined when we author the code; it gets determined while the code is running.

You might think I'm just playing word games here. Of course, when you write the program, you write out all the calls to each function, so you've already determined what the `this` is going to be when you authored the code, right? Right!?

Not so fast!

First of all, you don't always write all the code that invokes your function(s). Your `this`-aware function(s) might be passed as a callback(s) to some other code, either in your code base, or in a third-party framework/utility, or even inside a native built-in mechanism of the language or environment that's hosting your program.

But even aside from passing functions as callbacks, several mechanisms in JS allow for conditional runtime behaviors to determine which value (again, usually object) will be set for the `this` of a particular function invocation. So even though you may have written all that code, you *at best* will have to mentally execute the different conditions/paths that end up affecting the function invocation.

And why does all this matter?

Because it's not just you, the author of the code, that needs to figure this stuff out. It's *every single reader* of your code, forever. If anyone (even your future self) wants to read a piece of code that defines a `this`-aware function, that inevitably means that, to fully understand and predict its behavior, that person will have to find, read, and understand every single invocation of that function.

This Confuses Me

Now, in fairness, that's already partially true if we consider a function's parameters. To understand how a function is going to work, we need to know what is being passed into it. So any function with at least one parameter is, in a similar sense, *argument-aware* – meaning, what argument(s) is/are passed in and assigned to the parameter(s) of the function.

But with parameters, we often have a bit more of a hint from the function itself what the parameters will do and hold.

We often see the names of the parameters declared right in the function header, which goes a long way to explaining their nature/purpose. And if there are defaults for the parameters, we often see them declared inline with `= whatever` clauses. Moreover, depending on the code style of the author, we may see

in the first several lines of the function a set of logic that applies to these parameters; this could be assertions about the values (disallowed values, etc), or even modifications (type conversion, formatting, etc).

Actually, `this` is very much like a parameter to a function, but it's an implicit parameter rather than an explicit one. You don't see any signal that `this` is going to be used, in the function header anywhere. You have to read the entire function body to see if `this` appears anywhere.

The "parameter" name is always `this`, so we don't get much of a hint as to its nature/purpose from such a general name. In fact, there's historically a lot of confusion of what "this" even is supposed to mean. And we rarely see much if anything done to validate/convert/etc the `this` value applied to a function invocation. In fact, virtually all `this`-aware code I've seen just neatly assumes the `this` "parameter" is holding exactly what value is expected. Talk about **a trap for unexpected bugs!**

So What Is This?

If `this` is an implicit parameter, what's its purpose? What's being passed in?

Hopefully you have already read the "Scope & Closures" book of this series. If not, I strongly encourage you to circle back and read that one once you've finished this one. In that book, I explained at length how scopes (and closures!) work, an especially important characteristic of functions.

Lexical scope (including all the variables closed over) represents a *static* context for the function's lexical identifier references to be evaluated against. It's fixed/static because at author time, when you place functions and variable declarations in various (nested) scopes, those decisions are fixed, and unaffected by any runtime conditions.

By contrast, a different programming language might offer *dynamic* scope, where the context for a function's variable references is not determined by author-time decisions but by runtime conditions. Such a system would be undoubtedly more flexible than static context – though with flexibility often comes complexity.

To be clear: JS scope is always and only lexical and *static* (if we ignore non-strict mode cheats like `eval(...)` and `with`). However, one of the truly powerful things about JS is that it offers another mechanism with similar flexibility and capabilities to *dynamic* scope.

The `this` mechanism is, effectively, *dynamic* context (not scope); it's how a `this`-aware function can be dynamically invoked against different contexts – something that's impossible with closure and lexical scope identifiers!

Why Is This So Implicit?

You might wonder why something as important as a *dynamic* context is handled as an implicit input to a function, rather than being an explicit argument passed

in.

That's a very important question, but it's not one we can quite answer, yet. Hold onto that question though.

Can We Get On With This?

So why have I belabored *this* subject for a couple of pages now? You get it, right!? You're ready to move on.

My point is, you the author of code, and all other readers of the code even years or decades in the future, need to be **this**-aware. That's the choice, the burden, you place on the reading of such code. And yes, that goes for the choice to use **class** (see Chapter 3), as most class methods will be **this**-aware out of necessity.

Be aware of *this this* choice in code you write. Do it intentionally, and do it in such a way as to produce more outcome benefit than burden. Make sure **this** usage in your code *carries its own weight*.

Let me put it *this* way: don't use **this**-aware code unless you really can justify it, and you've carefully weighed the costs. Just because you've seen a lot of code examples slinging around **this** in others' code, doesn't mean that **this** belongs in *this* code you're writing.

The **this** mechanism in JS, paired with **[[Prototype]]** delegation, is an extremely powerful pillar of the language. But as the cliche goes: "with great power comes great responsibility". Anecdotally, even though I really like and appreciate *this* pillar of JS, probably less than 5% of the JS code I ever write uses it. And when I do, it's with restraint. It's not my default, go-to JS capability.

This Is It!

OK, enough of the wordy lecture. You're ready to dive into **this** code, right?

Let's revisit (and extend) **Point2d** from Chapter 3, but just as an object with data properties and functions on it, instead of using **class**:

```
var point = {
  x: null,
  y: null,

  init(x,y) {
    this.x = x;
    this.y = y;
  },
  rotate(angleRadians) {
    var rotatedX = this.x * Math.cos(angleRadians) -
      this.y * Math.sin(angleRadians);
```

```

    var rotatedY = this.x * Math.sin(angleRadians) +
      this.y * Math.cos(angleRadians);
    this.x = rotatedX;
    this.y = rotatedY;
  },
  toString() {
    return `(${this.x},${this.y})`;
  },
};

```

As you can see, the `init(..)`, `rotate(..)`, and `toString()` functions are `this`-aware. You might be in the habit of assuming that the `this` reference will obviously always hold the `point` object. But that's not guaranteed in any way.

Keep reminding yourself as you go through the rest of this chapter: the `this` value for a function is determined by *how* the function is invoked. That means you can't look at the function's definition, nor where the function is defined (not even the enclosing `class!`). In fact, it doesn't even matter where the function is called from.

We only need to look at *how* the functions are called; that's the only factor that matters.

Implicit Context Invocation

Consider this call:

```
point.init(3,4);
```

We're invoking the `init(..)` function, but notice the `point.` in front of it? This is an *implicit context* binding. It says to JS: invoke the `init(..)` function with `this` referencing `point`.

That is the *normal* way we'd expect a `this` to work, and that's also one of the most common ways we invoke functions. So the typical invocation gives us the intuitive outcome. That's a good thing!

Default Context Invocation

But what happens if we do this?

```
const init = point.init;
init(3,4);
```

You might assume that we'd get the same outcome as the previous snippet. But that's not how JS `this` assignment works.

The *call-site* for the function is `init(3,4)`, which is different than `point.init(3,4)`. When there's no *implicit context* (`point.`), nor any other kind of `this` assignment mechanism, the *default context* assignment occurs.

What will `this` reference when `init(3,4)` is invoked like that?

It depends.

Uh oh. Depends? That sounds confusing.

Don't worry, it's not as bad as it sounds. The *default context* assignment depends on whether the code is in strict-mode or not. But thankfully, virtually all JS code these days is running in strict-mode; for example, ESM (ES Modules) always run in strict-mode, as does code inside a `class` block. And virtually all transpiled JS code (via Babel, TypeScript, etc) is written to declare strict-mode.

So almost all of the time, modern JS code will be running in strict-mode, and thus the *default assignment* context won't "depend" on anything; it's pretty straightforward: `undefined`. That's it!

NOTE:

Keep in mind: `undefined` does not mean "not defined"; it means, "defined with the special empty `undefined` value". I know, I know... the name and meaning are mismatched. That's language legacy baggage, for you. (shrugging shoulders)

That means `init(3,4)`, if run in strict-mode, would throw an exception. Why? Because the `this.x` reference in `init(..)` is a `.x` property access on `undefined` (i.e., `undefined.x`), which is not allowed:

```
"use strict";

var point = { /* .. */ };

const init = point.init;
init(3,4);
// TypeError: Cannot set properties of
// undefined (setting 'x')
```

Stop for a moment and consider: why would JS choose to default the context to `undefined`, so that any *default context* invocation of a `this`-aware function will fail with such an exception?

Because a `this`-aware function **always needs a `this`**. The invocation `init(3,4)` isn't providing a `this`, so that is a mistake, and *should* raise an exception so the mistake can be corrected. The lesson: never invoke a `this`-aware function without providing it a `this`!

Just for completeness sake: in the less common non-strict mode, the *default context* is the global object – JS defines it as `globalThis`, which in browser JS is essentially an alias to `window`, and in Node it's `global`. So, when `init(3,4)` runs in non-strict mode, the `this.x` expression is `globalThis.x` – also known

as `window.x` in the browser, or `global.x` in Node. Thus, `globalThis.x` gets set as 3 and `globalThis.y` gets set as 4.

```
// no strict-mode here, beware!
```

```
var point = { /* .. */ };

const init = point.init;
init(3,4);

globalThis.x;    // 3
globalThis.y;    // 4
point.x;         // null
point.y;         // null
```

That's unfortunate, because it's almost certainly *not* the intended outcome. Not only is it bad if it's a global variable, but it's also *not* changing the property on our `point` object, so program bugs are guaranteed.

WARNING:

Ouch! Nobody wants accidental global variables implicitly created from all over the code. The lesson: always make sure your code is running in strict-mode!

Explicit Context Invocation

Functions can alternately be invoked with *explicit context*, using the built-in `call(..)` or `apply(..)` utilities:

```
var point = { /* .. */ };

const init = point.init;

init.call( point, 3, 4 );
// or: init.apply( point, [ 3, 4 ] )

point.x;         // 3
point.y;         // 4
```

`init.call(point,3,4)` is effectively the same as `point.init(3,4)`, in that both of them assign `point` as the `this` context for the `init(..)` invocation.

NOTE:

Both `call(..)` and `apply(..)` utilities take as their first argument a `this` context value; that's almost always an object, but can technically be any value (number, string, etc). The `call(..)` utility takes subsequent arguments and passes them through to the invoked function, whereas `apply(..)` expects its second argument to be an array of values to pass as arguments.

It might seem awkward to contemplate invoking a function with the *explicit context* assignment (`call(..)` / `apply(..)`) style in your program. But it's more useful than might be obvious at first glance.

Let's recall the original snippet:

```
var point = {
  x: null,
  y: null,

  init(x,y) {
    this.x = x;
    this.y = y;
  },
  rotate(angleRadians) { /* .. */ },
  toString() {
    return `(${this.x},${this.y})`;
  },
};

point.init(3,4);

var anotherPoint = {};
point.init.call( anotherPoint, 5, 6 );

point.x;           // 3
point.y;           // 4
anotherPoint.x;    // 5
anotherPoint.y;    // 6
```

Are you seeing what I did there?

I wanted to define `anotherPoint`, but I didn't want to repeat the definitions of those `init(..)` / `rotate(..)` / `toString()` functions from `point`. So I "borrowed" a function reference, `point.init`, and explicitly set the empty object `anotherPoint` as the `this` context, via `call(..)`.

When `init(..)` is running at that moment, `this` inside it will reference `anotherPoint`, and that's why the `x` / `y` properties (values 5 / 6, respectively) get set there.

Any `this`-aware functions can be borrowed like this: `point.rotate.call(anotherPoint, ..)`, `point.toString.call(anotherPoint)`.

Revisiting Implicit Context Invocation Another approach to share behavior between `point` and `anotherPoint` would have been:

```
var point = { /* .. */ };

var anotherPoint = {
  init: point.init,
  rotate: point.rotate,
  toString: point.toString,
};

anotherPoint.init(5,6);

anotherPoint.x;           // 5
anotherPoint.y;           // 6
```

This is another way of “borrowing” the functions, by adding shared references to the functions on any target object (e.g., `anotherPoint`). The call-site invocation `anotherPoint.init(5,6)` is the more natural/ergonomic style that relies on *implicit context* assignment.

It may seem this approach is a little cleaner, comparing `anotherPoint.init(5,6)` to `point.init.call(anotherPoint,5,6)`.

But the main downside is having to modify any target object with such shared function references, which can be verbose, manual, and error-prone. Sometimes such an approach is acceptable, but many other times, *explicit context* assignment with `call(..)` / `apply(..)` is more preferable.

New Context Invocation

We’ve so far seen three different ways of context assignment at the function call-site: *default*, *implicit*, and *explicit*.

A fourth way to call a function, and assign the `this` for that invocation, is with the `new` keyword:

```
var point = {
  // ..

  init: function() { /* .. */ }

  // ..
};

var anotherPoint = new point.init(3,4);
```

```
anotherPoint.x;      // 3
anotherPoint.y;      // 4
```

TIP:

This example has a bit of nuance to be explained. The `init: function() { ... }` form shown here – specifically, a function expression assigned to a property – is required for the function to be validly called with the `new` keyword. From previous snippets, the concise method form of `init() { ... }` defines a function that *cannot* be called with `new`.

You've typically seen `new` used with `class` for creating instances. But as an underlying mechanism of the JS language, `new` is not inherently a `class` operation.

In a sense, the `new` keyword hijacks a function and forces its behavior into a different mode than a normal invocation. Here are the 4 special steps that JS performs when a function is invoked with `new`:

1. create a brand new empty object, out of thin air.
2. link the `[[Prototype]]` of that new empty object to the function's `.prototype` object (see Chapter 2).
3. invoke the function with the `this` context set to that new empty object.
4. if the function doesn't return its own object value explicitly (with a `return ...` statement), assume the function call should instead return the new object (from steps 1-3).

WARNING:

Step 4 implies that if you `new` invoke a function that *does* return its own object – like `return { ... }`, etc – then the new object from steps 1-3 is *not* returned. That's a tricky gotcha to be aware of, in that it effectively discards that new object before the program has a chance to receive and store a reference to it. Essentially, `new` should never be used to invoke a function that has explicit `return ...` statement(s) in it.

To understand these 4 `new` steps more concretely, I'm going to illustrate them in code, as an alternate to using the `new` keyword:

```
// alternative to:
//   var anotherPoint = new point.init(3,4)

var anotherPoint;
// this is a bare block to hide local
```

```

// `let` declarations
{
  // (Step 1)
  let tmpObj = {};

  // (Step 2)
  Object.setPrototypeOf(
    tmpObj, point.init.prototype
  );
  // or: tmpObj.__proto__ = point.init.prototype

  // (Step 3)
  let res = point.init.call(tmpObj, 3, 4);

  // (Step 4)
  anotherPoint = (
    typeof res !== "object" ? tmpObj : res
  );
}

```

Clearly, the `new` invocation streamlines that set of manual steps!

TIP:

The `Object.setPrototypeOf(..)` in step 2 could also have been done via the `__proto__` property, such as `tmpObj.__proto__ = point.init.prototype`, or even as part of the object literal (step 1) with `tmpObj = { __proto__: point.init.prototype }`.

Skipping some of the formality of these steps, let's recall an earlier snippet and see how `new` approximates a similar outcome:

```

var point = { /* .. */ };

// this approach:
var anotherPoint = {};
point.init.call(anotherPoint, 5, 6);

// can instead be approximated as:
var yetAnotherPoint = new point.init(5, 6);

```

That's a bit nicer! But there's a caveat here.

Using the other functions that `point` holds against `anotherPoint` / `yetAnotherPoint`, we won't want to do with `new`. Why? Because `new` is creating a *new* object, but that's not what we want if we intend to invoke a function against an existing object.

Instead, we'll likely use *explicit context* assignment:

```
point.rotate.call( anotherPoint, /*angleRadians=*/Math.PI );  
  
point.toString.call( yetAnotherPoint );  
// (5,6)
```

Review This

We've seen four rules for `this` context assignment in function calls. Let's put them in order of precedence:

1. Is the function invoked with `new`, creating and setting a *new this*?
2. Is the function invoked with `call(..)` or `apply(..)`, *explicitly* setting `this`?
3. Is the function invoked with an object reference at the call-site (e.g., `point.init(..)`), *implicitly* setting `this`?
4. If none of the above... are we in non-strict mode? If so, *default* the `this` to `globalThis`. But if in strict-mode, *default* the `this` to `undefined`.

These rules, *in this order*, are how JS determines the `this` for a function invocation. If multiple rules match a call-site (e.g., `new point.init.call(..)`), the first rule from the list to match wins.

That's it, you're now master over the `this` keyword. Well, not quite. There's a bunch more nuance to cover. But you're well on your way!

An Arrow Points Somewhere

Everything I've asserted so far about `this` in functions, and how its determined based on the call-site, makes one giant assumption: that you're dealing with a *regular* function (or method).

So what's an *irregular* function?!? It looks like this:

```
const x = x => x <= x;
```

NOTE:

Yes, I'm being a tad sarcastic and unfair to call an arrow function "irregular" and to use such a contrived example. It's a joke, ok?

Here's a real example of an `=>` arrow function:

```
const clickHandler = evt =>  
  evt.target.matches("button") ?  
    this.theFormElem.submit() :  
    evt.stopPropagation();
```

For comparison sake, let me also show the non-arrow equivalent:

```
const clickHandler = function(evt) {
  evt.target.matches("button") ?
    this.theFormElem.submit() :
    evt.stopPropagation();
};
```

Or if we went a bit old-school about it – this is my jam! – we could try the standalone function declaration form:

```
function clickHandler(evt) {
  evt.target.matches("button") ?
    this.theFormElem.submit() :
    evt.stopPropagation();
}
```

Or if the function appeared as a method in a `class` definition, or as a concise method in an object literal, it would look like this:

```
// ...
clickHandler(evt) {
  evt.target.matches("button") ?
    this.theFormElem.submit() :
    evt.stopPropagation();
}
```

What I really want to focus on is how each of these forms of the function will behave with respect to their `this` reference, and whether the first `=>` form differs from the others (hint: it does!). But let's start with a little quiz to see if you've been paying attention.

For each of those function forms just shown, how do we know what each `this` will reference?

Where's The Call-site?

Hopefully, you responded with something like: “first, we need to see how the functions are called.”

Fair enough.

Let's say our program looks like this:

```
var infoForm = {
  theFormElem: null,
  theSubmitBtn: null,

  init() {
    this.theFormElem =
      document.getElementById("the-info-form");
```

```

    this.theSubmitBtn =
        theFormElem.querySelector("button[type=submit]");

    // is *this* the call-site?
    this.theSubmitBtn.addEventListener(
        "click",
        this.clickHandler,
        false
    );
}

// ...
}

```

Ah, interesting. Half of you readers have never seen actual DOM API code like `getElementById(..)`, `querySelector(..)`, and `addEventListener(..)` before. I heard the confusion bells whistle just now!

NOTE:

Sorry, I'm dating myself, here. I've been doing this stuff long enough that I remember when we did that kind of code long before we had utilities like jQuery cluttering up the code with `$` everywhere. And after many years of front-end evolution, we seem to have landed somewhere quite a bit more "modern" – at least, that's the prevailing presumption.

I'm guessing many of you these days are used to seeing component-framework code (React, etc) somewhat like this:

```

// ...

infoForm(props) {
    return (
        <form ref={this.theFormElem}>
            <button type=submit onClick={this.clickHandler}>
                Click Me
            </button>
        </form>
    );
}

// ...

```

Of course, there's a bunch of other ways that code might be shaped, depending on if you're using one framework or another, etc.

Or maybe you're not even using `class / this` style components anymore, because you've moved everything to hooks and closures. In any case, for our

discussion purposes, *this* chapter is all about **this**, so we need to stick to a coding style like the above, to have code related to the discussion.

And neither of those two previous code snippets show the `clickHandler` function being defined. But I've said repeatedly so far, that doesn't matter; all that matters is ... what? say it with me... all that matters is *how* the function is invoked.

So how is `clickHandler` being invoked? What's the call-site, and which context assignment rule does it match?

Hidden From Sight

If you're stuck, don't worry. I'm deliberately making this difficult, to point something very important out.

When the "click" or `onClick=` event handler bindings happen, in both cases, we specified `this.clickHandler`, which implies that there is a **this** context object with a property on it called `clickHandler`, which is holding our function definition.

So, is `this.clickHandler` the call-site? If it was, what assignment rule applies? The *implicit context* rule (#3)?

Unfortunately, no.

The problem is, **we cannot actually see the call-site** in this program. Uh oh.

If we can't see the call-site, how do we know *how* the function is going to actually get called?

That's the exact point I'm making.

It doesn't matter that we passed in `this.clickHandler`. That is merely a reference to a function object value. It's not a call-site.

Under the covers, somewhere inside a framework, library, or even the JS environment itself, when a user clicks the button, a reference to the `clickHandler(..)` function is going to be invoked. And as we've implied, that call-site is even going to pass in the DOM event object as the `evt` argument.

Since we can't see the call-site, we have to *imagine* it. Might it look like...?

```
// ...
eventCallback( domEventObj );
// ...
```

If it did, which **this** rule would apply? The *default context* rule (#4)?

Or, what if the call-site looked like this...?

```
// ...
eventCallback.call( domElement, domEventObj );
```

Now which `this` rule would apply? The *explicit context* rule (#2)?

Unless you open and view the source code for the framework/library, or read the documentation/specification, you won't *know* what to expect of that call-site. Which means that predicting, ultimately, what `this` points to in the `clickHandler` function you write, is... to put it mildly... a bit convoluted.

This Is Wrong

To spare you any more pain here, I'll cut to the chase.

Pretty much all implementations of a click-handler mechanism are going to do something like the `.call(..)`, and they're going to set the DOM element (e.g., button) the event listener is bound to, as the *explicit context* for the invocation.

Hmmm... is that ok, or is that going to be a problem?

Recall that our `clickHandler(..)` function is `this`-aware, and that its `this.theFormElem` reference implies referencing an object with a `theFormElem` property, which in turn is pointing at the parent `<form>` element. DOM buttons do not, by default, have a `theFormElem` property on them.

In other words, the `this` reference that our event handler will have set for it is almost certainly wrong. Oops.

Unless we want to rewrite the `clickHandler` function, we're going to need to fix that.

Fixing `this`

Let's consider some options to address the mis-assignment. To keep things focused, I'll stick to this style of event binding for the discussion:

```
this.submitBtn.addEventListener(
  "click",
  this.clickHandler,
  false
);
```

Here's one way to address it:

```
// store a fixed reference to the current
// `this` context
var context = this;

this.submitBtn.addEventListener(
  "click",
  function handler(evt){
```

```
        return context.clickHandler(evt);
    },
    false
);
```

TIP:

Most older JS code that uses this approach will say something like `var self = this` instead of the `context` name I'm giving it here. "Self" is a shorter word, and sounds cooler. But it's also entirely the wrong semantic meaning. The `this` keyword is not a "self" reference to the function, but rather the context for that current function invocation. Those may seem like the same thing at a glance, but they're completely different concepts, as different as apples and a Beatles song. So... to paraphrase them, "Hey developer, don't make it bad. Take a sad `self` and make it better `context`."

What's going on here? I recognized that the enclosing code, where the `addEventListener` call is going to run, has a current `this` context that is correct, and we need to ensure that same `this` context is applied when `clickHandler(..)` gets invoked.

I defined a surrounding function (`handler(..)`) and then forced the call-site to look like:

```
context.clickHandler(evt);
```

TIP:

Which `this` context assignment rule is applied here? That's right, the *implicit context* rule (#3).

Now, it doesn't matter what the internal call-site of the library/framework/environment looks like. But, why?

Because we're now *actually* in control of the call-site. It doesn't matter how `handler(..)` gets invoked, or what its `this` is assigned. It only matters than when `clickHandler(..)` is invoked, the `this` context is set to what we wanted.

I pulled off that trick not only by defining a surrounding function (`handler(..)`) so I can control the call-site, but... and this is important, so don't miss it... I defined `handler(..)` as a NON-`this`-aware function! There's no `this` keyword inside of `handler(..)`, so whatever `this` gets set (or not) by the library/framework/environment, is completely irrelevant.

The `var context = this` line is critical to the trick. It defines a lexical variable `context`, which is not some special keyword, holding a snapshot of the value in the outer `this`. Then inside `clickHandler`, we merely reference a lexical variable (`context`), no relative/magic `this` keyword.

Lexical This

The name for this pattern, by the way, is “lexical this”, meaning a `this` that behaves like a lexical scope variable instead of like a dynamic context binding.

But it turns out JS has an easier way of performing the “lexical this” magic trick. Are you ready for the trick reveal!?

...

The `=>` arrow function! Tada!

That’s right, the `=>` function is, unlike all other function forms, special, in that it’s not special at all. Or, rather, that it doesn’t define anything special for `this` behavior whatsoever.

In an `=>` function, the `this` keyword... **is not a keyword**. It’s absolutely no different from any other variable, like `context` or `happyFace` or `foobarbaz`.

Let me illustrate *this* point more directly:

```
function outer() {
  console.log(this.value);

  // define a return an "inner"
  // function
  var inner = () => {
    console.log(this.value);
  };

  return inner;
}

var one = {
  value: 42,
};

var two = {
  value: "sad face",
};

var innerFn = outer.call(one);
// 42

innerFn.call(two);
// 42  <-- not "sad face"
```

The `innerFn.call(two)` would, for any *regular* function definition, have resulted in “sad face” here. But since the `inner` function we defined and returned (and assigned to `innerFn`) was an *irregular* `=>` arrow function, it has no special `this` behavior, but instead has “lexical this” behavior.

When the `innerFn(..)` (aka `inner(..)`) function is invoked, even with an *explicit context* assignment via `.call(..)`, that assignment is ignored.

NOTE:

I'm not sure why `=>` arrow functions even have a `call(..)` / `apply(..)` on them, since they are silent no-op functions. I guess it's for consistency with normal functions. But as we'll see later, there are other inconsistencies between *regular* functions and *irregular* `=>` arrow functions.

When a `this` is encountered (`this.value`) inside an `=>` arrow function, `this` is treated like a normal lexical variable, not a special keyword. And since there is no `this` variable in that function itself, JS does what it always does with lexical variables: it goes up one level of lexical scope – in this case, to the surrounding `outer(..)` function, and it checks to see if there's any registered `this` in that scope.

Luckily, `outer(..)` is a *regular* function, which means it has a normal `this` keyword. And the `outer.call(one)` invocation assigned `one` to its `this`.

So, `innerFn.call(two)` is invoking `inner()`, but when `inner()` looks up a value for `this`, it gets... `one`, not `two`.

Back To The... Button You thought I was going to make a pun joke and say “future” there, didn’t you!?

A more direct and appropriate way of solving our earlier issue, where we had done `var context = this` to get a sort of faked “lexical this” behavior, is to use the `=>` arrow function, since its primary design feature is... “lexical this”.

```
this.submitBtn.addEventListener(  
  "click",  
  evt => this.clickHandler(evt),  
  false  
)
```

Boom! Problem solved! Mic drop!

Hear me on *this*: the `=>` arrow function is *not* – I repeat, *not* – about typing fewer characters. The primary point of the `=>` function being added to JS was to give us “lexical this” behavior without having to resort to `var context = this` (or worse, `var self = this`) style hacks.

TIP:

If you need “lexical this”, always prefer an `=>` arrow function. If you don’t need “lexical this”, well... the `=>` arrow function might not be the best tool for the job.

Confession Time I’ve said all along in this chapter, that how you write a function, and where you write the function, has *nothing* to do with how its `this` will be assigned.

For regular functions, that’s true. But when we consider an irregular `=>` arrow function, it’s not entirely accurate anymore.

Recall the original `=>` form of `clickHandler` from earlier in the chapter?

```
const clickHandler = evt =>
  evt.target.matches("button") ?
    this.theFormElem.submit() :
    evt.stopPropagation();
```

If we use that form, in the same context as our event binding, it could look like this:

```
const clickHandler = evt =>
  evt.target.matches("button") ?
    this.theFormElem.submit() :
    evt.stopPropagation();

this.submitBtn.addEventListener("click", clickHandler, false);
```

A lot of developers prefer to even further reduce it, to an inline `=>` arrow function:

```
this.submitBtn.addEventListener(
  "click",
  evt => evt.target.matches("button") ?
    this.theFormElem.submit() :
    evt.stopPropagation(),
  false
);
```

When we write an `=>` arrow function, we know for sure that its `this` binding will exactly be the current `this` binding of whatever surrounding function is running, regardless of what the call-site of the `=>` arrow function looks like. So in other words, *how* we wrote the `=>` arrow function, and *where* we wrote it, does matter.

That doesn’t fully answer the `this` question, though. It just shifts the question to *how the enclosing function was invoked*. Actually, the focus on the call-site is still the only thing that matters.

But the nuance I'm confessing to having omitted until *this* moment is: it matters *which* call-site we consider, not just *any* call-site in the current call stack. The call-site that matters is, the nearest function-invocation in the current call stack *that actually assigns a this context*.

Since an => arrow function never has a **this**-assigning call-site (no matter what), that call-site isn't relevant to the question. We have to keep stepping up the call stack until we find a function invocation that *is this*-assigning – even if such invoked function is not itself **this**-aware.

THAT is the only call-site that matters.

Find The Right Call-Site Let me illustrate, with a convoluted mess of a bunch of nested functions/calls:

```
globalThis.value = { result: "Sad face" };

function one() {
  function two() {
    var three = {
      value: { result: "Hmmm" },
      fn: () => {
        const four = () => this.value;
        return four.call({
          value: { result: "OK", },
        });
      },
    };
    return three.fn();
  };
  return two();
}

new one();           // ???
```

Can you run through that (nightmare) in your head and determine what will be returned from the `new one()` invocation?

It could be any of these:

```
// from `four.call(..)`:
{ result: "OK" }

// or, from `three` object:
{ result: "Hmmm" }

// or, from the `globalThis.value`:
```

```

{ result: "Sad face" }

// or, empty object from the `new` call:
{}

```

The call-stack for that `new one()` invocation is:

```

four      |
three.fn |
two       | (this = globalThis)
one      | (this = {})
[ global ] | (this = globalThis)

```

Since `four()` and `fn()` are both \Rightarrow arrow functions, the `three.fn()` and `four.call(..)` call-sites are not `this`-assigning; thus, they're irrelevant for our query. What's the next invocation to consider in the call-stack? `two()`. That's a regular function (it can accept `this`-assignment), and the call-site matches the *default context* assignment rule (#4). Since we're not in strict-mode, `this` is assigned `globalThis`.

When `four()` is running, `this` is just a normal variable. It looks then to its containing function (`three.fn()`), but it again finds a function with no `this`. So it goes up another level, and finds a `two()` *regular* function that has a `this` defined. And that `this` is `globalThis`. So the `this.value` expression resolves to `globalThis.value`, which returns us... { `result: "Sad face"` }.

...

Take a deep breath. I know that's a lot to mentally process. And in fairness, that's a super contrived example. You'll almost never see all that complexity mixed in one call-stack.

But you absolutely will find mixed call-stacks in real programs. You need to get comfortable with the analysis I just illustrated, to be able to unwind the call-stack until you find the most recent `this`-assigning call-site.

Remember the addage I quoted earlier: “with great power comes great responsibility”. Choosing `this`-oriented code (even `classes`) means choosing both the flexibility it affords us, as well as needing to be comfortable navigating the call-stack to understand how it will behave.

That's the only way to effectively write (and later read!) `this`-aware code.

This Is Bound To Come Up

Backing up a bit, there's another option if you don't want to use an \Rightarrow arrow function's “lexical `this`” behavior to address the button event handler functionality.

In addition to `call(..)` / `apply(..)` – these invoke functions, remember! – JS functions also have a third utility built in, called `bind(..)` – which does *not*

invoke the function, just to be clear.

The `bind(..)` utility defines a *new* wrapped/bound version of a function, where its `this` is preset and fixed, and cannot be overridden with a `call(..)` or `apply(..)`, or even an *implicit context* object at the call-site:

```
this.submitBtn.addEventListener(
  "click",
  this.clickHandler.bind(this),
  false
);
```

Since I'm passing in a `this`-bound function as the event handler, it similarly doesn't matter how that utility tries to set a `this`, because I've already forced the `this` to be what I wanted: the value of `this` from the surrounding function invocation context.

Hardly New This pattern is often referred to as "hard binding", since we're creating a function reference that is strongly bound to a particular `this`. A lot of JS writings have claimed that the `=>` arrow function is essentially just syntax for the `bind(this)` hard-binding. It's not. Let's dig in.

If you were going to create a `bind(..)` utility, it might look kinda like *this*:

```
function bind(fn,context) {
  return function bound(...args){
    return fn.apply(context,args);
  };
}
```

NOTE:

This is not actually how `bind(..)` is implemented. The behavior is more sophisticated and nuanced. I'm only illustrating one portion of its behavior in this snippet.

Does that look familiar? It's using the good ol' fake "lexical this" hack. And under the covers, it's an *explicit context* assignment, in this case via `apply(..)`.

So wait... doesn't that mean we could just do it with an `=>` arrow function?

```
function bind(fn,context) {
  return (...args) => fn.apply(context,args);
}
```

Eh... not quite. As with most things in JS, there's a bit of nuance. Let me illustrate:

```
// candidate implementation, for comparison
function fakeBind(fn,context) {
```

```

        return (...args) => fn.apply(context, args);
    }

    // test subject
    function thisAwareFn() {
        console.log(`Value: ${this.value}`);
    }

    // control data
    var obj = {
        value: 42,
    };

    // experiment
    var f = thisAwareFn.bind(obj);
    var g = fakeBind(thisAwareFn, obj);

    f();           // Value: 42
    g();           // Value: 42

    new f();       // Value: undefined
    new g();       // <--- ???

```

First, look at the `new f()` call. That's admittedly a strange usage, to call `new` on a hard-bound function. It's probably quite rare that you'd ever do so. But it shows something kind of interesting. Even though `f()` was hard-bound to a `this` context of `obj`, the `new` operator was able to hijack the hard-bound function's `this` and re-bind it to the newly created and empty object. That object has no `value` property, which is why we see "Value: undefined" printed out.

If that feels strange, I agree. It's a weird corner nuance. It's not something you'd likely ever exploit. But I point it out not just for trivia. Refer back to the four rules presented earlier in this chapter. Remember how I asserted their order-of-precedence, and `new` was at the top (#1), ahead of *explicit call(..) / apply(..) assignment rule* (#2)?

Since we can sort of think of `bind(..)` as a variation of that rule, we now see that order-of-precedence proven. `new` is more precedent than, and can override, even a hard-bound function. Sort of makes you think the hard-bound function is maybe not so "hard"-bound, huh?!

But... what's going to happen with the `new g()` call, which is invoking `new` on the returned `=>` arrow function? Do you predict the same outcome as `new f()`?

Sorry to disappoint.

That line will actually throw an exception, because an `=>` function cannot be used with the `new` keyword.

But why? My best answer, not being authoritative on TC39 myself, is that conceptually and actually, an `=>` arrow function is not a function with a hard-bound `this`, it's a function that has no `this` at all. As such, `new` makes no sense against such a function, so JS just disallows it.

NOTE:

Recall earlier, when I pointed out that `=>` arrow functions have `call(..)`, `apply(..)`, and indeed even a `bind(..)`. But we've seen that such functions basically ignore these utilities as no-ops. It's a bit strange, in my opinion, that `=>` arrow functions have all those utilities as pass-through no-ops, but for the `new` keyword, that's not just, again, a no-op pass-through, but rather disallowed with an exception.

But the main point is: an `=>` arrow function is *not* a syntactic form of `bind(this)`.

Losing This Battle

Returning once again to our button event handler example:

```
this.submitBtn.addEventListener(  
  "click",  
  this.clickHandler,  
  false  
)
```

There's a deeper concern we haven't yet addressed.

We've seen several different approaches to construct a different callback function reference to pass in there, in place of `this.clickHandler`.

But whichever of those ways we choose, they are producing literally a different function, not just an in-place modification to our existing `clickHandler` function.

Why does that matter?

Well, first of all, the more functions we create (and re-create), the more processing time (very slight) and the more memory (pretty small, usually) we're chewing up. And when we're re-creating a function reference, and throwing an old one away, that's also leaving un-reclaimed memory sitting around, which puts pressure on the garbage collector (GC) to more often, pause the universe of our program momentarily while it cleans up and reclaims that memory.

If hooking up this event listening is a one-time operation, no big deal. But if it's happening over and over again, the system-level performance effects *can* start to add up. Ever had an otherwise smooth animation jitter? That was probably the GC kicking in, cleaning up a bunch of reclaimable memory.

But another concern is, for things like event handlers, if we're going to remove an event listener at some later time, we need to keep a reference to the exact same function we attached originally. If we're using a library/framework, often (but not always!) they take care of that little dirty-work detail for you. But otherwise, it's on us to make sure that whatever function we plan to attach, we hold onto a reference just in case we need it later.

So the point I'm making is: presetting a `this` assignment, no matter how you do it, so that it's predictable, comes with a cost. A system level cost and a program maintenance/complexity cost. It is *never* free.

One way of reacting to that fact is to decide, OK, we're just going to manufacture all those `this`-assigned function references once, ahead of time, up-front. That way, we're sure to reduce both the system pressure, and the code pressure, to a minimum.

Sounds reasonable, right? Not so fast.

Pre-Binding Function Contexts If you have a one-off function reference that needs to be `this`-bound, and you use an `=>` arrow or a `bind(this)` call, I don't see any problems with that.

But if most or all of the `this`-aware functions in a segment of your code invoked in ways where the `this` isn't the predictable context you expect, and so you decide you need to hard-bind them all... I think that's a big warning signal that you're going about things the wrong way.

Please recall the discussion in the "Avoid This" section from Chapter 3, which started with this snippet of code:

```
class Point2d {
  x = null
  getDoubleX = () => this.x * 2

  constructor(x,y) {
    this.x = x;
    this.y = y;
  }
  toString() { /* .. */ }
}

var point = new Point2d(3,4);
```

Now imagine we did this with that code:

```
const getX = point.getDoubleX;

// later, elsewhere
```

```
getX();           // 6
```

As you can see, the problem we were trying to solve is the same as we've been dealing with here in this chapter. It's that we wanted to be able to invoke a function reference like `getX()`, and have that *mean* and *behave like point.getDoubleX()*. But `this` rules on *regular* functions don't work that way.

So we used an `=>` arrow function. No big deal, right!?

Wrong.

The real root problem is that we *want* two conflicting things out of our code, and we're trying to use the same *hammer* for both *nails*.

We want to have a `this`-aware method stored on the `class` prototype, so that there's only one definition for the function, and all our subclasses and instances nicely share that same function. And the way they all share is through the power of the dynamic `this` binding.

But at the same time, we *also* want those function references to magically stay `this`-assinged to our instance when we pass those function references around and other code is in charge of the call-site.

In other words, sometimes we want something like `point.getDoubleX` to mean, "give me a reference that's `this`-assigned to `point`", and other times we want the same expression `point.getDoubleX` to mean, give me a dynamic `this`-assignable function reference so it can properly get the context I need it to at this moment.

Perhaps JS could offer a different operator besides `..`, like `::` or `->` or something like that, which would let you distinguish what kind of function reference you're after. In fact, there's a long-standing proposal for a `this`-binding operator (`:::`), that picks up attention from time to time, and then seems to stall out. Who knows, maybe someday such an operator will finally land, and we'll have better options.

But I strongly suspect that even if it does land someday, it's going to vend a whole new function reference, exactly as the `=>` or `bind(this)` approaches we've already talked about. It won't come as a free and perfect solution. There will always be a tension between wanting the same function to sometimes be `this`-flexible and sometimes be `this`-predictable.

What JS authors of `class`-oriented code often run up against, sooner or later, is this exact tension. And you know what they do?

They don't consider the *costs* of simply pre-binding all the class's `this`-aware methods as instead `=>` arrow functions in member properties. They don't realize that it's completely defeated the entire purpose of the `[[Prototype]]` chain.

And they don't realize that if fixed-context is what they *really need*, there's an entirely different mechanism in JS that is better suited for that purpose.

Take A More Critical Look So when you do this sort of thing:

```
class Point2d {
  x = null
  y = null
  getDoubleX = () => this.x * 2
  toString = () => `(${this.x},${this.y})`

  constructor(x,y) {
    this.x = x;
    this.y = y;
  }
}

var point = new Point2d(3,4);
var anotherPoint = new Point2d(5,6);

var f = point.getDoubleX;
var g = anotherPoint.toString;

f();           // 6
g();           // (5,6)
```

I say, “ick!”, to the hard-bound `this`-aware methods `getDoubleX()` and `toString()` there. To me, that's a code smell. But here's an even *worse* approach that has been favored by many developers in the past:

```
class Point2d {
  x = null
  y = null

  constructor(x,y) {
    this.x = x;
    this.y = y;
    this.getDoubleX = this.getDoubleX.bind(this);
    this.toString = this.toString.bind(this);
  }
  getDoubleX() { return this.x * 2; }
  toString() { return `(${this.x},${this.y})`; }
}

var point = new Point2d(3,4);
var anotherPoint = new Point2d(5,6);
```

```

var f = point.getDoubleX;
var g = anotherPoint.toString;

f();           // 6
g();           // (5,6)

```

Double ick.

In both cases, you're using a `this` mechanism but completely betraying/neutering it, by taking away all the powerful dynamicism of `this`.

You really should at least be contemplating this alternate approach, which skips the whole `this` mechanism altogether:

```

function Point2d(px,py) {
  var x = px;
  var y = py;

  return {
    getDoubleX() { return x * 2; },
    toString() { return `(${x},${y})`; }
  };
}

var point = Point2d(3,4);
var anotherPoint = Point2d(5,6);

var f = point.getDoubleX;
var g = anotherPoint.toString;

f();           // 6
g();           // (5,6)

```

You see? No ugly or complex `this` to clutter up that code or worry about corner cases for. Lexical scope is super straightforward and intuitive.

When all we want is for most/all of our function behaviors to have a fixed and predictable context, the most appropriate solution, the most straightforward and even performant solution, is lexical variables and scope closure.

When you go to all the trouble of sprinkling `this` references all over a piece of code, and then you cut off the whole mechanism at the knees with => “lexical this” or `bind(this)`, you chose to make the code more verbose, more complex, more overwrought. And you got nothing out of it that was more beneficial, except to follow the `this` (and `class`) bandwagon.

...

Deep breath. Collect yourself.

I'm talking to myself, not you. But if what I just said bothers you, I'm talking to you, too!

OK, listen. That's just my opinion. If you don't agree, that's fine. But apply the same level of rigor to thinking about how these mechanisms work, as I have, when you decide what conclusions you want to arrive at.

Variations

Before we close out our lengthy discussion of `this`, there's a few irregular variations on function calls that we should discuss.

Indirect Function Calls

Recall this example from earlier in the chapter?

```
var point = {
  x: null,
  y: null,

  init(x,y) {
    this.x = x;
    this.y = y;
  },
  rotate(angleRadians) { /* .. */ },
  toString() { /* .. */ },
};

var init = point.init;
init(3,4);           // broken!
```

This is broken because the `init(3,4)` call-site doesn't provide the necessary `this`-assignment signal. But there's other ways to observe a similar breakage. For example:

```
(1,point.init)(3,4);           // broken!
```

This strange looking syntax is first evaluating an expression `(1,point.init)`, which is a comma series expression. The result of such an expression is the final evaluated value, which in this case is the function reference (held by `point.init`).

So the outcome puts that function reference onto the expression stack, and then invokes that value with `(3,4)`. That's an indirect invocation of the function. And what's the result? It actually matches the *default context* assignment rule (#4) we looked at earlier in the chapter.

Thus, in non-strict mode, the `this` for the `point.init(..)` call will be `globalThis`. Had we been in strict-mode, it would have been `undefined`, and

the `this.x = x` operation would then have thrown an exception for invalidly accessing the `x` property on the `undefined` value.

There's several different ways to get an indirect function invocation. For example:

```
(()=>point.init)()(3,4);    // broken!
```

And another example of indirect function invocation is the Immediately Invoked Function Expression (IIFE) pattern:

```
(function(){
  // `this` assigned via "default" rule
})();
```

As you can see, the function expression value is put onto the expression stack, and then it's invoked with the `()` on the end.

But what about this code:

```
(point.init)(3,4);
```

What will be the outcome of that code?

By the same reasoning we've seen in the previous examples, it stands to reason that the `point.init` expression puts the function value onto the expression stack, and then invoked indirectly with `(3,4)`.

Not quite, though! JS grammar has a special rule to handle the invocation form `(someIdentifier)(...)` as if it had been `someIdentifier(..)` (without the `(..)` around the identifier name).

Wondering why you might want to ever force the *default context* for `this` assignment via an indirect function invocation?

Accessing `globalThis`

Before we answer that, let's introduce another way of performing indirect function `this` assignment. Thus far, the indirect function invocation patterns shown are sensitive to strict-mode. But what if we wanted an indirect function `this` assignment that doesn't respect strict-mode.

The `Function(..)` constructor takes a string of code and dynamically defines the equivalent function. However, it always does so as if that function had been declared in the global scope. And furthermore, it ensures such function *does not* run in strict-mode, no matter the strict-mode status of the program. That's the same outcome as running an indirect

One niche usage of such strict-mode agnostic indirect function `this` assignment is for getting a reliable reference to the true global object prior to when the JS specification actually defined the `globalThis` identifier (for example, in a polyfill for it):

```
"use strict";

var gt = new Function("return this")();
gt === globalThis; // true
```

In fact, a similar outcome, using the comma operator trick (see previous section) and `eval(..)`:

```
"use strict";

function getGlobalThis() {
    return (1,eval)("this");
}

getGlobalThis() === globalThis; // true
```

NOTE:

`eval("this")` would be sensitive to strict-mode, but `(1,eval)("this")` is not, and therefore reliably gives us the `globalThis` in any program.

Unfortunately, the `new Function(..)` and `(1,eval)(..)` approaches both have an important limitation: that code will be blocked in browser-based JS code if the app is served with certain Content-Security-Policy (CSP) restrictions, disallowing dynamic code evaluation (for security reasons).

Can we get around this? Yes, mostly.¹

The JS specification says that a getter function defined on the global object, or on any object that inherits from it (like `Object.prototype`), runs the getter function with `this` context assigned to `globalThis`, regardless of the program's strict-mode.

```
// Adapted from: https://mathiasbynens.be/notes/globalthis#robust-polyfill
function getGlobalThis() {
    Object.defineProperty(Object.prototype, "__get_globalthis__", {
        get() { return this; },
        configurable: true
    });
    var gt = __get_globalthis__;
    delete Object.prototype.__get_globalthis__;
    return gt;
}

getGlobalThis() === globalThis; // true
```

¹“A horrifying globalThis polyfill in universal JavaScript”; Mathias Bynens; April 18 2019; <https://mathiasbynens.be/notes/globalthis#robust-polyfill> ; Accessed July 2022

Yeah, that's super gnarly. But that's JS `this` for you!

Template Tag Functions

There's one more unusual variation of function invocation we should cover: tagged template functions.

Template strings – what I prefer to call interpolated literals – can be “tagged” with a prefix function, which is invoked with the parsed contents of the template literal:

```
function tagFn(/* .. */) {  
  // ..  
}  
  
tagFn`actually a function invocation!`;
```

As you can see, there's no `(..)` invocation syntax, just the tag function (`tagFn`) appearing before the `template literal`; whitespace is allowed between them, but is very uncommon.

Despite the strange appearance, the function `tagFn(..)` will be invoked. It's passed the list of one or more string literals that were parsed from the template literal, along with any interpolated expression values that were encountered.

We're not going to cover all the ins and outs of tagged template functions – they're seriously one of the most powerful and interesting features ever added to JS – but since we're talking about `this` assignment in function invocations, for completeness sake we need to talk about how `this` will be assigned.

The other form for tag functions you may encounter is:

```
var someObj = {  
  tagFn() { /* .. */ }  
};  
  
someObj.tagFn`also a function invocation!`;
```

Here's the easy explanation: `tagFn`..`` and `someObj.tagFn`..`` will each have `this`-assignment behavior corresponding to call-sites as `tagFn(..)` and `someObj.tagFn(..)`, respectively. In other words, `tagFn`..`` behaves by the *default context* assignment rule (#4), and `someObj.tagFn`..`` behaves by the *implicit context* assignment rule (#3).

Luckily for us, we don't need to worry about the `new` or `call(..) / apply(..)` assignment rules, as those forms aren't possible with tag functions.

It should be pointed out that it's pretty rare for a tagged template literal function to be defined as `this`-aware, so it's fairly unlikely you'll need to apply these rules. But just in case, now you're in the *know*.

Stay Aware

So, that's `this`. I'm willing to bet for many of you, it was a bit more... shall we say, involved... than you might have been expecting.

The good news, perhaps, is that in practice you don't often trip over all these different complexities. But the more you use `this`, the more it requires you, and the readers of your code, to understand how it actually works.

The lesson here is that you should be intentional and aware of all aspects of `this` before you go sprinkling it about your code. Make sure you're using it most effectively and taking full advantage of this important pillar of JS.

You Don't Know JS Yet: Objects & Classes - 2nd Edition

Chapter 5: Delegation

NOTE:

Work in progress

We've thoroughly explored objects, prototypes, classes, and now the `this` keyword. But we're now going to revisit what we've learned so far from a bit of a different perspective.

What if you could leverage all the power of the objects, prototypes, and dynamic `this` mechanisms together, without ever using `class` or any of its descendants?

In fact, I would argue JS is inherently less class-oriented than the `class` keyword might appear. Because JS is a dynamic, prototypal language, its strong suit is actually... *delegation*.

Preamble

Before we begin looking at delegation, I want to offer a word of caution. This perspective on JS's object `[[Prototype]]` and `this` function context mechanisms is *not* mainstream. It's *not* how framework authors and libraries utilize JS. You won't, to my knowledge, find any big apps out there using this pattern.

So why on earth would I devote a chapter to such a pattern, if it's so unpopular?

Good question. The cheeky answer is: because it's my book and I can do what I feel like!

But the deeper answer is, because I think developing *this* understanding of one of the language's core pillars helps you *even if* all you ever do is use `class`-style JS patterns.

To be clear, delegation is not my invention. It's been around as a design pattern for decades. And for a long time, developers argued that prototypal delegation was *just* the dynamic form of inheritance.¹ But I think that was a mistake to conflate the two.²

¹“Treaty of Orlando”; Henry Lieberman, Lynn Andrea Stein, David Ungar; Oct 6, 1987; <https://web.media.mit.edu/~lieber/Publications/Treaty-of-Orlando-Treaty-Text.pdf> ; PDF; Accessed July 2022

²“Classes vs. Prototypes, Some Philosophical and Historical Observations”; Antero Taivalsaari; Apr 22, 1996; <https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.56.4713&rep=rep1&type=pdf> ; PDF; Accessed July 2022

For the purposes of this chapter, I'm going to present delegation, as implemented via JS mechanics, as an alternative design pattern, positioned somewhere between class-orientation and object-closure/module patterns.

The first step is to *de-construct* the `class` mechanism down to its individual parts. Then we'll cherry-pick and mix the parts a bit differently.

What's A Constructor, Anyway?

In Chapter 3, we saw `constructor(..)` as the main entry point for construction of a `class` instance. But the `constructor(..)` doesn't actually do any *creation* work, it's only *initialization* work. In other words, the instance is already created by the time the `constructor(..)` runs and initializes it – e.g., `this.whatever` types of assignments.

So where does the *creation* work actually happen? In the `new` operator. As the section “New Context Invocation” in Chapter 4 explains, there are four steps the `new` keyword performs; the first of those is the creation of a new empty object (the instance). The `constructor(..)` isn't even invoked until step 3 of `new`'s efforts.

But `new` is not the only – or perhaps even, best – way to *create* an object “instance”. Consider:

```
// a non-class "constructor"
function Point2d(x,y) {
  // create an object (1)
  var instance = {};

  // initialize the instance (3)
  instance.x = x;
  instance.y = y;

  // return the instance (4)
  return instance;
}

var point = Point2d(3,4);

point.x;           // 3
point.y;           // 4
```

There's no `class`, just a regular function definition (`Point2d(..)`). There's no `new` invocation, just a regular function call (`Point2d(3,4)`). And there's no `this` references, just regular object property assignments (`instance.x = ..`).

The term that's most often used to refer to this pattern of code is that `Point2d(..)` here is a *factory function*. Invoking it causes the construction

(creation and initialization) of an object, and returns that back to us. That's an extremely common pattern, at least as common as class-oriented code.

I comment-annotated (1), (3), and (4) in that snippet, which roughly correspond to steps 1, 3, and 4 of the `new` operation. But where's step 2?

If you recall, step 2 of `new` is about linking the object (created in step 1) to another object, via its `[[Prototype]]` slot (see Chapter 2). So what object might we want to link our `instance` object to? We could link it to an object that holds functions we'd like to associate/use with our instance.

Let's amend the previous snippet:

```
var prototypeObj = {
  toString() {
    return `(${this.x},${this.y})`;
  },
}

// a non-class "constructor"
function Point2d(x,y) {
  // create an object (1)
  var instance = {
    // link the instance's [[Prototype]] (2)
    __proto__: prototypeObj,
  };

  // initialize the instance (3)
  instance.x = x;
  instance.y = y;

  // return the instance (4)
  return instance;
}

var point = Point2d(3,4);
```

Now you see the `__proto__` assignment that's setting up the internal `[[Prototype]]` linkage, which was the missing step 2. I used the `__proto__` here merely for illustration purposes; using `setPrototypeOf(..)` as shown in Chapter 4 would have accomplished the same task.

New Factory Instance

What do you think would happen if we used `new` to invoke the `Point2d(..)` function as shown here?

```
var anotherPoint = new Point2d(5,6);

anotherPoint.toString();           // (5,6)
```

Wait! What's going on here? A regular, non-`class` factory function is invoked with the `new` keyword, as if it was a `class`. Does that change anything about the outcome of the code?

No... and yes. `anotherPoint` here is exactly the same object as it would have been had I not used `new`. But! The object that `new` creates, links, and assigns as `this` context? *That* object was completely ignored and thrown away, ultimately to be garbage collected by JS. Unfortunately, the JS engine cannot predict that you're not going to use the object that you asked `new` to create, so it always still gets created even if it goes unused.

That's right! Using a `new` keyword against a factory function might *feel* more ergonomic or familiar, but it's quite wasteful, in that it creates **two** objects, and wastefully throws one of them away.

Factory Initialization

In the current code example, the `Point2d(..)` function still looks an awful lot like a normal `constructor(..)` of a `class` definition. But what if we moved the initialization code to a separate function, say named `init(..)`:

```
var prototypeObj = {
  init(x,y) {
    // initialize the instance (3)
    this.x = x;
    this.y = y;
  },
  toString() {
    return `(${this.x},${this.y})`;
  },
}

// a non-class "constructor"
function Point2d(x,y) {
  // create an object (1)
  var instance = {
    // link the instance's [[Prototype]] (2)
    __proto__: prototypeObj,
  };

  // initialize the instance (3)
  instance.init(x,y);

  // return the instance (4)
}
```

```

        return instance;
    }

var point = Point2d(3,4);

point.toString();           // (3,4)

```

The `instance.init(..)` call makes use of the `[[Prototype]]` linkage set up via `__proto__` assignment. Thus, it *delegates* up the prototype chain to `prototypeObj.init(..)`, and invokes it with a `this` context of `instance` – via *implicit context* assignment (see Chapter 4).

Let's continue the deconstruction. Get ready for a switcheroo!

```

var Point2d = {
    init(x,y) {
        // initialize the instance (3)
        this.x = x;
        this.y = y;
    },
    toString() {
        return `(${this.x},${this.y})`;
    },
};

```

Whoa, what!? I discarded the `Point2d(..)` function, and instead renamed the `prototypeObj` as `Point2d`. Weird.

But let's look at the rest of the code now:

```

// steps 1, 2, and 4
var point = { __proto__: Point2d, };

// step 3
point.init(3,4);

point.toString();           // (3,4)

```

And one last refinement: let's use a built-in utility JS provides us, called `Object.create(..)`:

```

// steps 1, 2, and 4
var point = Object.create(Point2d);

// step 3
point.init(3,4);

point.toString();           // (3,4)

```

What operations does `Object.create(..)` perform?

1. create a brand new empty object, out of thin air.
2. link the `[[Prototype]]` of that new empty object to the function's `.prototype` object.

If those look familiar, it's because those are exactly the same first two steps of the `new` keyword (see Chapter 4).

Let's put this back together now:

```
var Point2d = {
  init(x,y) {
    this.x = x;
    this.y = y;
  },
  toString() {
    return `(${this.x},${this.y})`;
  },
};

var point = Object.create(Point2d);

point.init(3,4);

point.toString();           // (3,4)
```

Hmmm. Take a few moments to ponder what's been derived here. How does it compare to the `class` approach?

This pattern ditches the `class` and `new` keywords, but accomplishes the exact same outcome. The *cost*? The single `new` operation was broken up into two statements: `Object.create(Point2d)` and `point.init(3,4)`.

Help Me Reconstruct! If having those two operations separate bothers you – is it *too deconstructed*!? – they can always be recombined in a little factory helper:

```
function make(objType,...args) {
  var instance = Object.create(objType);
  instance.init(...args);
  return instance;
}

var point = make(Point2d,3,4);

point.toString();           // (3,4)
```

TIP:

Such a `make(..)` factory function helper works generally for any object-type, as long as you follow the implied convention that each `objType` you link to has a function named `init(..)` on it.

And of course, you can still create as many instances as you'd like:

```
var point = make(Point2d, 3, 4);  
  
var anotherPoint = make(Point2d, 5, 6);
```

Ditching Class Thinking

Quite frankly, the *deconstruction* we just went through only ends up in slightly different, and maybe slightly better or slightly worse, code as compared to the `class` style. If that's all delegation was about, it probably wouldn't even be useful enough for more than a footnote, much less a whole chapter.

But here's where we're going to really start pushing the class-oriented thinking itself, not just the syntax, aside.

Class-oriented design inherently creates a hierarchy of *classification*, meaning how we divide up and group characteristics, and then stack them vertically in an inheritance chain. Moreover, defining a subclass is a specialization of the generalized base class. Instantiating is a specialization of the generalized class.

Behavior in a traditional class hierarchy is a vertical composition through the layers of the inheritance chain. Attempts have been made over the decades, and even become rather popular at times, to flatten out deep hierarchies of inheritance, and favor a more horizontal composition through *mixins* and related ideas.

I'm not asserting there's anything wrong with those ways of approaching code. But I am saying that they aren't *naturally* how JS works, so adopting them in JS has been a long, winding, complicated road, and has variously accreted lots of nuanced syntax to retrofit on top of JS's core `[[Prototype]]` and `this` pillar.

For the rest of this chapter, I intend to discard both the syntax of `class` and the thinking of *class*.

Delegation Illustrated

So what is delegation about? At its core, it's about two or more *things* sharing the effort of completing a task.

Instead of defining a `Point2d` general parent *thing* that represents shared behavior that a set of one or more child `point` / `anotherPoint` *things* inherit

from, delegation moves us to building our program with discrete peer *things* that cooperate with each other.

I'll sketch that out in some code:

```
var Coordinates = {
  setX(x) {
    this.x = x;
  },
  setY(y) {
    this.y = y;
  },
  setXY(x,y) {
    this.setX(x);
    this.setY(y);
  },
};

var Inspect = {
  toString() {
    return `(${this.x},${this.y})`;
  },
};

var point = {};

Coordinates.setXY.call(point,3,4);
Inspect.toString.call(point); // (3,4)

var anotherPoint = Object.create(Coordinates);

anotherPoint.setXY(5,6);
Inspect.toString.call(anotherPoint); // (5,6)
```

Let's break down what's happening here.

I've defined `Coordinates` as a concrete object that holds some behaviors I associate with setting point coordinates (`x` and `y`). I've also defined `Inspect` as a concrete object that holds some debug inspection logic, such as `toString()`.

I then create two more concrete objects, `point` and `anotherPoint`.

`point` has no specific `[[Prototype]]` (default: `Object.prototype`). Using *explicit context* assignment (see Chapter 4), I invoke the `Coordinates.setXY(...)` and `Inspect.toString()` utilities in the context of `point`. That is what I call *explicit delegation*.

`anotherPoint` is `[[Prototype]]` linked to `Coordinates`, mostly for a bit of convenience. That lets me use *implicit context* assignment with

`anotherPoint.setXY(..)`. But I can still *explicitly* share `anotherPoint` as context for the `Inspect.toString()` call. That's what I call *implicit delegation*.

Don't miss `this`: We still accomplished composition: we composed the behaviors from `Coordinates` and `Inspect`, during runtime function invocations with `this` context sharing. We didn't have to author-combine those behaviors into a single `class` (or base-subclass `class` hierarchy) for `point` / `anotherPoint` to inherit from. I like to call this runtime composition, **virtual composition**.

The `point` here is: none of these four objects is a parent or child. They're all peers of each other, and they all have different purposes. We can organize our behavior in logical chunks (on each respective object), and share the context via `this` (and, optionally `[[Prototype]]` linkage), which ends up with the same composition outcomes as the other patterns we've examined thus far in the book.

That is the heart of the **delegation** pattern, as JS embodies it.

TIP:

In the first edition of this book series, this book ("this & Object Prototypes") coined a term, "OLOO", which stands for "Objects Linked to Other Objects" – to stand in contrast to "OO" ("Object Oriented"). In this preceding example, you can see the essence of OLOO: all we have are objects, linked to and cooperating with, other objects. I find this beautiful in its simplicity.

Composing Peer Objects

Let's take *this delegation* even further.

In the preceding snippet, `point` and `anotherPoint` merely held data, and the behaviors they delegated to were on other objects (`Coordinates` and `Inspect`). But we can add behaviors directly to any of the objects in a delegation chain, and those behaviors can even interact with each other, all through the magic of *virtual composition* (`this` context sharing).

To illustrate, we'll evolve our current `point` example a fair bit. And as a bonus we'll actually draw our points on a `<canvas>` element in the DOM. Let's take a look:

```
var Canvas = {
  setOrigin(x,y) {
    this.ctx.translate(x,y);

    // flip the canvas context vertically,
    // so coordinates work like on a normal
    // 2d (x,y) graph
```

```

        this.ctx.scale(1,-1);
},
pixel(x,y) {
    this.ctx.fillRect(x,y,1,1);
},
renderScene() {
    // clear the canvas
    var matrix = this.ctx.getTransform();
    this.ctx.resetTransform();
    this.ctx.clearRect(
        0, 0,
        this.ctx.canvas.width,
        this.ctx.canvas.height
    );
    this.ctx.setTransform(matrix);

    this.draw(); // <-- where is draw()?
},
};

var Coordinates = {
    setX(x) {
        this.x = Math.round(x);
    },
    setY(y) {
        this.y = Math.round(y);
    },
    setXY(x,y) {
        this.setX(x);
        this.setY(y);
        this.render(); // <-- where is render()?
    },
};

var ControlPoint = {
    // delegate to Coordinates
    __proto__: Coordinates,

    // NOTE: must have a <canvas id="my-canvas">
    // element in the DOM
    ctx: document.getElementById("my-canvas")
        .getContext("2d"),

    rotate(angleRadians) {
        var rotatedX = this.x * Math.cos(angleRadians) -
            this.y * Math.sin(angleRadians);
    }
};

```

```

        var rotatedY = this.x * Math.sin(angleRadians) +
                      this.y * Math.cos(angleRadians);
        this.setXY(rotatedX,rotatedY);
    },
    draw() {
        // plot the point
        Canvas.pixel.call(this,this.x,this.y);
    },
    render() {
        // clear the canvas, and re-render
        // our control-point
        Canvas.renderScene.call(this);
    },
};

// set the logical (0,0) origin at this
// physical location on the canvas
Canvas.setOrigin.call(ControlPoint,100,100);

ControlPoint.setXY(30,40);
// [renders point (30,40) on the canvas]

// ..
// later:

// rotate the point about the (0,0) origin
// 90 degrees counter-clockwise
ControlPoint.rotate(Math.PI / 2);
// [renders point (-40,30) on the canvas]

```

OK, that's a lot of code to digest. Take your time and re-read the snippet several times. I added a couple of new concrete objects (`Canvas` and `ControlPoint`) alongside the previous `Coordinates` object.

Make sure you see and understand the interactions between these three concrete objects.

`ControlPoint` is linked (via `__proto__`) to *implicitly delegate* ([[Prototype]] chain) to `Coordinates`.

Here's an *explicit delegation*: `Canvas.setOrigin.call(ControlPoint,100,100);`; I'm invoking the `Canvas.setOrigin(..)` call in the context of `ControlPoint`. That has the effect of sharing `ctx` with `setOrigin(..)`, via `this`.

`ControlPoint.setXY(..)` delegates *implicitly* to `Coordinates.setXY(..)`, but still in the context of `ControlPoint`. Here's a key detail that's easy to miss: see the `this.render()` inside of `Coordinates.setXY(..)?` Where does that come from? Since the `this` context is `ControlPoint` (not `Coordinates`),

it's invoking `ControlPoint.render()`.

`ControlPoint.render()` *explicitly delegates* to `Canvas.renderScene()`, again still in the `ControlPoint` context. `renderScene()` calls `this.draw()`, but where does that come from? Yep, still from `ControlPoint` (via `this` context).

And `ControlPoint.draw()`? It *explicitly delegates* to `Canvas.pixel(..)`, yet again still in the `ControlPoint` context.

All three objects have methods that end up invoking each other. But these calls aren't particularly hard-wired. `Canvas.renderScene()` doesn't call `ControlPoint.draw()`, it calls `this.draw()`. That's important, because it means that `Canvas.renderScene()` is more flexible to use in a different `this` context – e.g., against another kind of *point* object besides `ControlPoint`.

It's through the `this` context, and the `[[Prototype]]` chain, that these three objects basically are mixed (composed) virtually together, as needed at each step, so that they work together **as if they're one object rather than three separate objects**.

That's the *beauty* of virtual composition as realized by the delegation pattern in JS.

Flexible Context

I mentioned above that we can pretty easily add other concrete objects into the mix. Here's an example:

```
var Coordinates = { /* .. */ };

var Canvas = {
  /* .. */
  line(start,end) {
    this.ctx.beginPath();
    this.ctx.moveTo(start.x,start.y);
    this.ctx.lineTo(end.x,end.y);
    this.ctx.stroke();
  },
};

function lineAnchor(x,y) {
  var anchor = {
    __proto__: Coordinates,
    render() {},
  };
  anchor.setXY(x,y);
  return anchor;
}
```

```

var GuideLine = {
  // NOTE: must have a <canvas id="my-canvas">
  // element in the DOM
  ctx: document.getElementById("my-canvas")
    .getContext("2d"),

  setAnchors(sx,sy,ex,ey) {
    this.start = lineAnchor(sx,sy);
    this.end = lineAnchor(ex,ey);
    this.render();
  },
  draw() {
    // plot the point
    Canvas.line.call(this,this.start,this.end);
  },
  render() {
    // clear the canvas, and re-render
    // our line
    Canvas.renderScene.call(this);
  },
};

// set the logical (0,0) origin at this
// physical location on the canvas
Canvas.setOrigin.call(GuideLine,100,100);

GuideLine.setAnchors(-30,65,45,-17);
// [renders line from (-30,65) to (45,-17)
// on the canvas]

```

That's pretty nice, I think!

But I think another less-obvious benefit is that having objects linked dynamically via `this` context tends to make testing different parts of the program independently, somewhat easier.

For example, `Object.setPrototypeOf(..)` can be used to dynamically change the `[[Prototype]]` linkage of an object, delegating it to a different object such as a mock object. Or you could dynamically redefine `GuideLine.draw()` and `GuideLine.render()` to *explicitly delegate* to a `MockCanvas` instead of `Canvas`.

The `this` keyword, and the `[[Prototype]]` link, are a tremendously flexible mechanism when you understand and leverage them fully.

Why *This*?

OK, so it's hopefully clear that the delegation pattern leans heavily on implicit input, sharing context via `this` rather than through an explicit parameter.

You might rightly ask, why not just always pass around that context explicitly? We can certainly do so, but... to manually pass along the necessary context, we'll have to change pretty much every single function signature, and any corresponding call-sites.

Let's revisit the earlier `ControlPoint` delegation example, and implement it without any delegation-oriented `this` context sharing. Pay careful attention to the differences:

```
var Canvas = {
  setOrigin(ctx,x,y) {
    ctx.translate(x,y);
    ctx.scale(1,-1);
  },
  pixel(ctx,x,y) {
    ctx.fillRect(x,y,1,1);
  },
  renderScene(ctx,entity) {
    // clear the canvas
    var matrix = ctx.getTransform();
    ctx.resetTransform();
    ctx.clearRect(
      0, 0,
      ctx.canvas.width,
      ctx.canvas.height
    );
    ctx.setTransform(matrix);

    entity.draw();
  },
};

var Coordinates = {
  setX(entity,x) {
    entity.x = Math.round(x);
  },
  setY(entity,y) {
    entity.y = Math.round(y);
  },
  setXY(entity,x,y) {
    this.setX(entity,x);
    this.setY(entity,y);
    entity.render();
  },
};

var ControlPoint = {
```

```

// NOTE: must have a <canvas id="my-canvas">
// element in the DOM
ctx: document.getElementById("my-canvas")
  .getContext("2d"),

setXY(x,y) {
  Coordinates.setXY(this,x,y);
},
rotate(angleRadians) {
  var rotatedX = this.x * Math.cos(angleRadians) -
    this.y * Math.sin(angleRadians);
  var rotatedY = this.x * Math.sin(angleRadians) +
    this.y * Math.cos(angleRadians);
  this.setXY(rotatedX,rotatedY);
},
draw() {
  // plot the point
  Canvas.pixel(this.ctx,this.x,this.y);
},
render() {
  // clear the canvas, and re-render
  // our control-point
  Canvas.renderScene(this.ctx,this);
},
};

// set the logical (0,0) origin at this
// physical location on the canvas
Canvas.setOrigin(ControlPoint.ctx,100,100);

// ..

```

To be honest, some of you may prefer that style of code. And that's OK if you're in that camp. This snippet avoids `[[Prototype]]` entirely, and only relies on far fewer basic `this`-style references to properties and methods.

By contrast, the delegation style I'm advocating for in this chapter is unfamiliar and uses `[[Prototype]]` and `this` sharing in ways you're not likely familiar with. To use such a style effectively, you'll have to invest the time and practice to build a deeper familiarity.

But in my opinion, the “cost” of avoiding virtual composition through delegation can be felt across all the function signatures and call-sites; I find them way more cluttered. That explicit context passing is quite a tax.

In fact, I'd never advocate that style of code at all. If you want to avoid delegation, it's probably best to just stick to `class` style code, as seen in Chapter 3. As an exercise left to the reader, try to convert the earlier `ControlPoint` /

GuideLine code snippets to use `class`.

You Don't Know JS: Types & Grammar

Foreword

It was once said, “JavaScript is the only language developers don’t learn to use before using it.”

I laugh each time I hear that quote because it was true for me and I suspect it was for many other developers. JavaScript, and maybe even CSS and HTML, were not a core computer science language taught at college in the Internet’s early days, so personal development was very much based on the budding developer’s search and “view source” abilities to piece together these basic web languages.

I still remember my first high school website project. The task was to create any type of web store, and me being a James Bond fan, I decided to create a Goldeneye store. It had everything: the Goldeneye midi theme song playing in the background, a JavaScript-powered crosshairs following the mouse around the screen, and a gunshot sound that played upon every click. Q would have been proud of this masterpiece of a website.

I tell that story because I did back then what many developers are doing today: I copied and pasted chunks of JavaScript code into my project without having a clue what’s actually happening. The widespread use of JavaScript toolkits like jQuery have, in their own small way, perpetuated this pattern of nonlearning of core JavaScript.

I’m not disparaging JavaScript toolkit use; after all, I’m a member of the MooTools JavaScript team! But the reason JavaScript toolkits are as powerful as they are is because their developers know the fundamentals, and their “gotchas,” and apply them magnificently. As useful as these toolkits are, it’s still incredibly important to know the basics of the language, and with books like Kyle Simpson’s *You Don’t Know JS* series, there’s no excuse not to learn them.

Types and Grammar, the third installment of the series, is an excellent look at the core JavaScript fundamentals that copy and paste and JavaScript toolkits don’t and could never teach you. Coercion and its pitfalls, natives as constructors, and the whole gamut of JavaScript basics are thoroughly explained with focused code examples. Like the other books in this series, Kyle cuts straight to the point: no fluff and word-smithing – exactly the type of tech book I love.

Enjoy Types and Grammar and don’t let it get too far away from your desk!

David Walsh <http://davidwalsh.name>, [@davidwalshblog](https://twitter.com/davidwalshblog) Senior Web Developer, Mozilla

You Don't Know JS

Preface

I'm sure you noticed, but "JS" in the book series title is not an abbreviation for words used to curse about JavaScript, though cursing at the language's quirks is something we can probably all identify with!

From the earliest days of the web, JavaScript has been a foundational technology that drives interactive experience around the content we consume. While flickering mouse trails and annoying pop-up prompts may be where JavaScript started, nearly 2 decades later, the technology and capability of JavaScript has grown many orders of magnitude, and few doubt its importance at the heart of the world's most widely available software platform: the web.

But as a language, it has perpetually been a target for a great deal of criticism, owing partly to its heritage but even more to its design philosophy. Even the name evokes, as Brendan Eich once put it, "dumb kid brother" status next to its more mature older brother "Java". But the name is merely an accident of politics and marketing. The two languages are vastly different in many important ways. "JavaScript" is as related to "Java" as "Carnival" is to "Car".

Because JavaScript borrows concepts and syntax idioms from several languages, including proud C-style procedural roots as well as subtle, less obvious Scheme/Lisp-style functional roots, it is exceedingly approachable to a broad audience of developers, even those with just little to no programming experience. The "Hello World" of JavaScript is so simple that the language is inviting and easy to get comfortable with in early exposure.

While JavaScript is perhaps one of the easiest languages to get up and running with, its eccentricities make solid mastery of the language a vastly less common occurrence than in many other languages. Where it takes a pretty in-depth knowledge of a language like C or C++ to write a full-scale program, full-scale production JavaScript can, and often does, barely scratch the surface of what the language can do.

Sophisticated concepts which are deeply rooted into the language tend instead to surface themselves in *seemingly* simplistic ways, such as passing around functions as callbacks, which encourages the JavaScript developer to just use the language as-is and not worry too much about what's going on under the hood.

It is simultaneously a simple, easy-to-use language that has broad appeal, and a complex and nuanced collection of language mechanics which without careful study will elude *true understanding* even for the most seasoned of JavaScript developers.

Therein lies the paradox of JavaScript, the Achilles' Heel of the language, the challenge we are presently addressing. Because JavaScript *can* be used without understanding, the understanding of the language is often never attained.

Mission

If at every point that you encounter a surprise or frustration in JavaScript, your response is to add it to the blacklist, as some are accustomed to doing, you soon will be relegated to a hollow shell of the richness of JavaScript.

While this subset has been famously dubbed “The Good Parts”, I would implore you, dear reader, to instead consider it the “The Easy Parts”, “The Safe Parts”, or even “The Incomplete Parts”.

This *You Don’t Know JavaScript* book series offers a contrary challenge: learn and deeply understand *all* of JavaScript, even and especially “The Tough Parts”.

Here, we address head on the tendency of JS developers to learn “just enough” to get by, without ever forcing themselves to learn exactly how and why the language behaves the way it does. Furthermore, we eschew the common advice to *retreat* when the road gets rough.

I am not content, nor should you be, at stopping once something *just works*, and not really knowing *why*. I gently challenge you to journey down that bumpy “road less traveled” and embrace all that JavaScript is and can do. With that knowledge, no technique, no framework, no popular buzzword acronym of the week, will be beyond your understanding.

These books each take on specific core parts of the language which are most commonly misunderstood or under-understood, and dive very deep and exhaustively into them. You should come away from reading with a firm confidence in your understanding, not just of the theoretical, but the practical “what you need to know” bits.

The JavaScript you know *right now* is probably *parts* handed down to you by others who’ve been burned by incomplete understanding. *That* JavaScript is but a shadow of the true language. You don’t *really* know JavaScript, *yet*, but if you dig into this series, you *will*. Read on, my friends. JavaScript awaits you.

Summary

JavaScript is awesome. It’s easy to learn partially, and much harder to learn completely (or even *sufficiently*). When developers encounter confusion, they usually blame the language instead of their lack of understanding. These books aim to fix that, inspiring a strong appreciation for the language you can now, and *should*, deeply *know*.

Note: Many of the examples in this book assume modern (and future-reaching) JavaScript engine environments, such as ES6. Some code may not work as described if run in older (pre-ES6) engines.

You Don't Know JS: Types & Grammar

Chapter 1: Types

Most developers would say that a dynamic language (like JS) does not have *types*. Let's see what the ES5.1 specification (<http://www.ecma-international.org/ecma-262/5.1/>) has to say on the topic:

Algorithms within this specification manipulate values each of which has an associated type. The possible value types are exactly those defined in this clause. Types are further sub classified into ECMAScript language types and specification types.

An ECMAScript language type corresponds to values that are directly manipulated by an ECMAScript programmer using the ECMAScript language. The ECMAScript language types are Undefined, Null, Boolean, String, Number, and Object.

Now, if you're a fan of strongly typed (statically typed) languages, you may object to this usage of the word "type." In those languages, "type" means a whole lot *more* than it does here in JS.

Some people say JS shouldn't claim to have "types," and they should instead be called "tags" or perhaps "subtypes".

Bah! We're going to use this rough definition (the same one that seems to drive the wording of the spec): a *type* is an intrinsic, built-in set of characteristics that uniquely identifies the behavior of a particular value and distinguishes it from other values, both to the engine **and to the developer**.

In other words, if both the engine and the developer treat value `42` (the number) differently than they treat value `"42"` (the string), then those two values have different *types* – `number` and `string`, respectively. When you use `42`, you are *intending* to do something numeric, like math. But when you use `"42"`, you are *intending* to do something string'ish, like outputting to the page, etc. **These two values have different types.**

That's by no means a perfect definition. But it's good enough for this discussion. And it's consistent with how JS describes itself.

A Type By Any Other Name...

Beyond academic definition disagreements, why does it matter if JavaScript has *types* or not?

Having a proper understanding of each *type* and its intrinsic behavior is absolutely essential to understanding how to properly and accurately convert values to different types (see Coercion, Chapter 4). Nearly every JS program ever writ-

ten will need to handle value coercion in some shape or form, so it's important you do so responsibly and with confidence.

If you have the `number` value `42`, but you want to treat it like a `string`, such as pulling out the `"2"` as a character in position 1, you obviously must first convert (coerce) the value from `number` to `string`.

That seems simple enough.

But there are many different ways that such coercion can happen. Some of these ways are explicit, easy to reason about, and reliable. But if you're not careful, coercion can happen in very strange and surprising ways.

Coercion confusion is perhaps one of the most profound frustrations for JavaScript developers. It has often been criticized as being so *dangerous* as to be considered a flaw in the design of the language, to be shunned and avoided.

Armed with a full understanding of JavaScript types, we're aiming to illustrate why coercion's *bad reputation* is largely overhyped and somewhat undeserved – to flip your perspective, to seeing coercion's power and usefulness. But first, we have to get a much better grip on values and types.

Built-in Types

JavaScript defines seven built-in types:

- `null`
- `undefined`
- `boolean`
- `number`
- `string`
- `object`
- `symbol` – added in ES6!

Note: All of these types except `object` are called “primitives”.

The `typeof` operator inspects the type of the given value, and always returns one of seven string values – surprisingly, there's not an exact 1-to-1 match with the seven built-in types we just listed.

```
typeof undefined      === "undefined"; // true
typeof true           === "boolean";    // true
typeof 42             === "number";     // true
typeof "42"           === "string";    // true
typeof { life: 42 }   === "object";    // true

// added in ES6!
typeof Symbol()       === "symbol";    // true
```

These six listed types have values of the corresponding type and return a string value of the same name, as shown. `Symbol` is a new data type as of ES6, and

will be covered in Chapter 3.

As you may have noticed, I excluded `null` from the above listing. It's *special* – special in the sense that it's buggy when combined with the `typeof` operator:

```
typeof null === "object"; // true
```

It would have been nice (and correct!) if it returned `"null"`, but this original bug in JS has persisted for nearly two decades, and will likely never be fixed because there's too much existing web content that relies on its buggy behavior that "fixing" the bug would *create* more "bugs" and break a lot of web software.

If you want to test for a `null` value using its type, you need a compound condition:

```
var a = null;
```

```
(!a && typeof a === "object"); // true
```

`null` is the only primitive value that is "falsy" (aka false-like; see Chapter 4) but that also returns `"object"` from the `typeof` check.

So what's the seventh string value that `typeof` can return?

```
typeof function a(){ /* .. */ } === "function"; // true
```

It's easy to think that `function` would be a top-level built-in type in JS, especially given this behavior of the `typeof` operator. However, if you read the spec, you'll see it's actually a "subtype" of `object`. Specifically, a function is referred to as a "callable object" – an object that has an internal `[[Call]]` property that allows it to be invoked.

The fact that functions are actually objects is quite useful. Most importantly, they can have properties. For example:

```
function a(b,c) {  
  /* .. */  
}
```

The function object has a `length` property set to the number of formal parameters it is declared with.

```
a.length; // 2
```

Since you declared the function with two formal named parameters (`b` and `c`), the "length of the function" is 2.

What about arrays? They're native to JS, so are they a special type?

```
typeof [1,2,3] === "object"; // true
```

Nope, just objects. It's most appropriate to think of them also as a "subtype" of `object` (see Chapter 3), in this case with the additional characteristics of being

numerically indexed (as opposed to just being string-keyed like plain objects) and maintaining an automatically updated `.length` property.

Values as Types

In JavaScript, variables don't have types – **values have types**. Variables can hold any value, at any time.

Another way to think about JS types is that JS doesn't have “type enforcement,” in that the engine doesn't insist that a *variable* always holds values of the *same initial type* that it starts out with. A variable can, in one assignment statement, hold a `string`, and in the next hold a `number`, and so on.

The `value` 42 has an intrinsic type of `number`, and its *type* cannot be changed. Another value, like "42" with the `string` type, can be created *from* the `number` value 42 through a process called **coercion** (see Chapter 4).

If you use `typeof` against a variable, it's not asking “what's the type of the variable?” as it may seem, since JS variables have no types. Instead, it's asking “what's the type of the value *in* the variable?”

```
var a = 42;
typeof a; // "number"

a = true;
typeof a; // "boolean"
```

The `typeof` operator always returns a string. So:

```
typeof typeof 42; // "string"
```

The first `typeof` 42 returns "number", and `typeof` "number" is "string".

`undefined` vs “`undeclared`”

Variables that have no value *currently*, actually have the `undefined` value. Calling `typeof` against such variables will return "undefined":

```
var a;

typeof a; // "undefined"

var b = 42;
var c;

// later
b = c;

typeof b; // "undefined"
typeof c; // "undefined"
```

It's tempting for most developers to think of the word "undefined" and think of it as a synonym for "undeclared." However, in JS, these two concepts are quite different.

An "undefined" variable is one that has been declared in the accessible scope, but *at the moment* has no other value in it. By contrast, an "undeclared" variable is one that has not been formally declared in the accessible scope.

Consider:

```
var a;  
  
a; // undefined  
b; // ReferenceError: b is not defined
```

An annoying confusion is the error message that browsers assign to this condition. As you can see, the message is "b is not defined," which is of course very easy and reasonable to confuse with "b is undefined." Yet again, "undefined" and "is not defined" are very different things. It'd be nice if the browsers said something like "b is not found" or "b is not declared," to reduce the confusion!

There's also a special behavior associated with `typeof` as it relates to undeclared variables that even further reinforces the confusion. Consider:

```
var a;  
  
typeof a; // "undefined"  
  
typeof b; // "undefined"
```

The `typeof` operator returns "undefined" even for "undeclared" (or "not defined") variables. Notice that there was no error thrown when we executed `typeof b`, even though `b` is an undeclared variable. This is a special safety guard in the behavior of `typeof`.

Similar to above, it would have been nice if `typeof` used with an undeclared variable returned "undeclared" instead of conflating the result value with the different "undefined" case.

`typeof Undeclared`

Nevertheless, this safety guard is a useful feature when dealing with JavaScript in the browser, where multiple script files can load variables into the shared global namespace.

Note: Many developers believe there should never be any variables in the global namespace, and that everything should be contained in modules and private/separate namespaces. This is great in theory but nearly impossible in practicality; still it's a good goal to strive toward! Fortunately, ES6 added first-class support for modules, which will eventually make that much more practical.

As a simple example, imagine having a “debug mode” in your program that is controlled by a global variable (flag) called `DEBUG`. You’d want to check if that variable was declared before performing a debug task like logging a message to the console. A top-level global `var DEBUG = true` declaration would only be included in a “debug.js” file, which you only load into the browser when you’re in development/testing, but not in production.

However, you have to take care in how you check for the global `DEBUG` variable in the rest of your application code, so that you don’t throw a `ReferenceError`. The safety guard on `typeof` is our friend in this case.

```
// oops, this would throw an error!
if (DEBUG) {
  console.log( "Debugging is starting" );
}

// this is a safe existence check
if (typeof DEBUG !== "undefined") {
  console.log( "Debugging is starting" );
}
```

This sort of check is useful even if you’re not dealing with user-defined variables (like `DEBUG`). If you are doing a feature check for a built-in API, you may also find it helpful to check without throwing an error:

```
if (typeof atob === "undefined") {
  atob = function() { /*...*/ };
}
```

Note: If you’re defining a “polyfill” for a feature if it doesn’t already exist, you probably want to avoid using `var` to make the `atob` declaration. If you declare `var atob` inside the `if` statement, this declaration is hoisted (see the *Scope & Closures* title of this series) to the top of the scope, even if the `if` condition doesn’t pass (because the global `atob` already exists!). In some browsers and for some special types of global built-in variables (often called “host objects”), this duplicate declaration may throw an error. Omitting the `var` prevents this hoisted declaration.

Another way of doing these checks against global variables but without the safety guard feature of `typeof` is to observe that all global variables are also properties of the global object, which in the browser is basically the `window` object. So, the above checks could have been done (quite safely) as:

```
if (window.DEBUG) {
  // ...
}

if (!window.atob) {
  // ...
```

```
}
```

Unlike referencing undeclared variables, there is no `ReferenceError` thrown if you try to access an object property (even on the global `window` object) that doesn't exist.

On the other hand, manually referencing the global variable with a `window` reference is something some developers prefer to avoid, especially if your code needs to run in multiple JS environments (not just browsers, but server-side `node.js`, for instance), where the global object may not always be called `window`.

Technically, this safety guard on `typeof` is useful even if you're not using global variables, though these circumstances are less common, and some developers may find this design approach less desirable. Imagine a utility function that you want others to copy-and-paste into their programs or modules, in which you want to check to see if the including program has defined a certain variable (so that you can use it) or not:

```
function doSomethingCool() {
  var helper =
    (typeof FeatureXYZ !== "undefined") ?
    FeatureXYZ :
    function() { /*.. default feature ..*/ };

  var val = helper();
  // ..
}
```

`doSomethingCool()` tests for a variable called `FeatureXYZ`, and if found, uses it, but if not, uses its own. Now, if someone includes this utility in their module/program, it safely checks if they've defined `FeatureXYZ` or not:

```
// an IIFE (see "Immediately Invoked Function Expressions"
// discussion in the *Scope & Closures* title of this series)
(function(){
  function FeatureXYZ() { /*.. my XYZ feature ..*/ }

  // include `doSomethingCool(..)`
  function doSomethingCool() {
    var helper =
      (typeof FeatureXYZ !== "undefined") ?
      FeatureXYZ :
      function() { /*.. default feature ..*/ };

    var val = helper();
    // ..
  }
}
```

```
    doSomethingCool();
})();
```

Here, `FeatureXYZ` is not at all a global variable, but we're still using the safety guard of `typeof` to make it safe to check for. And importantly, here there is *no* object we can use (like we did for global variables with `window.____`) to make the check, so `typeof` is quite helpful.

Other developers would prefer a design pattern called “dependency injection,” where instead of `doSomethingCool()` inspecting implicitly for `FeatureXYZ` to be defined outside/around it, it would need to have the dependency explicitly passed in, like:

```
function doSomethingCool(FeatureXYZ) {
  var helper = FeatureXYZ ||
    function() { /*.. default feature ...*/ };

  var val = helper();
  // ...
}
```

There are lots of options when designing such functionality. No one pattern here is “correct” or “wrong” – there are various tradeoffs to each approach. But overall, it's nice that the `typeof` undeclared safety guard gives us more options.

Review

JavaScript has seven built-in *types*: `null`, `undefined`, `boolean`, `number`, `string`, `object`, `symbol`. They can be identified by the `typeof` operator.

Variables don't have types, but the values in them do. These types define intrinsic behavior of the values.

Many developers will assume “`undefined`” and “`undeclared`” are roughly the same thing, but in JavaScript, they're quite different. `undefined` is a value that a declared variable can hold. “`Undeclared`” means a variable has never been declared.

JavaScript unfortunately kind of conflates these two terms, not only in its error messages (“`ReferenceError: a is not defined`”) but also in the return values of `typeof`, which is “`undefined`” for both cases.

However, the safety guard (preventing an error) on `typeof` when used against an undeclared variable can be helpful in certain cases.

You Don't Know JS: Types & Grammar

Chapter 2: Values

arrays, strings, and numbers are the most basic building-blocks of any program, but JavaScript has some unique characteristics with these types that may either delight or confound you.

Let's look at several of the built-in value types in JS, and explore how we can more fully understand and correctly leverage their behaviors.

Arrays

As compared to other type-enforced languages, JavaScript arrays are just containers for any type of value, from `string` to `number` to `object` to even another `array` (which is how you get multidimensional `arrays`).

```
var a = [ 1, "2", [3] ];

a.length;      // 3
a[0] === 1;    // true
a[2][0] === 3; // true
```

You don't need to presize your `arrays` (see "Arrays" in Chapter 3), you can just declare them and add values as you see fit:

```
var a = [ ];

a.length; // 0

a[0] = 1;
a[1] = "2";
a[2] = [ 3 ];

a.length; // 3
```

Warning: Using `delete` on an `array` value will remove that slot from the `array`, but even if you remove the final element, it does **not** update the `length` property, so be careful! We'll cover the `delete` operator itself in more detail in Chapter 5.

Be careful about creating "sparse" `arrays` (leaving or creating empty/missing slots):

```
var a = [ ];

a[0] = 1;
// no `a[1]` slot set here
a[2] = [ 3 ];
```

```
a[1];      // undefined  
a.length;  // 3
```

While that works, it can lead to some confusing behavior with the “empty slots” you leave in between. While the slot appears to have the `undefined` value in it, it will not behave the same as if the slot is explicitly set (`a[1] = undefined`). See “Arrays” in Chapter 3 for more information.

`arrays` are numerically indexed (as you’d expect), but the tricky thing is that they also are objects that can have `string` keys/properties added to them (but which don’t count toward the `length` of the `array`):

```
var a = [ ];  
  
a[0] = 1;  
a["foobar"] = 2;  
  
a.length;    // 1  
a["foobar"]; // 2  
a.foobar;    // 2
```

However, a gotcha to be aware of is that if a `string` value intended as a key can be coerced to a standard base-10 `number`, then it is assumed that you wanted to use it as a `number` index rather than as a `string` key!

```
var a = [ ];  
  
a["13"] = 42;  
  
a.length; // 14
```

Generally, it’s not a great idea to add `string` keys/properties to `arrays`. Use `objects` for holding values in keys/properties, and save `arrays` for strictly numerically indexed values.

Array-Likes

There will be occasions where you need to convert an `array-like` value (a numerically indexed collection of values) into a true `array`, usually so you can call array utilities (like `indexOf(..)`, `concat(..)`, `forEach(..)`, etc.) against the collection of values.

For example, various DOM query operations return lists of DOM elements that are not true `arrays` but are `array-like` enough for our conversion purposes. Another common example is when functions expose the `arguments` (`array-like`) object (as of ES6, deprecated) to access the arguments as a list.

One very common way to make such a conversion is to borrow the `slice(..)` utility against the value:

```
function foo() {  
    var arr = Array.prototype.slice.call( arguments );  
    arr.push( "bam" );  
    console.log( arr );  
}  
  
foo( "bar", "baz" ); // [ "bar", "baz", "bam" ]
```

If `slice()` is called without any other parameters, as it effectively is in the above snippet, the default values for its parameters have the effect of duplicating the `array` (or, in this case, `array-like`).

As of ES6, there's also a built-in utility called `Array.from(..)` that can do the same task:

```
...  
var arr = Array.from( arguments );  
...
```

Note: `Array.from(..)` has several powerful capabilities, and will be covered in detail in the *ES6 & Beyond* title of this series.

Strings

It's a very common belief that `strings` are essentially just `arrays` of characters. While the implementation under the covers may or may not use `arrays`, it's important to realize that JavaScript `strings` are really not the same as `arrays` of characters. The similarity is mostly just skin-deep.

For example, let's consider these two values:

```
var a = "foo";  
var b = [ "f", "o", "o" ];
```

Strings do have a shallow resemblance to `arrays` – `array-likes`, as above – for instance, both of them having a `length` property, an `indexOf(..)` method (array version only as of ES5), and a `concat(..)` method:

```
a.length;                      // 3  
b.length;                      // 3  
  
a.indexOf( "o" );               // 1  
b.indexOf( "o" );               // 1  
  
var c = a.concat( "bar" );      // "foobar"  
var d = b.concat( [ "b", "a", "r" ] ); // [ "f", "o", "o", "b", "a", "r" ]
```

```

a === c;           // false
b === d;           // false

a;                 // "foo"
b;                 // ["f", "o", "o"]

```

So, they're both basically just “arrays of characters”, right? **Not exactly**:

```

a[1] = "O";
b[1] = "O";

a; // "foo"
b; // ["f", "O", "o"]

```

JavaScript **strings** are immutable, while **arrays** are quite mutable. Moreover, the `a[1]` character position access form was not always widely valid JavaScript. Older versions of IE did not allow that syntax (but now they do). Instead, the *correct* approach has been `a.charAt(1)`.

A further consequence of immutable **strings** is that none of the **string** methods that alter its contents can modify in-place, but rather must create and return new **strings**. By contrast, many of the methods that change **array** contents actually *do* modify in-place.

```

c = a.toUpperCase();
a === c; // false
a;       // "foo"
c;       // "FOO"

b.push( "!" );
b;         // ["f", "O", "o", "!"]

```

Also, many of the **array** methods that could be helpful when dealing with **strings** are not actually available for them, but we can “borrow” non-mutation **array** methods against our **string**:

```

a.join();        // undefined
a.map();         // undefined

var c = Array.prototype.join.call( a, "-" );
var d = Array.prototype.map.call( a, function(v){
    return v.toUpperCase() + ".";
} ).join( "" );

c;             // "f-o-o"
d;             // "F.O.O."

```

Let's take another example: reversing a **string** (incidentally, a common JavaScript interview trivia question!). **arrays** have a `reverse()` in-place mutator method, but **strings** do not:

```

a.reverse;      // undefined

b.reverse();    // ["!", "o", "D", "f"]
b;              // ["!", "o", "D", "f"]

```

Unfortunately, this “borrowing” doesn’t work with `array` mutators, because `strings` are immutable and thus can’t be modified in place:

```

Array.prototype.reverse.call( a );
// still returns a String object wrapper (see Chapter 3)
// for "foo" :(

```

Another workaround (aka hack) is to convert the `string` into an `array`, perform the desired operation, then convert it back to a `string`.

```

var c = a
    // split `a` into an array of characters
    .split( "" )
    // reverse the array of characters
    .reverse()
    // join the array of characters back to a string
    .join( "" );

c; // "oof"

```

If that feels ugly, it is. Nevertheless, *it works* for simple `strings`, so if you need something quick-n-dirty, often such an approach gets the job done.

Warning: Be careful! This approach **doesn’t work** for `strings` with complex (unicode) characters in them (astral symbols, multibyte characters, etc.). You need more sophisticated library utilities that are unicode-aware for such operations to be handled accurately. Consult Mathias Bynens’ work on the subject: *Esrever* (<https://github.com/mathiasbynens/esrever>).

The other way to look at this is: if you are more commonly doing tasks on your “strings” that treat them as basically *arrays of characters*, perhaps it’s better to just actually store them as `arrays` rather than as `strings`. You’ll probably save yourself a lot of hassle of converting from `string` to `array` each time. You can always call `join("")` on the `array of characters` whenever you actually need the `string` representation.

Numbers

JavaScript has just one numeric type: `number`. This type includes both “integer” values and fractional decimal numbers. I say “integer” in quotes because it’s long been a criticism of JS that there are not true integers, as there are in other languages. That may change at some point in the future, but for now, we just have `numbers` for everything.

So, in JS, an “integer” is just a value that has no fractional decimal value. That is, 42.0 is as much an “integer” as 42.

Like most modern languages, including practically all scripting languages, the implementation of JavaScript’s `numbers` is based on the “IEEE 754” standard, often called “floating-point.” JavaScript specifically uses the “double precision” format (aka “64-bit binary”) of the standard.

There are many great write-ups on the Web about the nitty-gritty details of how binary floating-point numbers are stored in memory, and the implications of those choices. Because understanding bit patterns in memory is not strictly necessary to understand how to correctly use `numbers` in JS, we’ll leave it as an exercise for the interested reader if you’d like to dig further into IEEE 754 details.

Numeric Syntax

Number literals are expressed in JavaScript generally as base-10 decimal literals. For example:

```
var a = 42;  
var b = 42.3;
```

The leading portion of a decimal value, if 0, is optional:

```
var a = 0.42;  
var b = .42;
```

Similarly, the trailing portion (the fractional) of a decimal value after the ., if 0, is optional:

```
var a = 42.0;  
var b = 42.;
```

Warning: 42. is pretty uncommon, and probably not a great idea if you’re trying to avoid confusion when other people read your code. But it is, nevertheless, valid.

By default, most `numbers` will be outputted as base-10 decimals, with trailing fractional 0s removed. So:

```
var a = 42.300;  
var b = 42.0;  
  
a; // 42.3  
b; // 42
```

Very large or very small `numbers` will by default be outputted in exponent form, the same as the output of the `toExponential()` method, like:

```
var a = 5E10;  
a; // 50000000000
```

```

a.toExponential(); // "5e+10"

var b = a * a;
b; // 2.5e+21

var c = 1 / a;
c; // 2e-11

```

Because `number` values can be boxed with the `Number` object wrapper (see Chapter 3), `number` values can access methods that are built into the `Number.prototype` (see Chapter 3). For example, the `toFixed(..)` method allows you to specify how many fractional decimal places you'd like the value to be represented with:

```

var a = 42.59;

a.toFixed( 0 ); // "43"
a.toFixed( 1 ); // "42.6"
a.toFixed( 2 ); // "42.59"
a.toFixed( 3 ); // "42.590"
a.toFixed( 4 ); // "42.5900"

```

Notice that the output is actually a `string` representation of the `number`, and that the value is 0-padded on the right-hand side if you ask for more decimals than the value holds.

`toPrecision(..)` is similar, but specifies how many *significant digits* should be used to represent the value:

```

var a = 42.59;

a.toPrecision( 1 ); // "4e+1"
a.toPrecision( 2 ); // "43"
a.toPrecision( 3 ); // "42.6"
a.toPrecision( 4 ); // "42.59"
a.toPrecision( 5 ); // "42.590"
a.toPrecision( 6 ); // "42.5900"

```

You don't have to use a variable with the value in it to access these methods; you can access these methods directly on `number` literals. But you have to be careful with the `.` operator. Since `.` is a valid numeric character, it will first be interpreted as part of the `number` literal, if possible, instead of being interpreted as a property accessor.

```

// invalid syntax:
42.toFixed( 3 ); // SyntaxError

// these are all valid:
(42).toFixed( 3 ); // "42.000"

```

```
0.42.toFixed( 3 ); // "0.420"  
42..toFixed( 3 ); // "42.000"
```

`42.toFixed(3)` is invalid syntax, because the `.` is swallowed up as part of the `42.` literal (which is valid – see above!), and so then there's no `.` property operator present to make the `.toFixed` access.

`42..toFixed(3)` works because the first `.` is part of the `number` and the second `.` is the property operator. But it probably looks strange, and indeed it's very rare to see something like that in actual JavaScript code. In fact, it's pretty uncommon to access methods directly on any of the primitive values. Uncommon doesn't mean *bad* or *wrong*.

Note: There are libraries that extend the built-in `Number.prototype` (see Chapter 3) to provide extra operations on/with `numbers`, and so in those cases, it's perfectly valid to use something like `10..makeItRain()` to set off a 10-second money raining animation, or something else silly like that.

This is also technically valid (notice the space):

```
42 .toFixed(3); // "42.000"
```

However, with the `number` literal specifically, **this is particularly confusing coding style** and will serve no other purpose but to confuse other developers (and your future self). Avoid it.

`numbers` can also be specified in exponent form, which is common when representing larger `numbers`, such as:

```
var onethousand = 1E3; // means 1 * 10^3  
var onemilliononehundredthousand = 1.1E6; // means 1.1 * 10^6
```

`number` literals can also be expressed in other bases, like binary, octal, and hexadecimal.

These formats work in current versions of JavaScript:

```
0xf3; // hexadecimal for: 243  
0Xf3; // ditto
```

```
0363; // octal for: 243
```

Note: Starting with ES6 + `strict` mode, the `0363` form of octal literals is no longer allowed (see below for the new form). The `0363` form is still allowed in `non-strict` mode, but you should stop using it anyway, to be future-friendly (and because you should be using `strict` mode by now!).

As of ES6, the following new forms are also valid:

```
0o363; // octal for: 243  
00363; // ditto
```

```
0b11110011; // binary for: 243
0B11110011; // ditto
```

Please do your fellow developers a favor: never use the `00363` form. `0` next to capital `O` is just asking for confusion. Always use the lowercase predicates `0x`, `0b`, and `0o`.

Small Decimal Values

The most (in)famous side effect of using binary floating-point numbers (which, remember, is true of **all** languages that use IEEE 754 – not *just* JavaScript as many assume/pretend) is:

```
0.1 + 0.2 === 0.3; // false
```

Mathematically, we know that statement should be `true`. Why is it `false`?

Simply put, the representations for `0.1` and `0.2` in binary floating-point are not exact, so when they are added, the result is not exactly `0.3`. It's **really** close: `0.3000000000000004`, but if your comparison fails, “close” is irrelevant.

Note: Should JavaScript switch to a different `number` implementation that has exact representations for all values? Some think so. There have been many alternatives presented over the years. None of them have been accepted yet, and perhaps never will. As easy as it may seem to just wave a hand and say, “fix that bug already!”, it's not nearly that easy. If it were, it most definitely would have been changed a long time ago.

Now, the question is, if some `numbers` can't be *trusted* to be exact, does that mean we can't use `numbers` at all? **Of course not.**

There are some applications where you need to be more careful, especially when dealing with fractional decimal values. There are also plenty of (maybe most?) applications that only deal with whole numbers (“integers”), and moreover, only deal with numbers in the millions or trillions at maximum. These applications have been, and always will be, **perfectly safe** to use numeric operations in JS.

What if we *did* need to compare two `numbers`, like `0.1 + 0.2` to `0.3`, knowing that the simple equality test fails?

The most commonly accepted practice is to use a tiny “rounding error” value as the *tolerance* for comparison. This tiny value is often called “machine epsilon,” which is commonly 2^{-52} (`2.220446049250313e-16`) for the kind of `numbers` in JavaScript.

As of ES6, `Number.EPSILON` is predefined with this tolerance value, so you'd want to use it, but you can safely polyfill the definition for pre-ES6:

```
if (!Number.EPSILON) {
  Number.EPSILON = Math.pow(2, -52);
}
```

We can use this `Number.EPSILON` to compare two `numbers` for “equality” (within the rounding error tolerance):

```
function numbersCloseEnoughToEqual(n1, n2) {
  return Math.abs( n1 - n2 ) < Number.EPSILON;
}

var a = 0.1 + 0.2;
var b = 0.3;

numbersCloseEnoughToEqual( a, b );           // true
numbersCloseEnoughToEqual( 0.0000001, 0.0000002 ); // false
```

The maximum floating-point value that can be represented is roughly `1.798e+308` (which is really, really, really huge!), predefined for you as `Number.MAX_VALUE`. On the small end, `Number.MIN_VALUE` is roughly `5e-324`, which isn’t negative but is really close to zero!

Safe Integer Ranges

Because of how `numbers` are represented, there is a range of “safe” values for the whole `number` “integers”, and it’s significantly less than `Number.MAX_VALUE`.

The maximum integer that can “safely” be represented (that is, there’s a guarantee that the requested value is actually representable unambiguously) is $2^{53} - 1$, which is `9007199254740991`. If you insert your commas, you’ll see that this is just over 9 quadrillion. So that’s pretty darn big for `numbers` to range up to.

This value is actually automatically predefined in ES6, as `Number.MAX_SAFE_INTEGER`. Unsurprisingly, there’s a minimum value, `-9007199254740991`, and it’s defined in ES6 as `Number.MIN_SAFE_INTEGER`.

The main way that JS programs are confronted with dealing with such large numbers is when dealing with 64-bit IDs from databases, etc. 64-bit numbers cannot be represented accurately with the `number` type, so must be stored in (and transmitted to/from) JavaScript using `string` representation.

Numeric operations on such large ID `number` values (besides comparison, which will be fine with `strings`) aren’t all that common, thankfully. But if you *do* need to perform math on these very large values, for now you’ll need to use a *big number* utility. Big numbers may get official support in a future version of JavaScript.

Testing for Integers

To test if a value is an integer, you can use the ES6-specified `Number.isInteger(..)`:

```
Number.isInteger( 42 );      // true
Number.isInteger( 42.000 ); // true
```

```
Number.isInteger( 42.3 ); // false
```

To polyfill `Number.isInteger(..)` for pre-ES6:

```
if (!Number.isInteger) {
  Number.isInteger = function(num) {
    return typeof num == "number" && num % 1 == 0;
  };
}
```

To test if a value is a *safe integer*, use the ES6-specified `Number.isSafeInteger(..)`:

```
Number.isSafeInteger( Number.MAX_SAFE_INTEGER ); // true
Number.isSafeInteger( Math.pow( 2, 53 ) ); // false
Number.isSafeInteger( Math.pow( 2, 53 ) - 1 ); // true
```

To polyfill `Number.isSafeInteger(..)` in pre-ES6 browsers:

```
if (!Number.isSafeInteger) {
  Number.isSafeInteger = function(num) {
    return Number.isInteger( num ) &&
      Math.abs( num ) <= Number.MAX_SAFE_INTEGER;
  };
}
```

32-bit (Signed) Integers

While integers can range up to roughly 9 quadrillion safely (53 bits), there are some numeric operations (like the bitwise operators) that are only defined for 32-bit numbers, so the “safe range” for numbers used in that way must be much smaller.

The range then is `Math.pow(-2,31)` (-2147483648, about -2.1 billion) up to `Math.pow(2,31)-1` (2147483647, about +2.1 billion).

To force a `number` value in `a` to a 32-bit signed integer value, use `a | 0`. This works because the `|` bitwise operator only works for 32-bit integer values (meaning it can only pay attention to 32 bits and any other bits will be lost). Then, “or’ing” with zero is essentially a no-op bitwise speaking.

Note: Certain special values (which we will cover in the next section) such as `NaN` and `Infinity` are not “32-bit safe,” in that those values when passed to a bitwise operator will pass through the abstract operation `ToInt32` (see Chapter 4) and become simply the `+0` value for the purpose of that bitwise operation.

Special Values

There are several special values spread across the various types that the *alert* JS developer needs to be aware of, and use properly.

The Non-value Values

For the `undefined` type, there is one and only one value: `undefined`. For the `null` type, there is one and only one value: `null`. So for both of them, the label is both its type and its value.

Both `undefined` and `null` are often taken to be interchangeable as either “empty” values or “non” values. Other developers prefer to distinguish between them with nuance. For example:

- `null` is an empty value
- `undefined` is a missing value

Or:

- `undefined` hasn't had a value yet
- `null` had a value and doesn't anymore

Regardless of how you choose to “define” and use these two values, `null` is a special keyword, not an identifier, and thus you cannot treat it as a variable to assign to (why would you!?). However, `undefined` is (unfortunately) an identifier. Uh oh.

Undefined

In non-strict mode, it's actually possible (though incredibly ill-advised!) to assign a value to the globally provided `undefined` identifier:

```
function foo() {
  undefined = 2; // really bad idea!
}

foo();

function foo() {
  "use strict";
  undefined = 2; // TypeError!
}

foo();
```

In both non-strict mode and strict mode, however, you can create a local variable of the name `undefined`. But again, this is a terrible idea!

```
function foo() {
  "use strict";
  var undefined = 2;
  console.log( undefined ); // 2
}

foo();
```

Friends don't let friends override `undefined`. Ever.

void Operator While `undefined` is a built-in identifier that holds (unless modified – see above!) the built-in `undefined` value, another way to get this value is the `void` operator.

The expression `void ___` “voids” out any value, so that the result of the expression is always the `undefined` value. It doesn’t modify the existing value; it just ensures that no value comes back from the operator expression.

```
var a = 42;

console.log( void a, a ); // undefined 42
```

By convention (mostly from C-language programming), to represent the `undefined` value stand-alone by using `void`, you’d use `void 0` (though clearly even `void true` or any other `void` expression does the same thing). There’s no practical difference between `void 0`, `void 1`, and `undefined`.

But the `void` operator can be useful in a few other circumstances, if you need to ensure that an expression has no result value (even if it has side effects).

For example:

```
function doSomething() {
  // note: `APP.ready` is provided by our application
  if (!APP.ready) {
    // try again later
    return void setTimeout( doSomething, 100 );
  }

  var result;

  // do some other stuff
  return result;
}

// were we able to do it right away?
if (doSomething()) {
  // handle next tasks right away
}
```

Here, the `setTimeout(..)` function returns a numeric value (the unique identifier of the timer interval, if you wanted to cancel it), but we want to `void` that out so that the return value of our function doesn’t give a false-positive with the `if` statement.

Many devs prefer to just do these actions separately, which works the same but doesn’t use the `void` operator:

```

if (!APP.ready) {
  // try again later
  setTimeout( doSomething, 100 );
  return;
}

```

In general, if there's ever a place where a value exists (from some expression) and you'd find it useful for the value to be `undefined` instead, use the `void` operator. That probably won't be terribly common in your programs, but in the rare cases you do need it, it can be quite helpful.

Special Numbers

The `number` type includes several special values. We'll take a look at each in detail.

The Not Number, `Number` Any mathematic operation you perform without both operands being `numbers` (or values that can be interpreted as regular `numbers` in base 10 or base 16) will result in the operation failing to produce a valid `number`, in which case you will get the `NaN` value.

`NaN` literally stands for “not a `number`”, though this label/description is very poor and misleading, as we'll see shortly. It would be much more accurate to think of `NaN` as being “invalid number,” “failed number,” or even “bad number,” than to think of it as “not a number.”

For example:

```

var a = 2 / "foo";      // NaN

typeof a === "number"; // true

```

In other words: “the type of not-a-number is ‘number’!” Hooray for confusing names and semantics.

`NaN` is a kind of “sentinel value” (an otherwise normal value that's assigned a special meaning) that represents a special kind of error condition within the `number` set. The error condition is, in essence: “I tried to perform a mathematic operation but failed, so here's the failed `number` result instead.”

So, if you have a value in some variable and want to test to see if it's this special failed-number `NaN`, you might think you could directly compare to `NaN` itself, as you can with any other value, like `null` or `undefined`. Nope.

```

var a = 2 / "foo";

a === NaN; // false
a === NaN; // false

```

`NaN` is a very special value in that it's never equal to another `NaN` value (i.e., it's never equal to itself). It's the only value, in fact, that is not reflexive (without the Identity characteristic `x === x`). So, `NaN !== NaN`. A bit strange, huh?

So how *do* we test for it, if we can't compare to `NaN` (since that comparison would always fail)?

```
var a = 2 / "foo";  
  
isNaN( a ); // true
```

Easy enough, right? We use the built-in global utility called `isNaN(..)` and it tells us if the value is `NaN` or not. Problem solved!

Not so fast.

The `isNaN(..)` utility has a fatal flaw. It appears it tried to take the meaning of `NaN` ("Not a Number") too literally – that its job is basically: "test if the thing passed in is either not a `number` or is a `number`." But that's not quite accurate.

```
var a = 2 / "foo";  
var b = "foo";  
  
a; // NaN  
b; // "foo"  
  
window.isNaN( a ); // true  
window.isNaN( b ); // true -- ouch!
```

Clearly, "`foo`" is literally *not a number*, but it's definitely not the `NaN` value either! This bug has been in JS since the very beginning (over 19 years of *ouch*).

As of ES6, finally a replacement utility has been provided: `Number.isNaN(..)`. A simple polyfill for it so that you can safely check `NaN` values *now* even in pre-ES6 browsers is:

```
if (!Number.isNaN) {  
    Number.isNaN = function(n) {  
        return (  
            typeof n === "number" &&  
            window.isNaN( n )  
        );  
    };  
}  
  
var a = 2 / "foo";  
var b = "foo";
```

```
Number.isNaN( a ); // true
Number.isNaN( b ); // false -- phew!
```

Actually, we can implement a `Number.isNaN(..)` polyfill even easier, by taking advantage of that peculiar fact that `NaN` isn't equal to itself. `NaN` is the *only* value in the whole language where that's true; every other value is always **equal to itself**.

So:

```
if (!Number.isNaN) {
  Number.isNaN = function(n) {
    return n !== n;
  };
}
```

Weird, huh? But it works!

`NaN`s are probably a reality in a lot of real-world JS programs, either on purpose or by accident. It's a really good idea to use a reliable test, like `Number.isNaN(..)` as provided (or polyfilled), to recognize them properly.

If you're currently using just `isNaN(..)` in a program, the sad reality is your program *has a bug*, even if you haven't been bitten by it yet!

Infinities Developers from traditional compiled languages like C are probably used to seeing either a compiler error or runtime exception, like "Divide by zero," for an operation like:

```
var a = 1 / 0;
```

However, in JS, this operation is well-defined and results in the value `Infinity` (aka `Number.POSITIVE_INFINITY`). Unsurprisingly:

```
var a = 1 / 0; // Infinity
var b = -1 / 0; // -Infinity
```

As you can see, `-Infinity` (aka `Number.NEGATIVE_INFINITY`) results from a divide-by-zero where either (but not both!) of the divide operands is negative.

JS uses finite numeric representations (IEEE 754 floating-point, which we covered earlier), so contrary to pure mathematics, it seems it *is* possible to overflow even with an operation like addition or subtraction, in which case you'd get `Infinity` or `-Infinity`.

For example:

```
var a = Number.MAX_VALUE; // 1.7976931348623157e+308
a + a; // Infinity
a + Math.pow( 2, 970 ); // Infinity
a + Math.pow( 2, 969 ); // 1.7976931348623157e+308
```

According to the specification, if an operation like addition results in a value that's too big to represent, the IEEE 754 "round-to-nearest" mode specifies what the result should be. So, in a crude sense, `Number.MAX_VALUE + Math.pow(2, 969)` is closer to `Number.MAX_VALUE` than to `Infinity`, so it "rounds down," whereas `Number.MAX_VALUE + Math.pow(2, 970)` is closer to `Infinity` so it "rounds up".

If you think too much about that, it's going to make your head hurt. So don't. Seriously, stop!

Once you overflow to either one of the *infinities*, however, there's no going back. In other words, in an almost poetic sense, you can go from finite to infinite but not from infinite back to finite.

It's almost philosophical to ask: "What is infinity divided by infinity". Our naive brains would likely say "1" or maybe "infinity." Turns out neither is true. Both mathematically and in JavaScript, `Infinity / Infinity` is not a defined operation. In JS, this results in `NaN`.

But what about any positive finite `number` divided by `Infinity`? That's easy! 0. And what about a negative finite `number` divided by `Infinity`? Keep reading!

Zeros While it may confuse the mathematics-minded reader, JavaScript has both a normal zero 0 (otherwise known as a positive zero `+0`) *and* a negative zero `-0`. Before we explain why the `-0` exists, we should examine how JS handles it, because it can be quite confusing.

Besides being specified literally as `-0`, negative zero also results from certain mathematic operations. For example:

```
var a = 0 / -3; // -0
var b = 0 * -3; // -0
```

Addition and subtraction cannot result in a negative zero.

A negative zero when examined in the developer console will usually reveal `-0`, though that was not the common case until fairly recently, so some older browsers you encounter may still report it as 0.

However, if you try to stringify a negative zero value, it will always be reported as "0", according to the spec.

```
var a = 0 / -3;

// (some browser) consoles at least get it right
a; // -0

// but the spec insists on lying to you!
a.toString(); // "0"
a + ""; // "0"
```

```

String( a );           // "0"

// strangely, even JSON gets in on the deception
JSON.stringify( a );    // "0"

```

Interestingly, the reverse operations (going from `string` to `number`) don't lie:

```

+"-0";           // -0
Number( "-0" ); // -0
JSON.parse( "-0" ); // -0

```

Warning: The `JSON.stringify(-0)` behavior of "0" is particularly strange when you observe that it's inconsistent with the reverse: `JSON.parse("-0")` reports `-0` as you'd correctly expect.

In addition to stringification of negative zero being deceptive to hide its true value, the comparison operators are also (intentionally) configured to *lie*.

```

var a = 0;
var b = 0 / -3;

a == b;      // true
-0 == 0;     // true

a === b;     // true
-0 === 0;    // true

0 > -0;     // false
a > b;       // false

```

Clearly, if you want to distinguish a `-0` from a `0` in your code, you can't just rely on what the developer console outputs, so you're going to have to be a bit more clever:

```

function isNegZero(n) {
  n = Number( n );
  return (n === 0) && (1 / n === -Infinity);
}

isNegZero( -0 );      // true
isNegZero( 0 / -3 ); // true
isNegZero( 0 );       // false

```

Now, why do we need a negative zero, besides academic trivia?

There are certain applications where developers use the magnitude of a value to represent one piece of information (like speed of movement per animation frame) and the sign of that `number` to represent another piece of information (like the direction of that movement).

In those applications, as one example, if a variable arrives at zero and it loses its sign, then you would lose the information of what direction it was moving in before it arrived at zero. Preserving the sign of the zero prevents potentially unwanted information loss.

Special Equality

As we saw above, the `NaN` value and the `-0` value have special behavior when it comes to equality comparison. `NaN` is never equal to itself, so you have to use ES6's `Number.isNaN(..)` (or a polyfill). Similarly, `-0` lies and pretends that it's equal (even `==` strict equal – see Chapter 4) to regular positive `0`, so you have to use the somewhat hackish `isNegZero(..)` utility we suggested above.

As of ES6, there's a new utility that can be used to test two values for absolute equality, without any of these exceptions. It's called `Object.is(..)`:

```
var a = 2 / "foo";
var b = -3 * 0;

Object.is( a, NaN );      // true
Object.is( b, -0 );      // true

Object.is( b, 0 );        // false
```

There's a pretty simple polyfill for `Object.is(..)` for pre-ES6 environments:

```
if (!Object.is) {
  Object.is = function(v1, v2) {
    // test for `-0`
    if (v1 === 0 && v2 === 0) {
      return 1 / v1 === 1 / v2;
    }
    // test for `NaN`
    if (v1 !== v1) {
      return v2 !== v2;
    }
    // everything else
    return v1 === v2;
  };
}
```

`Object.is(..)` probably shouldn't be used in cases where `==` or `==` are known to be *safe* (see Chapter 4 “Coercion”), as the operators are likely much more efficient and certainly are more idiomatic/common. `Object.is(..)` is mostly for these special cases of equality.

Value vs. Reference

In many other languages, values can either be assigned/passed by value-copy or by reference-copy depending on the syntax you use.

For example, in C++ if you want to pass a `number` variable into a function and have that variable's value updated, you can declare the function parameter like `int& myNum`, and when you pass in a variable like `x`, `myNum` will be a **reference to x**; references are like a special form of pointers, where you obtain a pointer to another variable (like an *alias*). If you don't declare a reference parameter, the value passed in will *always* be copied, even if it's a complex object.

In JavaScript, there are no pointers, and references work a bit differently. You cannot have a reference from one JS variable to another variable. That's just not possible.

A reference in JS points at a (shared) **value**, so if you have 10 different references, they are all always distinct references to a single shared value; **none of them are references/pointers to each other**.

Moreover, in JavaScript, there are no syntactic hints that control value vs. reference assignment/passing. Instead, the *type* of the value *solely* controls whether that value will be assigned by value-copy or by reference-copy.

Let's illustrate:

```
var a = 2;
var b = a; // `b` is always a copy of the value in `a`
b++;
a; // 2
b; // 3

var c = [1,2,3];
var d = c; // `d` is a reference to the shared `[1,2,3]` value
d.push( 4 );
c; // [1,2,3,4]
d; // [1,2,3,4]
```

Simple values (aka scalar primitives) are *always* assigned/passed by value-copy: `null`, `undefined`, `string`, `number`, `boolean`, and ES6's `symbol`.

Compound values – `objects` (including `arrays`, and all boxed object wrappers – see Chapter 3) and `functions` – *always* create a copy of the reference on assignment or passing.

In the above snippet, because `2` is a scalar primitive, `a` holds one initial copy of that value, and `b` is assigned another *copy* of the value. When changing `b`, you are in no way changing the value in `a`.

But **both c and d** are separate references to the same shared value `[1,2,3]`, which is a compound value. It's important to note that neither `c` nor `d` more

“owns” the `[1,2,3]` value – both are just equal peer references to the value. So, when using either reference to modify (`.push(4)`) the actual shared `array` value itself, it’s affecting just the one shared value, and both references will reference the newly modified value `[1,2,3,4]`.

Since references point to the values themselves and not to the variables, you cannot use one reference to change where another reference is pointed:

```
var a = [1,2,3];
var b = a;
a; // [1,2,3]
b; // [1,2,3]

// later
b = [4,5,6];
a; // [1,2,3]
b; // [4,5,6]
```

When we make the assignment `b = [4,5,6]`, we are doing absolutely nothing to affect `where a` is still referencing `([1,2,3])`. To do that, `b` would have to be a pointer to `a` rather than a reference to the `array` – but no such capability exists in JS!

The most common way such confusion happens is with function parameters:

```
function foo(x) {
  x.push( 4 );
  x; // [1,2,3,4]

  // later
  x = [4,5,6];
  x.push( 7 );
  x; // [4,5,6,7]
}

var a = [1,2,3];

foo( a );

a; // [1,2,3,4] not [4,5,6,7]
```

When we pass in the argument `a`, it assigns a copy of the `a` reference to `x`. `x` and `a` are separate references pointing at the same `[1,2,3]` value. Now, inside the function, we can use that reference to mutate the value itself (`push(4)`). But when we make the assignment `x = [4,5,6]`, this is in no way affecting where the initial reference `a` is pointing – still points at the (now modified) `[1,2,3,4]` value.

There is no way to use the `x` reference to change where `a` is pointing. We could

only modify the contents of the shared value that both `a` and `x` are pointing to.

To accomplish changing `a` to have the `[4,5,6,7]` value contents, you can't create a new `array` and assign – you must modify the existing `array` value:

```
function foo(x) {
  x.push( 4 );
  x; // [1,2,3,4]

  // later
  x.length = 0; // empty existing array in-place
  x.push( 4, 5, 6, 7 );
  x; // [4,5,6,7]
}

var a = [1,2,3];

foo( a );

a; // [4,5,6,7] not [1,2,3,4]
```

As you can see, `x.length = 0` and `x.push(4,5,6,7)` were not creating a new `array`, but modifying the existing shared `array`. So of course, `a` references the new `[4,5,6,7]` contents.

Remember: you cannot directly control/override value-copy vs. reference – those semantics are controlled entirely by the type of the underlying value.

To effectively pass a compound value (like an `array`) by value-copy, you need to manually make a copy of it, so that the reference passed doesn't still point to the original. For example:

```
foo( a.slice() );

slice(..) with no parameters by default makes an entirely new (shallow)
copy of the array. So, we pass in a reference only to the copied array, and
thus foo(..) cannot affect the contents of a.
```

To do the reverse – pass a scalar primitive value in a way where its value updates can be seen, kinda like a reference – you have to wrap the value in another compound value (`object`, `array`, etc) that *can* be passed by reference-copy:

```
function foo(wrapper) {
  wrapper.a = 42;
}

var obj = {
  a: 2
};
```

```
foo( obj );  
obj.a; // 42
```

Here, `obj` acts as a wrapper for the scalar primitive property `a`. When passed to `foo(..)`, a copy of the `obj` reference is passed in and set to the `wrapper` parameter. We now can use the `wrapper` reference to access the shared object, and update its property. After the function finishes, `obj.a` will see the updated value `42`.

It may occur to you that if you wanted to pass in a reference to a scalar primitive value like `2`, you could just box the value in its `Number` object wrapper (see Chapter 3).

It *is* true a copy of the reference to this `Number` object *will* be passed to the function, but unfortunately, having a reference to the shared object is not going to give you the ability to modify the shared primitive value, like you may expect:

```
function foo(x) {  
  x = x + 1;  
  x; // 3  
}  
  
var a = 2;  
var b = new Number(a); // or equivalently `Object(a)`  
  
foo(b);  
console.log(b); // 2, not 3
```

The problem is that the underlying scalar primitive value is *not mutable* (same goes for `String` and `Boolean`). If a `Number` object holds the scalar primitive value `2`, that exact `Number` object can never be changed to hold another value; you can only create a whole new `Number` object with a different value.

When `x` is used in the expression `x + 1`, the underlying scalar primitive value `2` is unboxed (extracted) from the `Number` object automatically, so the line `x = x + 1` very subtly changes `x` from being a shared reference to the `Number` object, to just holding the scalar primitive value `3` as a result of the addition operation `2 + 1`. Therefore, `b` on the outside still references the original unmodified/immutable `Number` object holding the value `2`.

You *can* add properties on top of the `Number` object (just not change its inner primitive value), so you could exchange information indirectly via those additional properties.

This is not all that common, however; it probably would not be considered a good practice by most developers.

Instead of using the wrapper object `Number` in this way, it's probably much better to use the manual object wrapper (`obj`) approach in the earlier snippet.

That's not to say that there's no clever uses for the boxed object wrappers like `Number` – just that you should probably prefer the scalar primitive value form in most cases.

References are quite powerful, but sometimes they get in your way, and sometimes you need them where they don't exist. The only control you have over reference vs. value-copy behavior is the type of the value itself, so you must indirectly influence the assignment/passing behavior by which value types you choose to use.

Review

In JavaScript, `arrays` are simply numerically indexed collections of any value-type. `strings` are somewhat “array-like”, but they have distinct behaviors and care must be taken if you want to treat them as `arrays`. Numbers in JavaScript include both “integers” and floating-point values.

Several special values are defined within the primitive types.

The `null` type has just one value: `null`, and likewise the `undefined` type has just the `undefined` value. `undefined` is basically the default value in any variable or property if no other value is present. The `void` operator lets you create the `undefined` value from any other value.

`numbers` include several special values, like `NaN` (supposedly “Not a Number”, but really more appropriately “invalid number”); `+Infinity` and `-Infinity`; and `-0`.

Simple scalar primitives (`strings`, `numbers`, etc.) are assigned/passed by value-copy, but compound values (`objects`, etc.) are assigned/passed by reference-copy. References are not like references/pointers in other languages – they're never pointed at other variables/references, only at the underlying values.

You Don't Know JS: Types & Grammar

Chapter 3: Natives

Several times in Chapters 1 and 2, we alluded to various built-ins, usually called “natives,” like `String` and `Number`. Let’s examine those in detail now.

Here’s a list of the most commonly used natives:

- `String()`
- `Number()`
- `Boolean()`
- `Array()`
- `Object()`
- `Function()`
- `RegExp()`
- `Date()`
- `Error()`
- `Symbol()` – added in ES6!

As you can see, these natives are actually built-in functions.

If you’re coming to JS from a language like Java, JavaScript’s `String()` will look like the `String(..)` constructor you’re used to for creating string values. So, you’ll quickly observe that you can do things like:

```
var s = new String( "Hello World!" );  
  
console.log( s.toString() ); // "Hello World!"
```

It *is* true that each of these natives can be used as a native constructor. But what’s being constructed may be different than you think.

```
var a = new String( "abc" );  
  
typeof a; // "object" ... not "String"  
  
a instanceof String; // true  
  
Object.prototype.toString.call( a ); // "[object String]"
```

The result of the constructor form of value creation (`new String("abc")`) is an object wrapper around the primitive ("abc") value.

Importantly, `typeof` shows that these objects are not their own special *types*, but more appropriately they are subtypes of the `object` type.

This object wrapper can further be observed with:

```
console.log( a );
```

The output of that statement varies depending on your browser, as developer consoles are free to choose however they feel it's appropriate to serialize the object for developer inspection.

Note: At the time of writing, the latest Chrome prints something like this:

```
String {0: "a", 1: "b", 2: "c", length: 3, [[PrimitiveValue]]:  
"abc"}. But older versions of Chrome used to just print this: String  
{0: "a", 1: "b", 2: "c"}. The latest Firefox currently prints String  
["a", "b", "c"], but used to print "abc" in italics, which was clickable to open  
the object inspector. Of course, these results are subject to rapid change and  
your experience may vary.
```

The point is, `new String("abc")` creates a string wrapper object around "abc", not just the primitive "abc" value itself.

Internal [[Class]]

Values that are `typeof "object"` (such as an array) are additionally tagged with an internal [[Class]] property (think of this more as an internal *classification* rather than related to classes from traditional class-oriented coding). This property cannot be accessed directly, but can generally be revealed indirectly by borrowing the default `Object.prototype.toString(..)` method called against the value. For example:

```
Object.prototype.toString.call( [1,2,3] ); // "[object Array]"  
  
Object.prototype.toString.call( /regex-literal/i ); // "[object RegExp]"
```

So, for the array in this example, the internal [[Class]] value is "Array", and for the regular expression, it's "RegExp". In most cases, this internal [[Class]] value corresponds to the built-in native constructor (see below) that's related to the value, but that's not always the case.

What about primitive values? First, `null` and `undefined`:

```
Object.prototype.toString.call( null ); // "[object Null]"  
Object.prototype.toString.call( undefined ); // "[object Undefined]"
```

You'll note that there are no `Null()` or `Undefined()` native constructors, but nevertheless the "Null" and "Undefined" are the internal [[Class]] values exposed.

But for the other simple primitives like `string`, `number`, and `boolean`, another behavior actually kicks in, which is usually called "boxing" (see "Boxing Wrappers" section next):

```
Object.prototype.toString.call( "abc" ); // "[object String]"  
Object.prototype.toString.call( 42 ); // "[object Number]"  
Object.prototype.toString.call( true ); // "[object Boolean]"
```

In this snippet, each of the simple primitives are automatically boxed by their respective object wrappers, which is why "String", "Number", and "Boolean" are revealed as the respective internal `[[Class]]` values.

Note: The behavior of `toString()` and `[[Class]]` as illustrated here has changed a bit from ES5 to ES6, but we cover those details in the *ES6 & Beyond* title of this series.

Boxing Wrappers

These object wrappers serve a very important purpose. Primitive values don't have properties or methods, so to access `.length` or `.toString()` you need an object wrapper around the value. Thankfully, JS will automatically *box* (aka wrap) the primitive value to fulfill such accesses.

```
var a = "abc";  
  
a.length; // 3  
a.toUpperCase(); // "ABC"
```

So, if you're going to be accessing these properties/methods on your string values regularly, like `a < a.length` condition in a `for` loop for instance, it might seem to make sense to just have the object form of the value from the start, so the JS engine doesn't need to implicitly create it for you.

But it turns out that's a bad idea. Browsers long ago performance-optimized the common cases like `.length`, which means your program will *actually go slower* if you try to "preoptimize" by directly using the object form (which isn't on the optimized path).

In general, there's basically no reason to use the object form directly. It's better to just let the boxing happen implicitly where necessary. In other words, never do things like `new String("abc")`, `new Number(42)`, etc – always prefer using the literal primitive values "abc" and 42.

Object Wrapper Gotchas

There are some gotchas with using the object wrappers directly that you should be aware of if you *do* choose to ever use them.

For example, consider `Boolean` wrapped values:

```
var a = new Boolean( false );  
  
if (!a) {  
    console.log( "Oops" ); // never runs  
}
```

The problem is that you've created an object wrapper around the `false` value, but objects themselves are "truthy" (see Chapter 4), so using the object behaves

oppositely to using the underlying `false` value itself, which is quite contrary to normal expectation.

If you want to manually box a primitive value, you can use the `Object(..)` function (no `new` keyword):

```
var a = "abc";
var b = new String( a );
var c = Object( a );

typeof a; // "string"
typeof b; // "object"
typeof c; // "object"

b instanceof String; // true
c instanceof String; // true

Object.prototype.toString.call( b ); // "[object String]"
Object.prototype.toString.call( c ); // "[object String]"
```

Again, using the boxed object wrapper directly (like `b` and `c` above) is usually discouraged, but there may be some rare occasions you'll run into where they may be useful.

Unboxing

If you have an object wrapper and you want to get the underlying primitive value out, you can use the `valueOf()` method:

```
var a = new String( "abc" );
var b = new Number( 42 );
var c = new Boolean( true );

a.valueOf(); // "abc"
b.valueOf(); // 42
c.valueOf(); // true
```

Unboxing can also happen implicitly, when using an object wrapper value in a way that requires the primitive value. This process (coercion) will be covered in more detail in Chapter 4, but briefly:

```
var a = new String( "abc" );
var b = a + ""; // `b` has the unboxed primitive value "abc"

typeof a; // "object"
typeof b; // "string"
```

Natives as Constructors

For `array`, `object`, `function`, and regular-expression values, it's almost universally preferred that you use the literal form for creating the values, but the literal form creates the same sort of object as the constructor form does (that is, there is no nonwrapped value).

Just as we've seen above with the other natives, these constructor forms should generally be avoided, unless you really know you need them, mostly because they introduce exceptions and gotchas that you probably don't really *want* to deal with.

```
Array(..)

var a = new Array( 1, 2, 3 );
a; // [1, 2, 3]

var b = [1, 2, 3];
b; // [1, 2, 3]
```

Note: The `Array(..)` constructor does not require the `new` keyword in front of it. If you omit it, it will behave as if you have used it anyway. So `Array(1,2,3)` is the same outcome as `new Array(1,2,3)`.

The `Array` constructor has a special form where if only one `number` argument is passed, instead of providing that value as *contents* of the array, it's taken as a `length` to "presize the array" (well, sorta).

This is a terrible idea. Firstly, you can trip over that form accidentally, as it's easy to forget.

But more importantly, there's no such thing as actually presizing the array. Instead, what you're creating is an otherwise empty array, but setting the `length` property of the array to the numeric value specified.

An array that has no explicit values in its slots, but has a `length` property that *implies* the slots exist, is a weird exotic type of data structure in JS with some very strange and confusing behavior. The capability to create such a value comes purely from old, deprecated, historical functionalities ("array-like objects" like the `arguments` object).

Note: An array with at least one "empty slot" in it is often called a "sparse array."

It doesn't help matters that this is yet another example where browser developer consoles vary on how they represent such an object, which breeds more confusion.

For example:

```
var a = new Array( 3 );  
  
a.length; // 3  
a;
```

The serialization of `a` in Chrome is (at the time of writing): `[undefined x 3]`. **This is really unfortunate.** It implies that there are three `undefined` values in the slots of this array, when in fact the slots do not exist (so-called “empty slots” – also a bad name!).

To visualize the difference, try this:

```
var a = new Array( 3 );  
var b = [ undefined, undefined, undefined ];  
var c = [];  
c.length = 3;  
  
a;  
b;  
c;
```

Note: As you can see with `c` in this example, empty slots in an array can happen after creation of the array. Changing the `length` of an array to go beyond its number of actually-defined slot values, you implicitly introduce empty slots. In fact, you could even call `delete b[1]` in the above snippet, and it would introduce an empty slot into the middle of `b`.

For `b` (in Chrome, currently), you’ll find `[undefined, undefined, undefined]` as the serialization, as opposed to `[undefined x 3]` for `a` and `c`. Confused? Yeah, so is everyone else.

Worse than that, at the time of writing, Firefox reports `[, , ,]` for `a` and `c`. Did you catch why that’s so confusing? Look closely. Three commas implies four slots, not three slots like we’d expect.

What!? Firefox puts an extra `,` on the end of their serialization here because as of ES5, trailing commas in lists (array values, property lists, etc.) are allowed (and thus dropped and ignored). So if you were to type in a `[, , ,]` value into your program or the console, you’d actually get the underlying value that’s like `[, ,]` (that is, an array with three empty slots). This choice, while confusing if reading the developer console, is defended as instead making copy-n-paste behavior accurate.

If you’re shaking your head or rolling your eyes about now, you’re not alone! Shrugs.

Unfortunately, it gets worse. More than just confusing console output, `a` and `b` from the above code snippet actually behave the same in some cases **but differently in others**:

```

a.join( "-" ); // "--"
b.join( "-" ); // "--"

a.map(function(v,i){ return i; }); // [ undefined x 3 ]
b.map(function(v,i){ return i; }); // [ 0, 1, 2 ]

```

Ugh.

The `a.map(..)` call *fails* because the slots don't actually exist, so `map(..)` has nothing to iterate over. `join(..)` works differently. Basically, we can think of it implemented sort of like this:

```

function fakeJoin(arr, connector) {
  var str = "";
  for (var i = 0; i < arr.length; i++) {
    if (i > 0) {
      str += connector;
    }
    if (arr[i] !== undefined) {
      str += arr[i];
    }
  }
  return str;
}

var a = new Array( 3 );
fakeJoin( a, "-" ); // "--"

```

As you can see, `join(..)` works by just *assuming* the slots exist and looping up to the `length` value. Whatever `map(..)` does internally, it (apparently) doesn't make such an assumption, so the result from the strange "empty slots" array is unexpected and likely to cause failure.

So, if you wanted to *actually* create an array of actual `undefined` values (not just "empty slots"), how could you do it (besides manually)?

```

var a = Array.apply( null, { length: 3 } );
a; // [ undefined, undefined, undefined ]

```

Confused? Yeah. Here's roughly how it works.

`apply(..)` is a utility available to all functions, which calls the function it's used with but in a special way.

The first argument is a `this` object binding (covered in the *this & Object Prototypes* title of this series), which we don't care about here, so we set it to `null`. The second argument is supposed to be an array (or something *like* an array – aka an "array-like object"). The contents of this "array" are "spread" out as arguments to the function in question.

So, `Array.apply(..)` is calling the `Array(..)` function and spreading out the values (of the `{ length: 3 }` object value) as its arguments.

Inside of `apply(..)`, we can envision there's another `for` loop (kinda like `join(..)` from above) that goes from 0 up to, but not including, `length` (3 in our case).

For each index, it retrieves that key from the object. So if the array-object parameter was named `arr` internally inside of the `apply(..)` function, the property access would effectively be `arr[0]`, `arr[1]`, and `arr[2]`. Of course, none of those properties exist on the `{ length: 3 }` object value, so all three of those property accesses would return the value `undefined`.

In other words, it ends up calling `Array(..)` basically like this: `Array(undefined,undefined,undefined)`, which is how we end up with an array filled with `undefined` values, and not just those (crazy) empty slots.

While `Array.apply(null, { length: 3 })` is a strange and verbose way to create an array filled with `undefined` values, it's **vastly** better and more reliable than what you get with the footgun'ish `Array(3)` empty slots.

Bottom line: **never ever, under any circumstances**, should you intentionally create and use these exotic empty-slot arrays. Just don't do it. They're nuts.

`Object(..)`, `Function(..)`, and `RegExp(..)`

The `Object(..)/Function(..)/RegExp(..)` constructors are also generally optional (and thus should usually be avoided unless specifically called for):

```
var c = new Object();
c.foo = "bar";
c; // { foo: "bar" }

var d = { foo: "bar" };
d; // { foo: "bar" }

var e = new Function( "a", "return a * 2;" );
var f = function(a) { return a * 2; };
function g(a) { return a * 2; }

var h = new RegExp( "^a*b+", "g" );
var i = /^a*b+/g;
```

There's practically no reason to ever use the `new Object()` constructor form, especially since it forces you to add properties one-by-one instead of many at once in the object literal form.

The `Function` constructor is helpful only in the rarest of cases, where you need to dynamically define a function's parameters and/or its function body. **Do**

not just treat `Function(..)` as an alternate form of `eval(..)`. You will almost never need to dynamically define a function in this way.

Regular expressions defined in the literal form (`/^a*b+/g`) are strongly preferred, not just for ease of syntax but for performance reasons – the JS engine precompiles and caches them before code execution. Unlike the other constructor forms we've seen so far, `RegExp(..)` has some reasonable utility: to dynamically define the pattern for a regular expression.

```
var name = "Kyle";
var namePattern = new RegExp( "\\b(?: " + name + ")+\\b", "ig" );

var matches = someText.match( namePattern );
```

This kind of scenario legitimately occurs in JS programs from time to time, so you'd need to use the `new RegExp("pattern","flags")` form.

`Date(..)` and `Error(..)`

The `Date(..)` and `Error(..)` native constructors are much more useful than the other natives, because there is no literal form for either.

To create a date object value, you must use `new Date()`. The `Date(..)` constructor accepts optional arguments to specify the date/time to use, but if omitted, the current date/time is assumed.

By far the most common reason you construct a date object is to get the current timestamp value (a signed integer number of milliseconds since Jan 1, 1970). You can do this by calling `getTime()` on a date object instance.

But an even easier way is to just call the static helper function defined as of ES5: `Date.now()`. And to polyfill that for pre-ES5 is pretty easy:

```
if (!Date.now) {
  Date.now = function(){
    return (new Date()).getTime();
  };
}
```

Note: If you call `Date()` without `new`, you'll get back a string representation of the date/time at that moment. The exact form of this representation is not specified in the language spec, though browsers tend to agree on something close to: "Fri Jul 18 2014 00:31:02 GMT-0500 (CDT)".

The `Error(..)` constructor (much like `Array()` above) behaves the same with the `new` keyword present or omitted.

The main reason you'd want to create an error object is that it captures the current execution stack context into the object (in most JS engines, revealed as a read-only `.stack` property once constructed). This stack context includes

the function call-stack and the line-number where the error object was created, which makes debugging that error much easier.

You would typically use such an error object with the `throw` operator:

```
function foo(x) {
  if (!x) {
    throw new Error( "x wasn't provided" );
  }
  // ...
}
```

Error object instances generally have at least a `message` property, and sometimes other properties (which you should treat as read-only), like `type`. However, other than inspecting the above-mentioned `stack` property, it's usually best to just call `toString()` on the error object (either explicitly, or implicitly through coercion – see Chapter 4) to get a friendly-formatted error message.

Tip: Technically, in addition to the general `Error(..)` native, there are several other specific-error-type natives: `EvalError(..)`, `RangeError(..)`, `ReferenceError(..)`, `SyntaxError(..)`, `TypeError(..)`, and `URIError(..)`. But it's very rare to manually use these specific error natives. They are automatically used if your program actually suffers from a real exception (such as referencing an undeclared variable and getting a `ReferenceError` error).

`Symbol(..)`

New as of ES6, an additional primitive value type has been added, called “`Symbol`”. `Symbol`s are special “unique” (not strictly guaranteed!) values that can be used as properties on objects with little fear of any collision. They're primarily designed for special built-in behaviors of ES6 constructs, but you can also define your own `Symbol`s.

`Symbol`s can be used as property names, but you cannot see or access the actual value of a `Symbol` from your program, nor from the developer console. If you evaluate a `Symbol` in the developer console, what's shown looks like `Symbol(Symbol.create)`, for example.

There are several predefined `Symbol`s in ES6, accessed as static properties of the `Symbol` function object, like `Symbol.create`, `Symbol.iterator`, etc. To use them, do something like:

```
obj[Symbol.iterator] = function(){ /*...*/ };
```

To define your own custom `Symbol`s, use the `Symbol(..)` native. The `Symbol(..)` native “constructor” is unique in that you're not allowed to use `new` with it, as doing so will throw an error.

```
var mysym = Symbol( "my own symbol" );
mysym; // Symbol(my own symbol)
```

```

mysym.toString();    // "Symbol(my own symbol)"
typeof mysym;       // "symbol"

var a = {};
a[mysym] = "foobar";

Object.getOwnPropertySymbols( a );
// [ Symbol(my own symbol) ]

```

While symbols are not actually private (`Object.getOwnPropertySymbols(..)` reflects on the object and reveals the symbols quite publicly), using them for private or special properties is likely their primary use-case. For most developers, they may take the place of property names with `_` underscore prefixes, which are almost always by convention signals to say, “hey, this is a private/special/internal property, so leave it alone!”

Note: Symbols are *not objects*, they are simple scalar primitives.

Native Prototypes

Each of the built-in native constructors has its own `.prototype` object – `Array.prototype`, `String.prototype`, etc.

These objects contain behavior unique to their particular object subtype.

For example, all string objects, and by extension (via boxing) `string` primitives, have access to default behavior as methods defined on the `String.prototype` object.

Note: By documentation convention, `String.prototype.XYZ` is shortened to `String#XYZ`, and likewise for all the other `.prototypes`.

- `String#indexOf(..)`: find the position in the string of another substring
- `String#charAt(..)`: access the character at a position in the string
- `String#substr(..)`, `String#substring(..)`, and `String#slice(..)`: extract a portion of the string as a new string
- `String#toUpperCase()` and `String#toLowerCase()`: create a new string that’s converted to either uppercase or lowercase
- `String#trim()`: create a new string that’s stripped of any trailing or leading whitespace

None of the methods modify the string *in place*. Modifications (like case conversion or trimming) create a new value from the existing value.

By virtue of prototype delegation (see the *this & Object Prototypes* title in this series), any string value can access these methods:

```

var a = " abc ";

a.indexOf( "c" ); // 3

```

```
a.toUpperCase(); // " ABC "
a.trim(); // "abc"
```

The other constructor prototypes contain behaviors appropriate to their types, such as `Number#toFixed(..)` (stringifying a number with a fixed number of decimal digits) and `Array#concat(..)` (merging arrays). All functions have access to `apply(..)`, `call(..)`, and `bind(..)` because `Function.prototype` defines them.

But, some of the native prototypes aren't *just* plain objects:

```
typeof Function.prototype;           // "function"
Function.prototype();               // it's an empty function!

RegExp.prototype.toString();        // "/(?:)/" -- empty regex
"abc".match( RegExp.prototype );   // [""]
```

A particularly bad idea, you can even modify these native prototypes (not just adding properties as you're probably familiar with):

```
Array.isArray( Array.prototype );   // true
Array.prototype.push( 1, 2, 3 );    // 3
Array.prototype;                  // [1,2,3]

// don't leave it that way, though, or expect weirdness!
// reset the `Array.prototype` to empty
Array.prototype.length = 0;
```

As you can see, `Function.prototype` is a function, `RegExp.prototype` is a regular expression, and `Array.prototype` is an array. Interesting and cool, huh?

Prototypes As Defaults `Function.prototype` being an empty function, `RegExp.prototype` being an "empty" (e.g., non-matching) regex, and `Array.prototype` being an empty array, make them all nice "default" values to assign to variables if those variables wouldn't already have had a value of the proper type.

For example:

```
function isThisCool(vals,fn,rx) {
  vals = vals || Array.prototype;
  fn = fn || Function.prototype;
  rx = rx || RegExp.prototype;

  return rx.test(
    vals.map( fn ).join( "" )
  );
}
```

```

isThisCool();           // true

isThisCool(
  ["a", "b", "c"],
  function(v){ return v.toUpperCase(); },
  /D/
);
// false

```

Note: As of ES6, we don't need to use the `vals = vals || ..` default value syntax trick (see Chapter 4) anymore, because default values can be set for parameters via native syntax in the function declaration (see Chapter 5).

One minor side-benefit of this approach is that the `.prototypes` are already created and built-in, thus created *only once*. By contrast, using `[]`, `function(){}, and /(?:)/ values themselves for those defaults would (likely, depending on engine implementations) be recreating those values (and probably garbage-collecting them later) for each call of isThisCool(..). That could be memory/CPU wasteful.`

Also, be very careful not to use `Array.prototype` as a default value **that will subsequently be modified**. In this example, `vals` is used read-only, but if you were to instead make in-place changes to `vals`, you would actually be modifying `Array.prototype` itself, which would lead to the gotchas mentioned earlier!

Note: While we're pointing out these native prototypes and some usefulness, be cautious of relying on them and even more wary of modifying them in any way. See Appendix A “Native Prototypes” for more discussion.

Review

JavaScript provides object wrappers around primitive values, known as natives (`String`, `Number`, `Boolean`, etc). These object wrappers give the values access to behaviors appropriate for each object subtype (`String#trim()` and `Array#concat(..)`).

If you have a simple scalar primitive value like `"abc"` and you access its `length` property or some `String.prototype` method, JS automatically “boxes” the value (wraps it in its respective object wrapper) so that the property/method accesses can be fulfilled.

You Don't Know JS: Types & Grammar

Chapter 4: Coercion

Now that we much more fully understand JavaScript's types and values, we turn our attention to a very controversial topic: coercion.

As we mentioned in Chapter 1, the debates over whether coercion is a useful feature or a flaw in the design of the language (or somewhere in between!) have raged since day one. If you've read other popular books on JS, you know that the overwhelmingly prevalent *message* out there is that coercion is magical, evil, confusing, and just downright a bad idea.

In the same overall spirit of this book series, rather than running away from coercion because everyone else does, or because you get bitten by some quirk, I think you should run toward that which you don't understand and seek to *get it* more fully.

Our goal is to fully explore the pros and cons (yes, there *are* pros!) of coercion, so that you can make an informed decision on its appropriateness in your program.

Converting Values

Converting a value from one type to another is often called "type casting," when done explicitly, and "coercion" when done implicitly (forced by the rules of how a value is used).

Note: It may not be obvious, but JavaScript coercions always result in one of the scalar primitive (see Chapter 2) values, like `string`, `number`, or `boolean`. There is no coercion that results in a complex value like `object` or `function`. Chapter 3 covers "boxing," which wraps scalar primitive values in their `object` counterparts, but this is not really coercion in an accurate sense.

Another way these terms are often distinguished is as follows: "type casting" (or "type conversion") occur in statically typed languages at compile time, while "type coercion" is a runtime conversion for dynamically typed languages.

However, in JavaScript, most people refer to all these types of conversions as *coercion*, so the way I prefer to distinguish is to say "implicit coercion" vs. "explicit coercion."

The difference should be obvious: "explicit coercion" is when it is obvious from looking at the code that a type conversion is intentionally occurring, whereas "implicit coercion" is when the type conversion will occur as a less obvious side effect of some other intentional operation.

For example, consider these two approaches to coercion:

```
var a = 42;
```

```
var b = a + "";           // implicit coercion  
  
var c = String(a);      // explicit coercion
```

For `b`, the coercion that occurs happens implicitly, because the `+` operator combined with one of the operands being a `string` value ("") will insist on the operation being a `string` concatenation (adding two strings together), which as a *hidden side effect* will force the `42` value in `a` to be coerced to its `string` equivalent: `"42"`.

By contrast, the `String(..)` function makes it pretty obvious that it's explicitly taking the value in `a` and coercing it to a `string` representation.

Both approaches accomplish the same effect: `"42"` comes from `42`. But it's the *how* that is at the heart of the heated debates over JavaScript coercion.

Note: Technically, there's some nuanced behavioral difference here beyond the stylistic difference. We cover that in more detail later in the chapter, in the "Implicitly: Strings \leftrightarrow Numbers" section.

The terms "explicit" and "implicit," or "obvious" and "hidden side effect," are *relative*.

If you know exactly what `a + ""` is doing and you're intentionally doing that to coerce to a `string`, you might feel the operation is sufficiently "explicit." Conversely, if you've never seen the `String(..)` function used for `string` coercion, its behavior might seem hidden enough as to feel "implicit" to you.

But we're having this discussion of "explicit" vs. "implicit" based on the likely opinions of an *average, reasonably informed, but not expert or JS specification devotee* developer. To whatever extent you do or do not find yourself fitting neatly in that bucket, you will need to adjust your perspective on our observations here accordingly.

Just remember: it's often rare that we write our code and are the only ones who ever read it. Even if you're an expert on all the ins and outs of JS, consider how a less experienced teammate of yours will feel when they read your code. Will it be "explicit" or "implicit" to them in the same way it is for you?

Abstract Value Operations

Before we can explore *explicit* vs *implicit* coercion, we need to learn the basic rules that govern how values *become* either a `string`, `number`, or `boolean`. The ES5 spec in section 9 defines several "abstract operations" (fancy spec-speak for "internal-only operation") with the rules of value conversion. We will specifically pay attention to: `ToString`, `ToNumber`, and `ToBoolean`, and to a lesser extent, `ToPrimitive`.

ToString

When any **non-string** value is coerced to a **string** representation, the conversion is handled by the **ToString** abstract operation in section 9.8 of the specification.

Built-in primitive values have natural stringification: `null` becomes `"null"`, `undefined` becomes `"undefined"` and `true` becomes `"true"`. **numbers** are generally expressed in the natural way you'd expect, but as we discussed in Chapter 2, very small or very large **numbers** are represented in exponent form:

```
// multiplying `1.07` by `1000`, seven times over
var a = 1.07 * 1000 * 1000 * 1000 * 1000 * 1000 * 1000;

// seven times three digits => 21 digits
a.toString(); // "1.07e21"
```

For regular objects, unless you specify your own, the default `toString()` (located in `Object.prototype.toString()`) will return the *internal [[Class]]* (see Chapter 3), like for instance `"[object Object]"`.

But as shown earlier, if an object has its own `toString()` method on it, and you use that object in a **string**-like way, its `toString()` will automatically be called, and the **string** result of that call will be used instead.

Note: The way an object is coerced to a **string** technically goes through the **ToPrimitive** abstract operation (ES5 spec, section 9.1), but those nuanced details are covered in more detail in the **ToNumber** section later in this chapter, so we will skip over them here.

Arrays have an overridden default `toString()` that stringifies as the (string) concatenation of all its values (each stringified themselves), with `,` in between each value:

```
var a = [1,2,3];

a.toString(); // "1,2,3"
```

Again, `toString()` can either be called explicitly, or it will automatically be called if a **non-string** is used in a **string** context.

JSON Stringification Another task that seems awfully related to **ToString** is when you use the `JSON.stringify(..)` utility to serialize a value to a JSON-compatible **string** value.

It's important to note that this stringification is not exactly the same thing as coercion. But since it's related to the **ToString** rules above, we'll take a slight diversion to cover JSON stringification behaviors here.

For most simple values, JSON stringification behaves basically the same as `toString()` conversions, except that the serialization result is *always a string*:

```
JSON.stringify( 42 ); // "42"
JSON.stringify( "42" ); // ""42"" (a string with a quoted string value in it)
JSON.stringify( null ); // "null"
JSON.stringify( true ); // "true"
```

Any *JSON-safe* value can be stringified by `JSON.stringify(..)`. But what is *JSON-safe*? Any value that can be represented validly in a JSON representation.

It may be easier to consider values that are **not** JSON-safe. Some examples: `undefineds`, `functions`, (ES6+) `symbols`, and `objects` with circular references (where property references in an object structure create a never-ending cycle through each other). These are all illegal values for a standard JSON structure, mostly because they aren't portable to other languages that consume JSON values.

The `JSON.stringify(..)` utility will automatically omit `undefined`, `function`, and `symbol` values when it comes across them. If such a value is found in an `array`, that value is replaced by `null` (so that the array position information isn't altered). If found as a property of an `object`, that property will simply be excluded.

Consider:

```
JSON.stringify( undefined ); // undefined
JSON.stringify( function(){} ); // undefined

JSON.stringify( [1,undefined,function(){},4] ); // "[1,null,null,4]"
JSON.stringify( { a:2, b:function(){} } ); // {"a":2}
```

But if you try to `JSON.stringify(..)` an `object` with circular reference(s) in it, an error will be thrown.

JSON stringification has the special behavior that if an `object` value has a `toJSON()` method defined, this method will be called first to get a value to use for serialization.

If you intend to JSON stringify an object that may contain illegal JSON value(s), or if you just have values in the `object` that aren't appropriate for the serialization, you should define a `toJSON()` method for it that returns a *JSON-safe* version of the `object`.

For example:

```
var o = { };

var a = {
  b: 42,
  c: o,
  d: function(){}
};
```

```

// create a circular reference inside `a`
o.e = a;

// would throw an error on the circular reference
// JSON.stringify( a );

// define a custom JSON value serialization
a.toJSON = function() {
  // only include the `b` property for serialization
  return { b: this.b };
};

JSON.stringify( a ); // {"b":42}

```

It's a very common misconception that `toJSON()` should return a JSON stringification representation. That's probably incorrect, unless you're wanting to actually stringify the `string` itself (usually not!). `toJSON()` should return the actual regular value (of whatever type) that's appropriate, and `JSON.stringify(..)` itself will handle the stringification.

In other words, `toJSON()` should be interpreted as “to a JSON-safe value suitable for stringification,” not “to a JSON string” as many developers mistakenly assume.

Consider:

```

var a = {
  val: [1,2,3],

  // probably correct!
  toJSON: function(){
    return this.val.slice( 1 );
  }
};

var b = {
  val: [1,2,3],

  // probably incorrect!
  toJSON: function(){
    return "[" +
      this.val.slice( 1 ).join() +
    "]";
  }
};

JSON.stringify( a ); // "[2,3]"

```

```
JSON.stringify( b ); // "[2,3]"
```

In the second call, we stringified the returned `string` rather than the `array` itself, which was probably not what we wanted to do.

While we're talking about `JSON.stringify(..)`, let's discuss some lesser-known functionalities that can still be very useful.

An optional second argument can be passed to `JSON.stringify(..)` that is called *replacer*. This argument can either be an `array` or a `function`. It's used to customize the recursive serialization of an `object` by providing a filtering mechanism for which properties should and should not be included, in a similar way to how `toJSON()` can prepare a value for serialization.

If *replacer* is an `array`, it should be an `array` of `strings`, each of which will specify a property name that is allowed to be included in the serialization of the `object`. If a property exists that isn't in this list, it will be skipped.

If *replacer* is a `function`, it will be called once for the `object` itself, and then once for each property in the `object`, and each time is passed two arguments, `key` and `value`. To skip a `key` in the serialization, return `undefined`. Otherwise, return the `value` provided.

```
var a = {
  b: 42,
  c: "42",
  d: [1,2,3]
};

JSON.stringify( a, ["b","c"] ); // {"b":42,"c":"42"}
```

```
JSON.stringify( a, function(k,v){
  if (k !== "c") return v;
} );
// {"b":42,"d":[1,2,3]}
```

Note: In the `function replacer` case, the `key` argument `k` is `undefined` for the first call (where the `a` object itself is being passed in). The `if` statement **filters out** the property named "c". Stringification is recursive, so the `[1,2,3]` array has each of its values (1, 2, and 3) passed as `v` to *replacer*, with indexes (0, 1, and 2) as `k`.

A third optional argument can also be passed to `JSON.stringify(..)`, called `space`, which is used as indentation for prettier human-friendly output. `space` can be a positive integer to indicate how many space characters should be used at each indentation level. Or, `space` can be a `string`, in which case up to the first ten characters of its value will be used for each indentation level.

```
var a = {
  b: 42,
```

```

    c: "42",
    d: [1,2,3]
};

JSON.stringify( a, null, 3 );
// "{  

//   "b": 42,  

//   "c": "42",  

//   "d": [  

//     1,  

//     2,  

//     3  

//   ]  

// }"  

JSON.stringify( a, null, "-----" );
// {"  

// -----"b": 42,  

// -----"c": "42",  

// -----"d": [  

// -----1,  

// -----2,  

// -----3  

// -----]  

// }"  


```

Remember, `JSON.stringify(..)` is not directly a form of coercion. We covered it here, however, for two reasons that relate its behavior to `ToString` coercion:

1. `string`, `number`, `boolean`, and `null` values all `stringify` for JSON basically the same as how they coerce to `string` values via the rules of the `ToString` abstract operation.
2. If you pass an `object` value to `JSON.stringify(..)`, and that `object` has a `toJSON()` method on it, `toJSON()` is automatically called to (sort of) “coerce” the value to be *JSON-safe* before stringification.

ToNumber

If any non-`number` value is used in a way that requires it to be a `number`, such as a mathematical operation, the ES5 spec defines the `ToNumber` abstract operation in section 9.3.

For example, `true` becomes `1` and `false` becomes `0`. `undefined` becomes `NaN`, but (curiously) `null` becomes `0`.

`ToNumber` for a `string` value essentially works for the most part like the rules/syntax for numeric literals (see Chapter 3). If it fails, the result is `NaN` (instead of a syntax error as with `number` literals). One example difference is

that 0-prefixed octal numbers are not handled as octals (just as normal base-10 decimals) in this operation, though such octals are valid as `number` literals (see Chapter 2).

Note: The differences between `number` literal grammar and `ToNumber` on a `string` value are subtle and highly nuanced, and thus will not be covered further here. Consult section 9.3.1 of the ES5 spec for more information.

Objects (and arrays) will first be converted to their primitive value equivalent, and the resulting value (if a primitive but not already a `number`) is coerced to a `number` according to the `ToNumber` rules just mentioned.

To convert to this primitive value equivalent, the `ToPrimitive` abstract operation (ES5 spec, section 9.1) will consult the value (using the internal `DefaultValue` operation – ES5 spec, section 8.12.8) in question to see if it has a `valueOf()` method. If `valueOf()` is available and it returns a primitive value, *that* value is used for the coercion. If not, but `toString()` is available, it will provide the value for the coercion.

If neither operation can provide a primitive value, a `TypeError` is thrown.

As of ES5, you can create such a noncoercible object – one without `valueOf()` and `toString()` – if it has a `null` value for its `[[Prototype]]`, typically created with `Object.create(null)`. See the *this & Object Prototypes* title of this series for more information on `[[Prototype]]`s.

Note: We cover how to coerce to `numbers` later in this chapter in detail, but for this next code snippet, just assume the `Number(..)` function does so.

Consider:

```
var a = {
  valueOf: function(){
    return "42";
  }
};

var b = {
  toString: function(){
    return "42";
  }
};

var c = [4,2];
c.toString = function(){
  return this.join( ""); // "42"
};

Number( a );           // 42
Number( b );           // 42
```

```
Number( c );           // 42
Number( "" );          // 0
Number( [] );          // 0
Number( [ "abc" ] );   // NaN
```

ToBoolean

Next, let's have a little chat about how `booleans` behave in JS. There's **lots of confusion and misconception** floating out there around this topic, so pay close attention!

First and foremost, JS has actual keywords `true` and `false`, and they behave exactly as you'd expect of `boolean` values. It's a common misconception that the values `1` and `0` are identical to `true/false`. While that may be true in other languages, in JS the `numbers` are `numbers` and the `booleans` are `booleans`. You can coerce `1` to `true` (and vice versa) or `0` to `false` (and vice versa). But they're not the same.

Falsy Values But that's not the end of the story. We need to discuss how values other than the two `booleans` behave whenever you coerce *to* their `boolean` equivalent.

All of JavaScript's values can be divided into two categories:

1. values that will become `false` if coerced to `boolean`
2. everything else (which will obviously become `true`)

I'm not just being facetious. The JS spec defines a specific, narrow list of values that will coerce to `false` when coerced to a `boolean` value.

How do we know what the list of values is? In the ES5 spec, section 9.2 defines a `ToBoolean` abstract operation, which says exactly what happens for all the possible values when you try to coerce them "to boolean."

From that table, we get the following as the so-called "falsy" values list:

- `undefined`
- `null`
- `false`
- `+0, -0, and NaN`
- `""`

That's it. If a value is on that list, it's a "falsy" value, and it will coerce to `false` if you force a `boolean` coercion on it.

By logical conclusion, if a value is *not* on that list, it must be on *another list*, which we call the "truthy" values list. But JS doesn't really define a "truthy" list per se. It gives some examples, such as saying explicitly that all objects are truthy, but mostly the spec just implies: **anything not explicitly on the falsy list is therefore truthy**.

Falsy Objects Wait a minute, that section title even sounds contradictory. I literally *just said* the spec calls all objects truthy, right? There should be no such thing as a “falsy object.”

What could that possibly even mean?

You might be tempted to think it means an object wrapper (see Chapter 3) around a falsy value (such as "", 0 or `false`). But don’t fall into that *trap*.

Note: That’s a subtle specification joke some of you may get.

Consider:

```
var a = new Boolean( false );
var b = new Number( 0 );
var c = new String( "" );
```

We know all three values here are objects (see Chapter 3) wrapped around obviously falsy values. But do these objects behave as `true` or as `false`? That’s easy to answer:

```
var d = Boolean( a && b && c );

d; // true
```

So, all three behave as `true`, as that’s the only way `d` could end up as `true`.

Tip: Notice the `Boolean(...)` wrapped around the `a && b && c` expression – you might wonder why that’s there. We’ll come back to that later in this chapter, so make a mental note of it. For a sneak-peek (trivia-wise), try for yourself what `d` will be if you just do `d = a && b && c` without the `Boolean(...)` call!

So, if “falsy objects” are **not just objects wrapped around falsy values**, what the heck are they?

The tricky part is that they can show up in your JS program, but they’re not actually part of JavaScript itself.

What!?

There are certain cases where browsers have created their own sort of *exotic* values behavior, namely this idea of “falsy objects,” on top of regular JS semantics.

A “falsy object” is a value that looks and acts like a normal object (properties, etc.), but when you coerce it to a `boolean`, it coerces to a `false` value.

Why!?

The most well-known case is `document.all`: an array-like (object) provided to your JS program *by the DOM* (not the JS engine itself), which exposes elements in your page to your JS program. It *used* to behave like a normal object—it would act truthy. But not anymore.

`document.all` itself was never really “standard” and has long since been deprecated/abandoned.

“Can’t they just remove it, then?” Sorry, nice try. Wish they could. But there’s far too many legacy JS code bases out there that rely on using it.

So, why make it act falsy? Because coercions of `document.all` to `boolean` (like in `if` statements) were almost always used as a means of detecting old, nonstandard IE.

IE has long since come up to standards compliance, and in many cases is pushing the web forward as much or more than any other browser. But all that old `if (document.all) { /* it's IE */ }` code is still out there, and much of it is probably never going away. All this legacy code is still assuming it’s running in decade-old IE, which just leads to bad browsing experience for IE users.

So, we can’t remove `document.all` completely, but IE doesn’t want `if (document.all) { .. }` code to work anymore, so that users in modern IE get new, standards-compliant code logic.

“What should we do?” **“I’ve got it! Let’s bastardize the JS type system and pretend that `document.all` is falsy!”

Ugh. That sucks. It’s a crazy gotcha that most JS developers don’t understand. But the alternative (doing nothing about the above no-win problems) sucks *just a little bit more*.

So... that’s what we’ve got: crazy, nonstandard “falsy objects” added to JavaScript by the browsers. Yay!

Truthy Values Back to the truthy list. What exactly are the truthy values? Remember: **a value is truthy if it’s not on the falsy list**.

Consider:

```
var a = "false";
var b = "0";
var c = " ";

var d = Boolean( a && b && c );

d;
```

What value do you expect `d` to have here? It’s gotta be either `true` or `false`.

It’s `true`. Why? Because despite the contents of those `string` values looking like falsy values, the `string` values themselves are all truthy, because `""` is the only `string` value on the falsy list.

What about these?

```

var a = [] ;           // empty array -- truthy or falsy?
var b = {} ;           // empty object -- truthy or falsy?
var c = function(){}; // empty function -- truthy or falsy?

var d = Boolean( a && b && c );

d;

```

Yep, you guessed it, `d` is still `true` here. Why? Same reason as before. Despite what it may seem like, `[]`, `{}`, and `function(){}` are *not* on the falsy list, and thus are truthy values.

In other words, the truthy list is infinitely long. It's impossible to make such a list. You can only make a finite falsy list and consult *it*.

Take five minutes, write the falsy list on a post-it note for your computer monitor, or memorize it if you prefer. Either way, you'll easily be able to construct a virtual truthy list whenever you need it by simply asking if it's on the falsy list or not.

The importance of truthy and falsy is in understanding how a value will behave if you coerce it (either explicitly or implicitly) to a `boolean` value. Now that you have those two lists in mind, we can dive into coercion examples themselves.

Explicit Coercion

Explicit coercion refers to type conversions that are obvious and explicit. There's a wide range of type conversion usage that clearly falls under the *explicit* coercion category for most developers.

The goal here is to identify patterns in our code where we can make it clear and obvious that we're converting a value from one type to another, so as to not leave potholes for future developers to trip into. The more explicit we are, the more likely someone later will be able to read our code and understand without undue effort what our intent was.

It would be hard to find any salient disagreements with *explicit* coercion, as it most closely aligns with how the commonly accepted practice of type conversion works in statically typed languages. As such, we'll take for granted (for now) that *explicit* coercion can be agreed upon to not be evil or controversial. We'll revisit this later, though.

Explicitly: Strings <-> Numbers

We'll start with the simplest and perhaps most common coercion operation: coercing values between `string` and `number` representation.

To coerce between `strings` and `numbers`, we use the built-in `String(...)` and `Number(...)` functions (which we referred to as "native constructors" in Chapter

3), but **very importantly**, we do not use the `new` keyword in front of them. As such, we're not creating object wrappers.

Instead, we're actually *explicitly coercing* between the two types:

```
var a = 42;
var b = String( a );

var c = "3.14";
var d = Number( c );

b; // "42"
d; // 3.14
```

`String(..)` coerces from any other value to a primitive `string` value, using the rules of the `ToString` operation discussed earlier. `Number(..)` coerces from any other value to a primitive `number` value, using the rules of the `ToNumber` operation discussed earlier.

I call this *explicit* coercion because in general, it's pretty obvious to most developers that the end result of these operations is the applicable type conversion.

In fact, this usage actually looks a lot like it does in some other statically typed languages.

For example, in C/C++, you can say either `(int)x` or `int(x)`, and both will convert the value in `x` to an integer. Both forms are valid, but many prefer the latter, which kinda looks like a function call. In JavaScript, when you say `Number(x)`, it looks awfully similar. Does it matter that it's *actually* a function call in JS? Not really.

Besides `String(..)` and `Number(..)`, there are other ways to “explicitly” convert these values between `string` and `number`:

```
var a = 42;
var b = a.toString();

var c = "3.14";
var d = +c;

b; // "42"
d; // 3.14
```

Calling `a.toString()` is ostensibly explicit (pretty clear that “`toString`” means “to a string”), but there's some hidden implicitness here. `toString()` cannot be called on a *primitive* value like `42`. So JS automatically “boxes” (see Chapter 3) `42` in an object wrapper, so that `toString()` can be called against the object. In other words, you might call it “explicitly implicit.”

`+c` here is showing the *unary operator* form (operator with only one operand) of the `+` operator. Instead of performing mathematic addition (or string concat-

nation – see below), the unary `+` explicitly coerces its operand (`c`) to a `number` value.

Is `+c` *explicit* coercion? Depends on your experience and perspective. If you know (which you do, now!) that unary `+` is explicitly intended for `number` coercion, then it's pretty explicit and obvious. However, if you've never seen it before, it can seem awfully confusing, implicit, with hidden side effects, etc.

Note: The generally accepted perspective in the open-source JS community is that unary `+` is an accepted form of *explicit* coercion.

Even if you really like the `+c` form, there are definitely places where it can look awfully confusing. Consider:

```
var c = "3.14";
var d = 5+ +c;
```

```
d; // 8.14
```

The unary `-` operator also coerces like `+` does, but it also flips the sign of the number. However, you cannot put two `--` next to each other to unflip the sign, as that's parsed as the decrement operator. Instead, you would need to do: `- -"3.14"` with a space in between, and that would result in coercion to `3.14`.

You can probably dream up all sorts of hideous combinations of binary operators (like `+` for addition) next to the unary form of an operator. Here's another crazy example:

```
1 + - + + + - + 1; // 2
```

You should strongly consider avoiding unary `+` (or `-`) coercion when it's immediately adjacent to other operators. While the above works, it would almost universally be considered a bad idea. Even `d = +c` (or `d += c` for that matter!) can far too easily be confused for `d += c`, which is entirely different!

Note: Another extremely confusing place for unary `+` to be used adjacent to another operator would be the `++` increment operator and `--` decrement operator. For example: `a +++b`, `a + ++b`, and `a + + +b`. See “Expression Side-Effects” in Chapter 5 for more about `++`.

Remember, we're trying to be explicit and **reduce** confusion, not make it much worse!

Date To number Another common usage of the unary `+` operator is to coerce a `Date` object into a `number`, because the result is the unix timestamp (milliseconds elapsed since 1 January 1970 00:00:00 UTC) representation of the date/time value:

```
var d = new Date( "Mon, 18 Aug 2014 08:53:06 CDT" );
+d; // 1408369986000
```

The most common usage of this idiom is to get the current *now* moment as a timestamp, such as:

```
var timestamp = +new Date();
```

Note: Some developers are aware of a peculiar syntactic “trick” in JavaScript, which is that the () set on a constructor call (a function called with `new`) is *optional* if there are no arguments to pass. So you may run across the `var timestamp = +new Date;` form. However, not all developers agree that omitting the () improves readability, as it’s an uncommon syntax exception that only applies to the `new fn()` call form and not the regular `fn()` call form.

But coercion is not the only way to get the timestamp out of a `Date` object. A noncoercion approach is perhaps even preferable, as it’s even more explicit:

```
var timestamp = new Date().getTime();
// var timestamp = (new Date()).getTime();
// var timestamp = (new Date).getTime();
```

But an *even more* preferable noncoercion option is to use the ES5 added `Date.now()` static function:

```
var timestamp = Date.now();
```

And if you want to polyfill `Date.now()` into older browsers, it’s pretty simple:

```
if (!Date.now) {
  Date.now = function() {
    return +new Date();
  };
}
```

I’d recommend skipping the coercion forms related to dates. Use `Date.now()` for current *now* timestamps, and `new Date(...).getTime()` for getting a timestamp of a specific *non-now* date/time that you need to specify.

The Curious Case of the ~ One coercive JS operator that is often overlooked and usually very confused is the tilde ~ operator (aka “bitwise NOT”). Many of those who even understand what it does will often times still want to avoid it. But sticking to the spirit of our approach in this book and series, let’s dig into it to find out if ~ has anything useful to give us.

In the “32-bit (Signed) Integers” section of Chapter 2, we covered how bitwise operators in JS are defined only for 32-bit operations, which means they force their operands to conform to 32-bit value representations. The rules for how this happens are controlled by the `ToInt32` abstract operation (ES5 spec, section 9.5).

`ToInt32` first does a `ToNumber` coercion, which means if the value is "123", it’s going to first become 123 before the `ToInt32` rules are applied.

While not *technically* coercion itself (since the type doesn't change!), using bitwise operators (like `|` or `~`) with certain special `number` values produces a coercive effect that results in a different `number` value.

For example, let's first consider the `|` "bitwise OR" operator used in the otherwise no-op idiom `0 | x`, which (as Chapter 2 showed) essentially only does the `ToInt32` conversion:

```
0 | -0;          // 0
0 | NaN;         // 0
0 | Infinity;   // 0
0 | -Infinity;  // 0
```

These special numbers aren't 32-bit representable (since they come from the 64-bit IEEE 754 standard – see Chapter 2), so `ToInt32` just specifies 0 as the result from these values.

It's debatable if `0 | __` is an *explicit* form of this coercive `ToInt32` operation or if it's more *implicit*. From the spec perspective, it's unquestionably *explicit*, but if you don't understand bitwise operations at this level, it can seem a bit more *implicitly* magical. Nevertheless, consistent with other assertions in this chapter, we will call it *explicit*.

So, let's turn our attention back to `~`. The `~` operator first "coerces" to a 32-bit `number` value, and then performs a bitwise negation (flipping each bit's parity).

Note: This is very similar to how `!` not only coerces its value to `boolean` but also flips its parity (see discussion of the "unary `!`" later).

But... what!? Why do we care about bits being flipped? That's some pretty specialized, nuanced stuff. It's pretty rare for JS developers to need to reason about individual bits.

Another way of thinking about the definition of `~` comes from old-school computer science/discrete Mathematics: `~` performs two's-complement. Great, thanks, that's totally clearer!

Let's try again: `~x` is roughly the same as `-(x+1)`. That's weird, but slightly easier to reason about. So:

```
~42;    // -(42+1) ==> -43
```

You're probably still wondering what the heck all this `~` stuff is about, or why it really matters for a coercion discussion. Let's quickly get to the point.

Consider `-(x+1)`. What's the only value that you can perform that operation on that will produce a 0 (or `-0` technically!) result? `-1`. In other words, `~` used with a range of `number` values will produce a falsy (easily coercible to `false`) 0 value for the `-1` input value, and any other truthy `number` otherwise.

Why is that relevant?

`-1` is commonly called a “sentinel value,” which basically means a value that’s given an arbitrary semantic meaning within the greater set of values of its same type (`numbers`). The C-language uses `-1` sentinel values for many functions that return `>= 0` values for “success” and `-1` for “failure.”

JavaScript adopted this precedent when defining the `string` operation `indexOf(..)`, which searches for a substring and if found returns its zero-based index position, or `-1` if not found.

It’s pretty common to try to use `indexOf(..)` not just as an operation to get the position, but as a `boolean` check of presence/absence of a substring in another `string`. Here’s how developers usually perform such checks:

```
var a = "Hello World";

if (a.indexOf( "lo" ) >= 0) {    // true
    // found it!
}
if (a.indexOf( "lo" ) != -1) {    // true
    // found it
}

if (a.indexOf( "ol" ) < 0) {    // true
    // not found!
}
if (a.indexOf( "ol" ) == -1) {    // true
    // not found!
}
```

I find it kind of gross to look at `>= 0` or `== -1`. It’s basically a “leaky abstraction,” in that it’s leaking underlying implementation behavior – the usage of sentinel `-1` for “failure” – into my code. I would prefer to hide such a detail.

And now, finally, we see why `~` could help us! Using `~` with `indexOf()` “coerces” (actually just transforms) the value **to be appropriately boolean-coercible**:

```
var a = "Hello World";

~a.indexOf( "lo" );           // -4    <-- truthy!

if (~a.indexOf( "lo" )) {    // true
    // found it!
}

~a.indexOf( "ol" );           // 0    <-- falsy!
!~a.indexOf( "ol" );         // true

if (!~a.indexOf( "ol" )) {    // true
```

```
// not found!
}
```

`~` takes the return value of `indexOf(..)` and transforms it: for the “failure” `-1` we get the falsy `0`, and every other value is truthy.

Note: The `-(x+1)` pseudo-algorithm for `~` would imply that `~-1` is `-0`, but actually it produces `0` because the underlying operation is actually bitwise, not mathematic.

Technically, `if (~a.indexOf(..))` is still relying on *implicit* coercion of its resultant `0` to `false` or nonzero to `true`. But overall, `~` still feels to me more like an *explicit* coercion mechanism, as long as you know what it’s intended to do in this idiom.

I find this to be cleaner code than the previous `>= 0 / == -1` clutter.

Truncating Bits There’s one more place `~` may show up in code you run across: some developers use the double tilde `~~` to truncate the decimal part of a `number` (i.e., “coerce” it to a whole number “integer”). It’s commonly (though mistakenly) said this is the same result as calling `Math.floor(..)`.

How `~~` works is that the first `~` applies the `ToInt32` “coercion” and does the bitwise flip, and then the second `~` does another bitwise flip, flipping all the bits back to the original state. The end result is just the `ToInt32` “coercion” (aka truncation).

Note: The bitwise double-flip of `~~` is very similar to the parity double-negate `!!` behavior, explained in the “Explicitly: `* -> Boolean`” section later.

However, `~~` needs some caution/clarification. First, it only works reliably on 32-bit values. But more importantly, it doesn’t work the same on negative numbers as `Math.floor(..)` does!

```
Math.floor( -49.6 );      // -50
~~-49.6;                  // -49
```

Setting the `Math.floor(..)` difference aside, `~~x` can truncate to a (32-bit) integer. But so does `x | 0`, and seemingly with (slightly) *less effort*.

So, why might you choose `~~x` over `x | 0`, then? Operator precedence (see Chapter 5):

```
~~1E20 / 10;             // 166199296
1E20 | 0 / 10;           // 1661992960
(1E20 | 0) / 10;         // 166199296
```

Just as with all other advice here, use `~` and `~~` as explicit mechanisms for “coercion” and value transformation only if everyone who reads/writes such code is properly aware of how these operators work!

Explicitly: Parsing Numeric Strings

A similar outcome to coercing a `string` to a `number` can be achieved by parsing a `number` out of a `string`'s character contents. There are, however, distinct differences between this parsing and the type conversion we examined above.

Consider:

```
var a = "42";
var b = "42px";

Number( a );    // 42
parseInt( a ); // 42

Number( b );    // NaN
parseInt( b ); // 42
```

Parsing a numeric value out of a string is *tolerant* of non-numeric characters – it just stops parsing left-to-right when encountered – whereas coercion is *not tolerant* and fails resulting in the `NaN` value.

Parsing should not be seen as a substitute for coercion. These two tasks, while similar, have different purposes. Parse a `string` as a `number` when you don't know/care what other non-numeric characters there may be on the right-hand side. Coerce a `string` (to a `number`) when the only acceptable values are numeric and something like "42px" should be rejected as a `number`.

Tip: `parseInt(..)` has a twin, `parseFloat(..)`, which (as it sounds) pulls out a floating-point number from a string.

Don't forget that `parseInt(..)` operates on `string` values. It makes absolutely no sense to pass a `number` value to `parseInt(..)`. Nor would it make sense to pass any other type of value, like `true`, `function(){..}` or `[1,2,3]`.

If you pass a non-`string`, the value you pass will automatically be coerced to a `string` first (see "ToString" earlier), which would clearly be a kind of hidden *implicit* coercion. It's a really bad idea to rely upon such a behavior in your program, so never use `parseInt(..)` with a non-`string` value.

Prior to ES5, another gotcha existed with `parseInt(..)`, which was the source of many JS programs' bugs. If you didn't pass a second argument to indicate which numeric base (aka radix) to use for interpreting the numeric `string` contents, `parseInt(..)` would look at the beginning character(s) to make a guess.

If the first two characters were "0x" or "0X", the guess (by convention) was that you wanted to interpret the `string` as a hexadecimal (base-16) `number`. Otherwise, if the first character was "0", the guess (again, by convention) was that you wanted to interpret the `string` as an octal (base-8) `number`.

Hexadecimal `strings` (with the leading `0x` or `0X`) aren't terribly easy to get

mixed up. But the octal number guessing proved devilishly common. For example:

```
var hour = parseInt( selectedHour.value );
var minute = parseInt( selectedMinute.value );

console.log( "The time you selected was: " + hour + ":" + minute);
```

Seems harmless, right? Try selecting 08 for the hour and 09 for the minute. You'll get 0:0. Why? because neither 8 nor 9 are valid characters in octal base-8.

The pre-ES5 fix was simple, but so easy to forget: **always pass 10 as the second argument**. This was totally safe:

```
var hour = parseInt( selectedHour.value, 10 );
var minute = parseInt( selectedMinute.value, 10 );
```

As of ES5, `parseInt(..)` no longer guesses octal. Unless you say otherwise, it assumes base-10 (or base-16 for "0x" prefixes). That's much nicer. Just be careful if your code has to run in pre-ES5 environments, in which case you still need to pass 10 for the radix.

Parsing Non-Strings One somewhat infamous example of `parseInt(..)`'s behavior is highlighted in a sarcastic joke post a few years ago, poking fun at this JS behavior:

```
parseInt( 1/0, 19 ); // 18
```

The assumptive (but totally invalid) assertion was, "If I pass in `Infinity`, and parse an integer out of that, I should get `Infinity` back, not 18." Surely, JS must be crazy for this outcome, right?

Though this example is obviously contrived and unreal, let's indulge the madness for a moment and examine whether JS really is that crazy.

First off, the most obvious sin committed here is to pass a **non-string** to `parseInt(..)`. That's a no-no. Do it and you're asking for trouble. But even if you do, JS politely coerces what you pass in into a **string** that it can try to parse.

Some would argue that this is unreasonable behavior, and that `parseInt(..)` should refuse to operate on a **non-string** value. Should it perhaps throw an error? That would be very Java-like, frankly. I shudder at thinking JS should start throwing errors all over the place so that `try..catch` is needed around almost every line.

Should it return `NaN`? Maybe. But... what about:

```
parseInt( new String( "42" ) );
```

Should that fail, too? It's a non-**string** value. If you want that **String** object wrapper to be unboxed to "42", then is it really so unusual for 42 to first become "42" so that 42 can be parsed back out?

I would argue that this half-*explicit*, half-*implicit* coercion that can occur can often be a very helpful thing. For example:

```
var a = {  
  num: 21,  
  toString: function() { return String( this.num * 2 ); }  
};  
  
parseInt( a ); // 42
```

The fact that **parseInt**(...) forcibly coerces its value to a **string** to perform the parse on is quite sensible. If you pass in garbage, and you get garbage back out, don't blame the trash can – it just did its job faithfully.

So, if you pass in a value like **Infinity** (the result of 1 / 0 obviously), what sort of **string** representation would make the most sense for its coercion? Only two reasonable choices come to mind: "Infinity" and "∞". JS chose "Infinity". I'm glad it did.

I think it's a good thing that **all values** in JS have some sort of default **string** representation, so that they aren't mysterious black boxes that we can't debug and reason about.

Now, what about base-19? Obviously, completely bogus and contrived. No real JS programs use base-19. It's absurd. But again, let's indulge the ridiculousness. In base-19, the valid numeric characters are 0 - 9 and a - i (case insensitive).

So, back to our **parseInt(1/0, 19)** example. It's essentially **parseInt("Infinity", 19)**. How does it parse? The first character is "I", which is value 18 in the silly base-19. The second character "n" is not in the valid set of numeric characters, and as such the parsing simply politely stops, just like when it ran across "p" in "42px".

The result? 18. Exactly like it sensibly should be. The behaviors involved to get us there, and not to an error or to **Infinity** itself, are **very important** to JS, and should not be so easily discarded.

Other examples of this behavior with **parseInt(...)** that may be surprising but are quite sensible include:

```
parseInt( 0.000008 );      // 0  ("0" from "0.000008")  
parseInt( 0.000008 );      // 8  ("8" from "8e-7")  
parseInt( false, 16 );      // 250  ("fa" from "false")  
parseInt( parseInt, 16 );    // 15  ("f" from "function..")  
  
parseInt( "0x10" );        // 16  
parseInt( "103", 2 );       // 2
```

`parseInt(..)` is actually pretty predictable and consistent in its behavior. If you use it correctly, you'll get sensible results. If you use it incorrectly, the crazy results you get are not the fault of JavaScript.

Explicitly: `* -> Boolean`

Now, let's examine coercing from any non-`boolean` value to a `boolean`.

Just like with `String(..)` and `Number(..)` above, `Boolean(..)` (without the `new`, of course!) is an explicit way of forcing the `ToBoolean` coercion:

```
var a = "0";
var b = [];
var c = {};  
  
var d = "";
var e = 0;
var f = null;
var g;  
  
Boolean( a ); // true
Boolean( b ); // true
Boolean( c ); // true  
  
Boolean( d ); // false
Boolean( e ); // false
Boolean( f ); // false
Boolean( g ); // false
```

While `Boolean(..)` is clearly explicit, it's not at all common or idiomatic.

Just like the unary `+` operator coerces a value to a `number` (see above), the unary `!` negate operator explicitly coerces a value to a `boolean`. The *problem* is that it also flips the value from truthy to falsy or vice versa. So, the most common way JS developers explicitly coerce to `boolean` is to use the `!!` double-negate operator, because the second `!` will flip the parity back to the original:

```
var a = "0";
var b = [];
var c = {};  
  
var d = "";
var e = 0;
var f = null;
var g;  
  
!!a; // true
!!b; // true
```

```
!!c;      // true
!!d;      // false
!!e;      // false
!!f;      // false
!!g;      // false
```

Any of these `ToBoolean` coercions would happen *implicitly* without the `Boolean(..)` or `!!`, if used in a `boolean` context such as an `if (..) ..` statement. But the goal here is to explicitly force the value to a `boolean` to make it clearer that the `ToBoolean` coercion is intended.

Another example use-case for explicit `ToBoolean` coercion is if you want to force a `true/false` value coercion in the JSON serialization of a data structure:

```
var a = [
  1,
  function(){ /*...*/ },
  2,
  function(){ /*...*/ }
];

JSON.stringify( a ); // "[1,null,2,null]"

JSON.stringify( a, function(key,val){
  if (typeof val == "function") {
    // force `ToBoolean` coercion of the function
    return !!val;
  }
  else {
    return val;
  }
} );
// "[1,true,2,true]"
```

If you come to JavaScript from Java, you may recognize this idiom:

```
var a = 42;

var b = a ? true : false;
```

The `? :` ternary operator will test `a` for truthiness, and based on that test will either assign `true` or `false` to `b`, accordingly.

On its surface, this idiom looks like a form of *explicit* `ToBoolean`-type coercion, since it's obvious that only either `true` or `false` come out of the operation.

However, there's a hidden *implicit* coercion, in that the `a` expression has to first be coerced to `boolean` to perform the truthiness test. I'd call this idiom "explicitly implicit." Furthermore, I'd suggest **you should avoid this idiom**

completely in JavaScript. It offers no real benefit, and worse, masquerades as something it's not.

`Boolean(a)` and `!!a` are far better as *explicit* coercion options.

Implicit Coercion

Implicit coercion refers to type conversions that are hidden, with non-obvious side-effects that implicitly occur from other actions. In other words, *implicit coercions* are any type conversions that aren't obvious (to you).

While it's clear what the goal of *explicit* coercion is (making code explicit and more understandable), it might be *too* obvious that *implicit* coercion has the opposite goal: making code harder to understand.

Taken at face value, I believe that's where much of the ire towards coercion comes from. The majority of complaints about "JavaScript coercion" are actually aimed (whether they realize it or not) at *implicit* coercion.

Note: Douglas Crockford, author of "*JavaScript: The Good Parts*", has claimed in many conference talks and writings that JavaScript coercion should be avoided. But what he seems to mean is that *implicit* coercion is bad (in his opinion). However, if you read his own code, you'll find plenty of examples of coercion, both *implicit* and *explicit*! In truth, his angst seems to primarily be directed at the `==` operation, but as you'll see in this chapter, that's only part of the coercion mechanism.

So, is **implicit coercion** evil? Is it dangerous? Is it a flaw in JavaScript's design? Should we avoid it at all costs?

I bet most of you readers are inclined to enthusiastically cheer, "Yes!"

Not so fast. Hear me out.

Let's take a different perspective on what *implicit* coercion is, and can be, than just that it's "the opposite of the good explicit kind of coercion." That's far too narrow and misses an important nuance.

Let's define the goal of *implicit* coercion as: to reduce verbosity, boilerplate, and/or unnecessary implementation detail that clutters up our code with noise that distracts from the more important intent.

Simplifying Implicitly

Before we even get to JavaScript, let me suggest something pseudo-code-ish from some theoretical strongly typed language to illustrate:

```
SomeType x = SomeType( AnotherType( y ) )
```

In this example, I have some arbitrary type of value in `y` that I want to convert to the `SomeType` type. The problem is, this language can't go directly from

whatever `y` currently is to `SomeType`. It needs an intermediate step, where it first converts to `AnotherType`, and then from `AnotherType` to `SomeType`.

Now, what if that language (or definition you could create yourself with the language) *did* just let you say:

```
SomeType x = SomeType( y )
```

Wouldn't you generally agree that we simplified the type conversion here to reduce the unnecessary "noise" of the intermediate conversion step? I mean, is it *really* all that important, right here at this point in the code, to see and deal with the fact that `y` goes to `AnotherType` first before then going to `SomeType`?

Some would argue, at least in some circumstances, yes. But I think an equal argument can be made of many other circumstances that here, the simplification **actually aids in the readability of the code** by abstracting or hiding away such details, either in the language itself or in our own abstractions.

Undoubtedly, behind the scenes, somewhere, the intermediate conversion step is still happening. But if that detail is hidden from view here, we can just reason about getting `y` to type `SomeType` as a generic operation and hide the messy details.

While not a perfect analogy, what I'm going to argue throughout the rest of this chapter is that JS *implicit* coercion can be thought of as providing a similar aid to your code.

But, **and this is very important**, that is not an unbounded, absolute statement. There are definitely plenty of *evils* lurking around *implicit* coercion, that will harm your code much more than any potential readability improvements. Clearly, we have to learn how to avoid such constructs so we don't poison our code with all manner of bugs.

Many developers believe that if a mechanism can do some useful thing **A** but can also be abused or misused to do some awful thing **Z**, then we should throw out that mechanism altogether, just to be safe.

My encouragement to you is: don't settle for that. Don't "throw the baby out with the bathwater." Don't assume *implicit* coercion is all bad because all you think you've ever seen is its "bad parts." I think there are "good parts" here, and I want to help and inspire more of you to find and embrace them!

Implicitly: Strings <-> Numbers

Earlier in this chapter, we explored *explicitly* coercing between `string` and `number` values. Now, let's explore the same task but with *implicit* coercion approaches. But before we do, we have to examine some nuances of operations that will *implicitly* force coercion.

The `+` operator is overloaded to serve the purposes of both `number` addition and `string` concatenation. So how does JS know which type of operation you want

to use? Consider:

```
var a = "42";
var b = "0";

var c = 42;
var d = 0;

a + b; // "420"
c + d; // 42
```

What's different that causes "420" vs 42? It's a common misconception that the difference is whether one or both of the operands is a **string**, as that means **+** will assume **string** concatenation. While that's partially true, it's more complicated than that.

Consider:

```
var a = [1,2];
var b = [3,4];

a + b; // "1,23,4"
```

Neither of these operands is a **string**, but clearly they were both coerced to **strings** and then the **string** concatenation kicked in. So what's really going on?

(**Warning:** deeply nitty gritty spec-speak coming, so skip the next two paragraphs if that intimidates you!)

According to ES5 spec section 11.6.1, the **+** algorithm (when an **object** value is an operand) will concatenate if either operand is either already a **string**, or if the following steps produce a **string** representation. So, when **+** receives an **object** (including **array**) for either operand, it first calls the **ToPrimitive** abstract operation (section 9.1) on the value, which then calls the **[[DefaultValue]]** algorithm (section 8.12.8) with a context hint of **number**.

If you're paying close attention, you'll notice that this operation is now identical to how the **ToNumber** abstract operation handles **objects** (see the "**ToNumber**" section earlier). The **valueOf()** operation on the **array** will fail to produce a simple primitive, so it then falls to a **toString()** representation. The two **arrays** thus become "1,2" and "3,4", respectively. Now, **+** concatenates the two **strings** as you'd normally expect: "1,23,4".

Let's set aside those messy details and go back to an earlier, simplified explanation: if either operand to **+** is a **string** (or becomes one with the above steps!),

the operation will be `string` concatenation. Otherwise, it's always numeric addition.

Note: A commonly cited coercion gotcha is `[] + {}` vs. `{} + []`, as those two expressions result, respectively, in `"[object Object]"` and `0`. There's more to it, though, and we cover those details in "Blocks" in Chapter 5.

What's that mean for *implicit* coercion?

You can coerce a `number` to a `string` simply by "adding" the `number` and the `""` empty `string`:

```
var a = 42;
var b = a + "";

b; // "42"
```

Tip: Numeric addition with the `+` operator is commutative, which means `2 + 3` is the same as `3 + 2`. String concatenation with `+` is obviously not generally commutative, **but** with the specific case of `""`, it's effectively commutative, as `a + ""` and `"" + a` will produce the same result.

It's extremely common/idiomatic to (*implicitly*) coerce `number` to `string` with `a + ""` operation. In fact, interestingly, even some of the most vocal critics of *implicit* coercion still use that approach in their own code, instead of one of its *explicit* alternatives.

I think this is a great example of a useful form in *implicit* coercion, despite how frequently the mechanism gets criticized!

Comparing this *implicit* coercion of `a + ""` to our earlier example of `String(a)` *explicit* coercion, there's one additional quirk to be aware of. Because of how the `ToPrimitive` abstract operation works, `a + ""` invokes `valueOf()` on the `a` value, whose return value is then finally converted to a `string` via the internal `ToString` abstract operation. But `String(a)` just invokes `toString()` directly.

Both approaches ultimately result in a `string`, but if you're using an `object` instead of a regular primitive `number` value, you may not necessarily get the *same* `string` value!

Consider:

```
var a = {
  valueOf: function() { return 42; },
  toString: function() { return 4; }
};

a + ""; // "42"

String(a); // "4"
```

Generally, this sort of gotcha won't bite you unless you're really trying to create confusing data structures and operations, but you should be careful if you're defining both your own `valueOf()` and `toString()` methods for some `object`, as how you coerce the value could affect the outcome.

What about the other direction? How can we *implicitly coerce* from `string` to `number`?

```
var a = "3.14";
var b = a - 0;

b; // 3.14
```

The `-` operator is defined only for numeric subtraction, so `a - 0` forces `a`'s value to be coerced to a `number`. While far less common, `a * 1` or `a / 1` would accomplish the same result, as those operators are also only defined for numeric operations.

What about `object` values with the `-` operator? Similar story as for `+` above:

```
var a = [3];
var b = [1];

a - b; // 2
```

Both `array` values have to become `numbers`, but they end up first being coerced to `strings` (using the expected `toString()` serialization), and then are coerced to `numbers`, for the `-` subtraction to perform on.

So, is *implicit* coercion of `string` and `number` values the ugly evil you've always heard horror stories about? I don't personally think so.

Compare `b = String(a)` (*explicit*) to `b = a + ""` (*implicit*). I think cases can be made for both approaches being useful in your code. Certainly `b = a + ""` is quite a bit more common in JS programs, proving its own utility regardless of *feelings* about the merits or hazards of *implicit* coercion in general.

Implicitly: Booleans → Numbers

I think a case where *implicit* coercion can really shine is in simplifying certain types of complicated `boolean` logic into simple numeric addition. Of course, this is not a general-purpose technique, but a specific solution for specific cases.

Consider:

```
function onlyOne(a,b,c) {
  return (!((a && !b && !c) ||
            (!a && b && !c) || (!a && !b && c));
}

var a = true;
```

```

var b = false;

onlyOne( a, b, b ); // true
onlyOne( b, a, b ); // true

onlyOne( a, b, a ); // false

```

This `onlyOne(..)` utility should only return `true` if exactly one of the arguments is `true` / truthy. It's using *implicit* coercion on the truthy checks and *explicit* coercion on the others, including the final return value.

But what if we needed that utility to be able to handle four, five, or twenty flags in the same way? It's pretty difficult to imagine implementing code that would handle all those permutations of comparisons.

But here's where coercing the `boolean` values to `numbers` (0 or 1, obviously) can greatly help:

```

function onlyOne() {
  var sum = 0;
  for (var i=0; i < arguments.length; i++) {
    // skip falsy values. same as treating
    // them as 0's, but avoids NaN's.
    if (arguments[i]) {
      sum += arguments[i];
    }
  }
  return sum == 1;
}

var a = true;
var b = false;

onlyOne( b, a ); // true
onlyOne( b, a, b, b, b ); // true

onlyOne( b, b ); // false
onlyOne( b, a, b, b, b, a ); // false

```

Note: Of course, instead of the `for` loop in `onlyOne(..)`, you could more tersely use the ES5 `reduce(..)` utility, but I didn't want to obscure the concepts.

What we're doing here is relying on the 1 for `true`/truthy coercions, and numerically adding them all up. `sum += arguments[i]` uses *implicit* coercion to make that happen. If one and only one value in the `arguments` list is `true`, then the numeric sum will be 1, otherwise the sum will not be 1 and thus the desired condition is not met.

We could of course do this with *explicit* coercion instead:

```

function onlyOne() {
  var sum = 0;
  for (var i=0; i < arguments.length; i++) {
    sum += Number( !!arguments[i] );
  }
  return sum === 1;
}

```

We first use `!!arguments[i]` to force the coercion of the value to `true` or `false`. That's so you could pass non-`boolean` values in, like `onlyOne("42", 0)`, and it would still work as expected (otherwise you'd end up with `string` concatenation and the logic would be incorrect).

Once we're sure it's a `boolean`, we do another *explicit* coercion with `Number(..)` to make sure the value is 0 or 1.

Is the *explicit* coercion form of this utility “better”? It does avoid the `NaN` trap as explained in the code comments. But, ultimately, it depends on your needs. I personally think the former version, relying on *implicit* coercion is more elegant (if you won't be passing `undefined` or `NaN`), and the *explicit* version is needlessly more verbose.

But as with almost everything we're discussing here, it's a judgment call.

Note: Regardless of *implicit* or *explicit* approaches, you could easily make `onlyTwo(..)` or `onlyFive(..)` variations by simply changing the final comparison from 1, to 2 or 5, respectively. That's drastically easier than adding a bunch of `&&` and `||` expressions. So, generally, coercion is very helpful in this case.

Implicitly: * → Boolean

Now, let's turn our attention to *implicit* coercion to `boolean` values, as it's by far the most common and also by far the most potentially troublesome.

Remember, *implicit* coercion is what kicks in when you use a value in such a way that it forces the value to be converted. For numeric and `string` operations, it's fairly easy to see how the coercions can occur.

But, what sort of expression operations require/force (*implicitly*) a `boolean` coercion?

1. The test expression in an `if (..)` statement.
2. The test expression (second clause) in a `for (.. ; .. ; ..)` header.
3. The test expression in `while (..)` and `do..while(..)` loops.
4. The test expression (first clause) in `? : ternary expressions`.
5. The left-hand operand (which serves as a test expression – see below!) to the `||` (“logical or”) and `&&` (“logical and”) operators.

Any value used in these contexts that is not already a `boolean` will be *implicitly* coerced to a `boolean` using the rules of the `ToBoolean` abstract operation covered earlier in this chapter.

Let's look at some examples:

```
var a = 42;
var b = "abc";
var c;
var d = null;

if (a) {
    console.log( "yep" );           // yep
}

while (c) {
    console.log( "nope, never runs" );
}

c = d ? a : b;                  // "abc"

if ((a && d) || c) {
    console.log( "yep" );           // yep
}
```

In all these contexts, the non-`boolean` values are *implicitly coerced* to their `boolean` equivalents to make the test decisions.

Operators `||` and `&&`

It's quite likely that you have seen the `||` ("logical or") and `&&` ("logical and") operators in most or all other languages you've used. So it'd be natural to assume that they work basically the same in JavaScript as in other similar languages.

There's some very little known, but very important, nuance here.

In fact, I would argue these operators shouldn't even be called "logical _____ operators", as that name is incomplete in describing what they do. If I were to give them a more accurate (if more clumsy) name, I'd call them "selector operators," or more completely, "operand selector operators."

Why? Because they don't actually result in a *logic* value (aka `boolean`) in JavaScript, as they do in some other languages.

So what *do* they result in? They result in the value of one (and only one) of their two operands. In other words, **they select one of the two operand's values**.

Quoting the ES5 spec from section 11.11:

The value produced by a `&&` or `||` operator is not necessarily of type Boolean. The value produced will always be the value of one of the two operand expressions.

Let's illustrate:

```
var a = 42;
var b = "abc";
var c = null;

a || b;      // 42
a && b;      // "abc"

c || b;      // "abc"
c && b;      // null
```

Wait, what!? Think about that. In languages like C and PHP, those expressions result in `true` or `false`, but in JS (and Python and Ruby, for that matter!), the result comes from the values themselves.

Both `||` and `&&` operators perform a `boolean` test on the **first operand** (`a` or `c`). If the operand is not already `boolean` (as it's not, here), a normal `ToBoolean` coercion occurs, so that the test can be performed.

For the `||` operator, if the test is `true`, the `||` expression results in the value of the *first operand* (`a` or `c`). If the test is `false`, the `||` expression results in the value of the *second operand* (`b`).

Inversely, for the `&&` operator, if the test is `true`, the `&&` expression results in the value of the *second operand* (`b`). If the test is `false`, the `&&` expression results in the value of the *first operand* (`a` or `c`).

The result of a `||` or `&&` expression is always the underlying value of one of the operands, **not** the (possibly coerced) result of the test. In `c && b`, `c` is `null`, and thus falsy. But the `&&` expression itself results in `null` (the value in `c`), not in the coerced `false` used in the test.

Do you see how these operators act as “operand selectors”, now?

Another way of thinking about these operators:

```
a || b;
// roughly equivalent to:
a ? a : b;

a && b;
// roughly equivalent to:
a ? b : a;
```

Note: I call `a || b` “roughly equivalent” to `a ? a : b` because the outcome is identical, but there’s a nuanced difference. In `a ? a : b`, if `a` was a more complex expression (like for instance one that might have side effects like calling a `function`, etc.), then the `a` expression would possibly be evaluated twice (if the first evaluation was truthy). By contrast, for `a || b`, the `a` expression is evaluated only once, and that value is used both for the coercive test as well as the result value (if appropriate). The same nuance applies to the `a && b` and `a ? b : a` expressions.

An extremely common and helpful usage of this behavior, which there’s a good chance you may have used before and not fully understood, is:

```
function foo(a,b) {  
  a = a || "hello";  
  b = b || "world";  
  
  console.log( a + " " + b );  
}  
  
foo();           // "hello world"  
foo( "yeah", "yeah!" ); // "yeah yeah!"
```

The `a = a || "hello"` idiom (sometimes said to be JavaScript’s version of the C# “null coalescing operator”) acts to test `a` and if it has no value (or only an undesired falsy value), provides a backup default value (“`hello`”).

Be careful, though!

```
foo( "That's it!", "" ); // "That's it! world" <-- Oops!
```

See the problem? `""` as the second argument is a falsy value (see `ToBoolean` earlier in this chapter), so the `b = b || "world"` test fails, and the “`world`” default value is substituted, even though the intent probably was to have the explicitly passed `""` be the value assigned to `b`.

This `||` idiom is extremely common, and quite helpful, but you have to use it only in cases where *all falsy values* should be skipped. Otherwise, you’ll need to be more explicit in your test, and probably use `a ? : ternary` instead.

This *default value assignment* idiom is so common (and useful!) that even those who publicly and vehemently decry JavaScript coercion often use it in their own code!

What about `&&`?

There’s another idiom that is quite a bit less commonly authored manually, but which is used by JS minifiers frequently. The `&&` operator “selects” the second operand if and only if the first operand tests as truthy, and this usage is sometimes called the “guard operator” (also see “Short Circuited” in Chapter 5) – the first expression test “guards” the second expression:

```

function foo() {
  console.log( a );
}

var a = 42;

a && foo(); // 42

```

`foo()` gets called only because `a` tests as truthy. If that test failed, this `a && foo()` expression statement would just silently stop – this is known as “short circuiting” – and never call `foo()`.

Again, it’s not nearly as common for people to author such things. Usually, they’d do `if (a) { foo(); }` instead. But JS minifiers choose `a && foo()` because it’s much shorter. So, now, if you ever have to decipher such code, you’ll know what it’s doing and why.

OK, so `||` and `&&` have some neat tricks up their sleeve, as long as you’re willing to allow the *implicit* coercion into the mix.

Note: Both the `a = b || "something"` and `a && b()` idioms rely on short circuiting behavior, which we cover in more detail in Chapter 5.

The fact that these operators don’t actually result in `true` and `false` is possibly messing with your head a little bit by now. You’re probably wondering how all your `if` statements and `for` loops have been working, if they’ve included compound logical expressions like `a && (b || c)`.

Don’t worry! The sky is not falling. Your code is (probably) just fine. It’s just that you probably never realized before that there was an *implicit* coercion to `boolean` going on **after** the compound expression was evaluated.

Consider:

```

var a = 42;
var b = null;
var c = "foo";

if (a && (b || c)) {
  console.log( "yep" );
}

```

This code still works the way you always thought it did, except for one subtle extra detail. The `a && (b || c)` expression *actually* results in `"foo"`, not `true`. So, the `if` statement *then* forces the `"foo"` value to coerce to a `boolean`, which of course will be `true`.

See? No reason to panic. Your code is probably still safe. But now you know more about how it does what it does.

And now you also realize that such code is using *implicit* coercion. If you’re in

the “avoid (implicit) coercion camp” still, you’re going to need to go back and make all of those tests *explicit*:

```
if (!a && (!b || !c)) {  
  console.log( "yep" );  
}
```

Good luck with that! ... Sorry, just teasing.

Symbol Coercion

Up to this point, there’s been almost no observable outcome difference between *explicit* and *implicit* coercion – only the readability of code has been at stake.

But ES6 Symbols introduce a gotcha into the coercion system that we need to discuss briefly. For reasons that go well beyond the scope of what we’ll discuss in this book, *explicit* coercion of a `symbol` to a `string` is allowed, but *implicit* coercion of the same is disallowed and throws an error.

Consider:

```
var s1 = Symbol( "cool" );  
String( s1 ); // "Symbol(cool)"  
  
var s2 = Symbol( "not cool" );  
s2 + ""; // TypeError
```

`symbol` values cannot coerce to `number` at all (throws an error either way), but strangely they can both *explicitly* and *implicitly* coerce to `boolean` (always `true`).

Consistency is always easier to learn, and exceptions are never fun to deal with, but we just need to be careful around the new ES6 `symbol` values and how we coerce them.

The good news: it’s probably going to be exceedingly rare for you to need to coerce a `symbol` value. The way they’re typically used (see Chapter 3) will probably not call for coercion on a normal basis.

Loose Equals vs. Strict Equals

Loose equals is the `==` operator, and strict equals is the `===` operator. Both operators are used for comparing two values for “equality,” but the “loose” vs. “strict” indicates a **very important** difference in behavior between the two, specifically in how they decide “equality.”

A very common misconception about these two operators is: “`==` checks values for equality and `===` checks both values and types for equality.” While that sounds nice and reasonable, it’s inaccurate. Countless well-respected JavaScript books and blogs have said exactly that, but unfortunately they’re all *wrong*.

The correct description is: “`==` allows coercion in the equality comparison and `===` disallows coercion.”

Equality Performance

Stop and think about the difference between the first (inaccurate) explanation and this second (accurate) one.

In the first explanation, it seems obvious that `===` is *doing more work* than `==`, because it has to *also* check the type. In the second explanation, `==` is the one *doing more work* because it has to follow through the steps of coercion if the types are different.

Don’t fall into the trap, as many have, of thinking this has anything to do with performance, though, as if `==` is going to be slower than `===` in any relevant way. While it’s measurable that coercion does take *a little bit* of processing time, it’s mere microseconds (yes, that’s millionths of a second!).

If you’re comparing two values of the same types, `==` and `===` use the identical algorithm, and so other than minor differences in engine implementation, they should do the same work.

If you’re comparing two values of different types, the performance isn’t the important factor. What you should be asking yourself is: when comparing these two values, do I want coercion or not?

If you want coercion, use `==` loose equality, but if you don’t want coercion, use `===` strict equality.

Note: The implication here then is that both `==` and `===` check the types of their operands. The difference is in how they respond if the types don’t match.

Abstract Equality

The `==` operator’s behavior is defined as “The Abstract Equality Comparison Algorithm” in section 11.9.3 of the ES5 spec. What’s listed there is a comprehensive but simple algorithm that explicitly states every possible combination of types, and how the coercions (if necessary) should happen for each combination.

Warning: When (*implicit*) coercion is maligned as being too complicated and too flawed to be a *useful good part*, it is these rules of “abstract equality” that are being condemned. Generally, they are said to be too complex and too unintuitive for developers to practically learn and use, and that they are prone more to causing bugs in JS programs than to enabling greater code readability. I believe this is a flawed premise – that you readers are competent developers who write (and read and understand!) algorithms (aka code) all day long. So, what follows is a plain exposition of the “abstract equality” in simple terms. But I implore you to also read the ES5 spec section 11.9.3. I think you’ll be surprised at just how reasonable it is.

Basically, the first clause (11.9.3.1) says, if the two values being compared are of the same type, they are simply and naturally compared via Identity as you'd expect. For example, 42 is only equal to 42, and "abc" is only equal to "abc".

Some minor exceptions to normal expectation to be aware of:

- `NaN` is never equal to itself (see Chapter 2)
- `+0` and `-0` are equal to each other (see Chapter 2)

The final provision in clause 11.9.3.1 is for `==` loose equality comparison with `objects` (including `functions` and `arrays`). Two such values are only *equal* if they are both references to *the exact same value*. No coercion occurs here.

Note: The `==` strict equality comparison is defined identically to 11.9.3.1, including the provision about two `object` values. It's a very little known fact that `==` and `===` behave identically in the case where two `objects` are being compared!

The rest of the algorithm in 11.9.3 specifies that if you use `==` loose equality to compare two values of different types, one or both of the values will need to be *implicitly* coerced. This coercion happens so that both values eventually end up as the same type, which can then directly be compared for equality using simple value Identity.

Note: The `!=` loose not-equality operation is defined exactly as you'd expect, in that it's literally the `==` operation comparison performed in its entirety, then the negation of the result. The same goes for the `!==` strict not-equality operation.

Comparing: strings to numbers To illustrate `==` coercion, let's first build off the `string` and `number` examples earlier in this chapter:

```
var a = 42;
var b = "42";

a === b;    // false
a == b;     // true
```

As we'd expect, `a === b` fails, because no coercion is allowed, and indeed the `42` and `"42"` values are different.

However, the second comparison `a == b` uses loose equality, which means that if the types happen to be different, the comparison algorithm will perform *implicit* coercion on one or both values.

But exactly what kind of coercion happens here? Does the `a` value of `42` become a `string`, or does the `b` value of `"42"` become a `number`?

In the ES5 spec, clauses 11.9.3.4-5 say:

4. If `Type(x)` is `Number` and `Type(y)` is `String`, return the result of the comparison `x == ToNumber(y)`.

5. If `Type(x)` is `String` and `Type(y)` is `Number`, return the result of the comparison `ToNumber(x) == y`.

Warning: The spec uses `Number` and `String` as the formal names for the types, while this book prefers `number` and `string` for the primitive types. Do not let the capitalization of `Number` in the spec confuse you for the `Number()` native function. For our purposes, the capitalization of the type name is irrelevant – they have basically the same meaning.

Clearly, the spec says the "42" value is coerced to a `number` for the comparison. The *how* of that coercion has already been covered earlier, specifically with the `ToNumber` abstract operation. In this case, it's quite obvious then that the resulting two `42` values are equal.

Comparing: anything to boolean One of the biggest gotchas with the *implicit* coercion of `==` loose equality pops up when you try to compare a value directly to `true` or `false`.

Consider:

```
var a = "42";
var b = true;

a == b; // false
```

Wait, what happened here!? We know that "42" is a *truthy* value (see earlier in this chapter). So, how come it's not `==` loose equal to `true`?

The reason is both simple and deceptively tricky. It's so easy to misunderstand, many JS developers never pay close enough attention to fully grasp it.

Let's again quote the spec, clauses 11.9.3.6-7:

6. If `Type(x)` is `Boolean`, return the result of the comparison `ToNumber(x) == y`.
7. If `Type(y)` is `Boolean`, return the result of the comparison `x == ToNumber(y)`.

Let's break that down. First:

```
var x = true;
var y = "42";

x == y; // false
```

The `Type(x)` is indeed `Boolean`, so it performs `ToNumber(x)`, which coerces `true` to `1`. Now, `1 == "42"` is evaluated. The types are still different, so (essentially recursively) we reconsult the algorithm, which just as above will coerce "42" to `42`, and `1 == 42` is clearly `false`.

Reverse it, and we still get the same outcome:

```

var x = "42";
var y = false;

x == y; // false

```

The `Type(y)` is Boolean this time, so `ToNumber(y)` yields 0. `"42" == 0` recursively becomes `42 == 0`, which is of course `false`.

In other words, **the value "42" is neither == true nor == false**. At first, that statement might seem crazy. How can a value be neither truthy nor falsy?

But that's the problem! You're asking the wrong question, entirely. It's not your fault, really. Your brain is tricking you.

`"42"` is indeed truthy, but `"42" == true` is not performing a boolean test/coercion at all, no matter what your brain says. `"42"` is not being coerced to a boolean (`true`), but instead `true` is being coerced to a 1, and then `"42"` is being coerced to 42.

Whether we like it or not, `ToBoolean` is not even involved here, so the truthiness or falsiness of `"42"` is irrelevant to the `==` operation!

What *is* relevant is to understand how the `==` comparison algorithm behaves with all the different type combinations. As it regards a boolean value on either side of the `==`, a boolean always coerces to a number *first*.

If that seems strange to you, you're not alone. I personally would recommend to never, ever, under any circumstances, use `== true` or `== false`. Ever.

But remember, I'm only talking about `==` here. `== true` and `== false` wouldn't allow the coercion, so they're safe from this hidden `ToNumber` coercion.

Consider:

```

var a = "42";

// bad (will fail!):
if (a == true) {
    // ..
}

// also bad (will fail!):
if (a === true) {
    // ..
}

// good enough (works implicitly):
if (a) {
    // ..
}

```

```

// better (works explicitly):
if (!!a) {
  // ..
}

// also great (works explicitly):
if (Boolean( a )) {
  // ..
}

```

If you avoid ever using `== true` or `== false` (aka loose equality with `booleans`) in your code, you'll never have to worry about this truthiness/falsiness mental gotcha.

Comparing: `nulls` to `undefineds` Another example of *implicit* coercion can be seen with `==` loose equality between `null` and `undefined` values. Yet again quoting the ES5 spec, clauses 11.9.3.2-3:

2. If `x` is `null` and `y` is `undefined`, return `true`.
3. If `x` is `undefined` and `y` is `null`, return `true`.

`null` and `undefined`, when compared with `==` loose equality, equate to (aka coerce to) each other (as well as themselves, obviously), and no other values in the entire language.

What this means is that `null` and `undefined` can be treated as indistinguishable for comparison purposes, if you use the `==` loose equality operator to allow their mutual *implicit* coercion.

```

var a = null;
var b;

a == b;      // true
a == null;   // true
b == null;   // true

a == false;  // false
b == false;  // false
a == "";     // false
b == "";     // false
a == 0;      // false
b == 0;      // false

```

The coercion between `null` and `undefined` is safe and predictable, and no other values can give false positives in such a check. I recommend using this coercion to allow `null` and `undefined` to be indistinguishable and thus treated as the same value.

For example:

```
var a = doSomething();  
  
if (a == null) {  
    // ..  
}
```

The `a == null` check will pass only if `doSomething()` returns either `null` or `undefined`, and will fail with any other value, even other falsy values like `0`, `false`, and `""`.

The *explicit* form of the check, which disallows any such coercion, is (I think) unnecessarily much uglier (and perhaps a tiny bit less performant!):

```
var a = doSomething();  
  
if (a === undefined || a === null) {  
    // ..  
}
```

In my opinion, the form `a == null` is yet another example where *implicit* coercion improves code readability, but does so in a reliably safe way.

Comparing: objects to non-objects If an `object/function/array` is compared to a simple scalar primitive (`string`, `number`, or `boolean`), the ES5 spec says in clauses 11.9.3.8-9:

8. If `Type(x)` is either `String` or `Number` and `Type(y)` is `Object`, return the result of the comparison `x == ToPrimitive(y)`.
9. If `Type(x)` is `Object` and `Type(y)` is either `String` or `Number`, return the result of the comparison `ToPrimitive(x) == y`.

Note: You may notice that these clauses only mention `String` and `Number`, but not `Boolean`. That's because, as quoted earlier, clauses 11.9.3.6-7 take care of coercing any `Boolean` operand presented to a `Number` first.

Consider:

```
var a = 42;  
var b = [ 42 ];  
  
a == b; // true
```

The `[42]` value has its `ToPrimitive` abstract operation called (see the “Abstract Value Operations” section earlier), which results in the `"42"` value. From there, it's just `42 == "42"`, which as we've already covered becomes `42 == 42`, so `a` and `b` are found to be coercively equal.

Tip: All the quirks of the `ToPrimitive` abstract operation that we discussed earlier in this chapter (`toString()`, `valueOf()`) apply here as you'd expect.

This can be quite useful if you have a complex data structure that you want to define a custom `valueOf()` method on, to provide a simple value for equality comparison purposes.

In Chapter 3, we covered “unboxing,” where an `Object` wrapper around a primitive value (like from `new String("abc")`, for instance) is unwrapped, and the underlying primitive value (“abc”) is returned. This behavior is related to the `ToPrimitive` coercion in the `==` algorithm:

```
var a = "abc";
var b = Object( a );      // same as `new String( a )`  
  
a === b;                  // false
a == b;                   // true
```

`a == b` is `true` because `b` is coerced (aka “unboxed,” unwrapped) via `ToPrimitive` to its underlying “abc” simple scalar primitive value, which is the same as the value in `a`.

There are some values where this is not the case, though, because of other overriding rules in the `==` algorithm. Consider:

```
var a = null;
var b = Object( a );      // same as `Object()`  
a == b;                  // false  
  
var c = undefined;
var d = Object( c );      // same as `Object()`  
c == d;                  // false  
  
var e = NaN;
var f = Object( e );      // same as `new Number( e )`  
e == f;                  // false
```

The `null` and `undefined` values cannot be boxed – they have no `Object` wrapper equivalent – so `Object(null)` is just like `Object()` in that both just produce a normal object.

`NaN` can be boxed to its `Number` object wrapper equivalent, but when `==` causes an unboxing, the `NaN == NaN` comparison fails because `NaN` is never equal to itself (see Chapter 2).

Edge Cases

Now that we’ve thoroughly examined how the *implicit* coercion of `==` loose equality works (in both sensible and surprising ways), let’s try to call out the worst, craziest corner cases so we can see what we need to avoid to not get bitten with coercion bugs.

First, let's examine how modifying the built-in native prototypes can produce crazy results:

A Number By Any Other Value Would...

```
Number.prototype.valueOf = function() {
    return 3;
};

new Number( 2 ) == 3;    // true
```

Warning: `2 == 3` would not have fallen into this trap, because neither `2` nor `3` would have invoked the built-in `Number.prototype.valueOf()` method because both are already primitive `number` values and can be compared directly. However, `new Number(2)` must go through the `ToPrimitive` coercion, and thus invoke `valueOf()`.

Evil, huh? Of course it is. No one should ever do such a thing. The fact that you *can* do this is sometimes used as a criticism of coercion and `==`. But that's misdirected frustration. JavaScript is not *bad* because you can do such things, a developer is *bad if they do such things*. Don't fall into the "my programming language should protect me from myself" fallacy.

Next, let's consider another tricky example, which takes the evil from the previous example to another level:

```
if (a == 2 && a == 3) {
    // ..
}
```

You might think this would be impossible, because `a` could never be equal to both `2` and `3` *at the same time*. But "at the same time" is inaccurate, since the first expression `a == 2` happens strictly *before* `a == 3`.

So, what if we make `a.valueOf()` have side effects each time it's called, such that the first time it returns `2` and the second time it's called it returns `3`? Pretty easy:

```
var i = 2;

Number.prototype.valueOf = function() {
    return i++;
};

var a = new Number( 42 );

if (a == 2 && a == 3) {
    console.log( "Yep, this happened." );
}
```

Again, these are evil tricks. Don't do them. But also don't use them as complaints against coercion. Potential abuses of a mechanism are not sufficient evidence to condemn the mechanism. Just avoid these crazy tricks, and stick only with valid and proper usage of coercion.

False-y Comparisons The most common complaint against *implicit* coercion in `==` comparisons comes from how falsy values behave surprisingly when compared to each other.

To illustrate, let's look at a list of the corner-cases around falsy value comparisons, to see which ones are reasonable and which are troublesome:

```
"0" == null;          // false
"0" == undefined;    // false
"0" == false;         // true -- UH OH!
"0" == NaN;          // false
"0" == 0;             // true
"0" == "";            // false

false == null;         // false
false == undefined;   // false
false == NaN;          // false
false == 0;             // true -- UH OH!
false == "";            // true -- UH OH!
false == [];            // true -- UH OH!
false == {};            // false

"" == null;           // false
"" == undefined;      // false
"" == NaN;            // false
"" == 0;               // true -- UH OH!
"" == [];              // true -- UH OH!
"" == {};              // false

0 == null;            // false
0 == undefined;       // false
0 == NaN;             // false
0 == [];               // true -- UH OH!
0 == {};              // false
```

In this list of 24 comparisons, 17 of them are quite reasonable and predictable. For example, we know that `""` and `NaN` are not at all equatable values, and indeed they don't coerce to be loose equals, whereas `"0"` and `0` are reasonably equatable and *do* coerce as loose equals.

However, seven of the comparisons are marked with "UH OH!" because as false positives, they are much more likely gotchas that could trip you up. `""` and `0`

are definitely distinctly different values, and it's rare you'd want to treat them as equatable, so their mutual coercion is troublesome. Note that there aren't any false negatives here.

The Crazy Ones We don't have to stop there, though. We can keep looking for even more troublesome coercions:

```
[] == ![];      // true
```

Oooo, that seems at a higher level of crazy, right!? Your brain may likely trick you that you're comparing a truthy to a falsy value, so the `true` result is surprising, as we *know* a value can never be truthy and falsy at the same time!

But that's not what's actually happening. Let's break it down. What do we know about the `!` unary operator? It explicitly coerces to a `boolean` using the `ToBoolean` rules (and it also flips the parity). So before `[] == ![]` is even processed, it's actually already translated to `[] == false`. We already saw that form in our above list (`false == []`), so its surprise result is *not new* to us.

How about other corner cases?

```
2 == [2];      // true
"" == [null];  // true
```

As we said earlier in our `ToNumber` discussion, the right-hand side `[2]` and `[null]` values will go through a `ToPrimitive` coercion so they can be more readily compared to the simple primitives (2 and `"`, respectively) on the left-hand side. Since the `valueOf()` for `array` values just returns the `array` itself, coercion falls to stringifying the `array`.

`[2]` will become `"2"`, which then is `ToNumber` coerced to 2 for the right-hand side value in the first comparison. `[null]` just straight becomes `"`.

So, `2 == 2` and `"" == ""` are completely understandable.

If your instinct is to still dislike these results, your frustration is not actually with coercion like you probably think it is. It's actually a complaint against the default `array` values' `ToPrimitive` behavior of coercing to a `string` value. More likely, you'd just wish that `[2].toString()` didn't return `"2"`, or that `[null].toString()` didn't return `"`.

But what exactly *should* these `string` coercions result in? I can't really think of any other appropriate `string` coercion of `[2]` than `"2"`, except perhaps `"[2]"` – but that could be very strange in other contexts!

You could rightly make the case that since `String(null)` becomes `"null"`, then `String([null])` should also become `"null"`. That's a reasonable assertion. So, that's the real culprit.

Implicit coercion itself isn't the evil here. Even an *explicit* coercion of `[null]` to a `string` results in `"`. What's at odds is whether it's sensible at all for `array`

values to stringify to the equivalent of their contents, and exactly how that happens. So, direct your frustration at the rules for `String([...])`, because that's where the craziness stems from. Perhaps there should be no stringification coercion of `arrays` at all? But that would have lots of other downsides in other parts of the language.

Another famously cited gotcha:

```
0 == "\n";      // true
```

As we discussed earlier with empty "", "\n" (or " " or any other whitespace combination) is coerced via `ToNumber`, and the result is 0. What other `number` value would you expect whitespace to coerce to? Does it bother you that *explicit Number(" ")* yields 0?

Really the only other reasonable `number` value that empty strings or whitespace strings could coerce to is the `NaN`. But would that *really* be better? The comparison "< " == `NaN` would of course fail, but it's unclear that we'd have really *fixed* any of the underlying concerns.

The chances that a real-world JS program fails because `0 == "\n"` are awfully rare, and such corner cases are easy to avoid.

Type conversions **always** have corner cases, in any language – nothing specific to coercion. The issues here are about second-guessing a certain set of corner cases (and perhaps rightly so!?), but that's not a salient argument against the overall coercion mechanism.

Bottom line: almost any crazy coercion between *normal values* that you're likely to run into (aside from intentionally tricky `valueOf()` or `toString()` hacks as earlier) will boil down to the short seven-item list of gotcha coercions we've identified above.

To contrast against these 24 likely suspects for coercion gotchas, consider another list like this:

```
42 == "43";          // false
"foo" == 42;          // false
"true" == true;       // false

42 == "42";          // true
"foo" == [ "foo" ];   // true
```

In these nonfalsy, noncorner cases (and there are literally an infinite number of comparisons we could put on this list), the coercion results are totally safe, reasonable, and explainable.

Sanity Check OK, we've definitely found some crazy stuff when we've looked deeply into *implicit* coercion. No wonder that most developers claim coercion is evil and should be avoided, right!?

But let's take a step back and do a sanity check.

By way of magnitude comparison, we have *a list* of seven troublesome gotcha coercions, but we have *another list* of (at least 17, but actually infinite) coercions that are totally sane and explainable.

If you're looking for a textbook example of "throwing the baby out with the bathwater," this is it: discarding the entirety of coercion (the infinitely large list of safe and useful behaviors) because of a list of literally just seven gotchas.

The more prudent reaction would be to ask, "how can I use the countless *good parts* of coercion, but avoid the few *bad parts*?"

Let's look again at the *bad* list:

```
"0" == false;          // true -- UH OH!
false == 0;            // true -- UH OH!
false == "";           // true -- UH OH!
false == [];           // true -- UH OH!
"" == 0;               // true -- UH OH!
"" == [];              // true -- UH OH!
0 == [];               // true -- UH OH!
```

Four of the seven items on this list involve `== false` comparison, which we said earlier you should **always, always** avoid. That's a pretty easy rule to remember.

Now the list is down to three.

```
"" == 0;               // true -- UH OH!
"" == [];              // true -- UH OH!
0 == [];               // true -- UH OH!
```

Are these reasonable coercions you'd do in a normal JavaScript program? Under what conditions would they really happen?

I don't think it's terribly likely that you'd literally use `== []` in a `boolean` test in your program, at least not if you know what you're doing. You'd probably instead be doing `== ""` or `== 0`, like:

```
function doSomething(a) {
  if (a == "") {
    // ...
  }
}
```

You'd have an oops if you accidentally called `doSomething(0)` or `doSomething([])`. Another scenario:

```
function doSomething(a,b) {
  if (a == b) {
    // ...
  }
}
```

```
    }  
}
```

Again, this could break if you did something like `doSomething("",0)` or `doSomething([], "")`.

So, while the situations *can* exist where these coercions will bite you, and you'll want to be careful around them, they're probably not super common on the whole of your code base.

Safely Using Implicit Coercion The most important advice I can give you: examine your program and reason about what values can show up on either side of an `==` comparison. To effectively avoid issues with such comparisons, here's some heuristic rules to follow:

1. If either side of the comparison can have `true` or `false` values, don't ever, EVER use `==`.
2. If either side of the comparison can have `[]`, `""`, or `0` values, seriously consider not using `==`.

In these scenarios, it's almost certainly better to use `==` instead of `==`, to avoid unwanted coercion. Follow those two simple rules and pretty much all the coercion gotchas that could reasonably hurt you will effectively be avoided.

Being more explicit/verbose in these cases will save you from a lot of headaches.

The question of `==` vs. `==` is really appropriately framed as: should you allow coercion for a comparison or not?

There's lots of cases where such coercion can be helpful, allowing you to more tersely express some comparison logic (like with `null` and `undefined`, for example).

In the overall scheme of things, there's relatively few cases where *implicit* coercion is truly dangerous. But in those places, for safety sake, definitely use `==`.

Tip: Another place where coercion is guaranteed *not* to bite you is with the `typeof` operator. `typeof` is always going to return you one of seven strings (see Chapter 1), and none of them are the empty `""` string. As such, there's no case where checking the type of some value is going to run afoul of *implicit* coercion. `typeof x == "function"` is 100% as safe and reliable as `typeof x === "function"`. Literally, the spec says the algorithm will be identical in this situation. So, don't just blindly use `==` everywhere simply because that's what your code tools tell you to do, or (worst of all) because you've been told in some book to **not think about it**. You own the quality of your code.

Is *implicit* coercion evil and dangerous? In a few cases, yes, but overwhelmingly, no.

Be a responsible and mature developer. Learn how to use the power of coercion (both *explicit* and *implicit*) effectively and safely. And teach those around you to do the same.

Here's a handy table made by Alex Dorey (@dorey on GitHub) to visualize a variety of comparisons:

Source: <https://github.com/dorey/JavaScript-Equality-Table>

Abstract Relational Comparison

While this part of *implicit* coercion often gets a lot less attention, it's important nonetheless to think about what happens with `a < b` comparisons (similar to how we just examined `a == b` in depth).

The “Abstract Relational Comparison” algorithm in ES5 section 11.8.5 essentially divides itself into two parts: what to do if the comparison involves both `string` values (second half), or anything else (first half).

Note: The algorithm is only defined for `a < b`. So, `a > b` is handled as `b < a`.

The algorithm first calls `ToPrimitive` coercion on both values, and if the return result of either call is not a `string`, then both values are coerced to `number` values using the `ToNumber` operation rules, and compared numerically.

For example:

```
var a = [ 42 ];
var b = [ "43" ];

a < b; // true
b < a; // false
```

Note: Similar caveats for `-0` and `NaN` apply here as they did in the `==` algorithm discussed earlier.

However, if both values are `strings` for the `<` comparison, simple lexicographic (natural alphabetic) comparison on the characters is performed:

```
var a = [ "42" ];
var b = [ "043" ];

a < b; // false
```

`a` and `b` are *not* coerced to `numbers`, because both of them end up as `strings` after the `ToPrimitive` coercion on the two `arrays`. So, `"42"` is compared character by character to `"043"`, starting with the first characters `"4"` and `"0"`, respectively. Since `"0"` is lexicographically *less than* `"4"`, the comparison returns `false`.

The exact same behavior and reasoning goes for:

```

var a = [ 4, 2 ];
var b = [ 0, 4, 3 ];

a < b; // false

```

Here, **a** becomes "4,2" and **b** becomes "0,4,3", and those lexicographically compare identically to the previous snippet.

What about:

```

var a = { b: 42 };
var b = { b: 43 };

a < b; // ?

```

a < b is also **false**, because **a** becomes `[object Object]` and **b** becomes `[object Object]`, and so clearly **a** is not lexicographically less than **b**.

But strangely:

```

var a = { b: 42 };
var b = { b: 43 };

a < b; // false
a == b; // false
a > b; // false

a <= b; // true
a >= b; // true

```

Why is **a == b** not **true**? They're the same **string** value ("`[object Object]`"), so it seems they should be equal, right? Nope. Recall the previous discussion about how **==** works with **object** references.

But then how are **a <= b** and **a >= b** resulting in **true**, if **a < b** and **a == b** and **a > b** are all **false**?

Because the spec says for **a <= b**, it will actually evaluate **b < a** first, and then negate that result. Since **b < a** is *also false*, the result of **a <= b** is **true**.

That's probably awfully contrary to how you might have explained what **<=** does up to now, which would likely have been the literal: "less than *or* equal to." JS more accurately considers **<=** as "not greater than" `!(a > b)`, which JS treats as `!(b < a)`. Moreover, **a >= b** is explained by first considering it as **b <= a**, and then applying the same reasoning.

Unfortunately, there is no "strict relational comparison" as there is for equality. In other words, there's no way to prevent *implicit* coercion from occurring with relational comparisons like **a < b**, other than to ensure that **a** and **b** are of the same type explicitly before making the comparison.

Use the same reasoning from our earlier `==` vs. `===` sanity check discussion. If coercion is helpful and reasonably safe, like in a `42 < "43"` comparison, **use it**. On the other hand, if you need to be safe about a relational comparison, *explicitly coerce* the values first, before using `<` (or its counterparts).

```
var a = [ 42 ];
var b = "043";

a < b;                      // false -- string comparison!
Number( a ) < Number( b );  // true -- number comparison!
```

Review

In this chapter, we turned our attention to how JavaScript type conversions happen, called **coercion**, which can be characterized as either *explicit* or *implicit*.

Coercion gets a bad rap, but it's actually quite useful in many cases. An important task for the responsible JS developer is to take the time to learn all the ins and outs of coercion to decide which parts will help improve their code, and which parts they really should avoid.

Explicit coercion is code which is obvious that the intent is to convert a value from one type to another. The benefit is improvement in readability and maintainability of code by reducing confusion.

Implicit coercion is coercion that is “hidden” as a side-effect of some other operation, where it's not as obvious that the type conversion will occur. While it may seem that *implicit* coercion is the opposite of *explicit* and is thus bad (and indeed, many think so!), actually *implicit* coercion is also about improving the readability of code.

Especially for *implicit*, coercion must be used responsibly and consciously. Know why you're writing the code you're writing, and how it works. Strive to write code that others will easily be able to learn from and understand as well.

You Don't Know JS: Types & Grammar

Chapter 5: Grammar

The last major topic we want to tackle is how JavaScript's language syntax works (aka its grammar). You may think you know how to write JS, but there's an awful lot of nuance to various parts of the language grammar that lead to confusion and misconception, so we want to dive into those parts and clear some things up.

Note: The term "grammar" may be a little less familiar to readers than the term "syntax." In many ways, they are similar terms, describing the *rules* for how the language works. There are nuanced differences, but they mostly don't matter for our discussion here. The grammar for JavaScript is a structured way to describe how the syntax (operators, keywords, etc.) fits together into well-formed, valid programs. In other words, discussing syntax without grammar would leave out a lot of the important details. So our focus here in this chapter is most accurately described as *grammar*, even though the raw syntax of the language is what developers directly interact with.

Statements & Expressions

It's fairly common for developers to assume that the term "statement" and "expression" are roughly equivalent. But here we need to distinguish between the two, because there are some very important differences in our JS programs.

To draw the distinction, let's borrow from terminology you may be more familiar with: the English language.

A "sentence" is one complete formation of words that expresses a thought. It's comprised of one or more "phrases," each of which can be connected with punctuation marks or conjunction words ("and," "or," etc). A phrase can itself be made up of smaller phrases. Some phrases are incomplete and don't accomplish much by themselves, while other phrases can stand on their own. These rules are collectively called the *grammar* of the English language.

And so it goes with JavaScript grammar. Statements are sentences, expressions are phrases, and operators are conjunctions/punctuation.

Every expression in JS can be evaluated down to a single, specific value result. For example:

```
var a = 3 * 6;  
var b = a;  
b;
```

In this snippet, `3 * 6` is an expression (evaluates to the value `18`). But `a` on the second line is also an expression, as is `b` on the third line. The `a` and `b` expressions

both evaluate to the values stored in those variables at that moment, which also happens to be 18.

Moreover, each of the three lines is a statement containing expressions. `var a = 3 * 6` and `var b = a` are called “declaration statements” because they each declare a variable (and optionally assign a value to it). The `a = 3 * 6` and `b = a` assignments (minus the `vars`) are called assignment expressions.

The third line contains just the expression `b`, but it’s also a statement all by itself (though not a terribly interesting one!). This is generally referred to as an “expression statement.”

Statement Completion Values

It’s a fairly little known fact that statements all have completion values (even if that value is just `undefined`).

How would you even go about seeing the completion value of a statement?

The most obvious answer is to type the statement into your browser’s developer console, because when you execute it, the console by default reports the completion value of the most recent statement it executed.

Let’s consider `var b = a`. What’s the completion value of that statement?

The `b = a` assignment expression results in the value that was assigned (18 above), but the `var` statement itself results in `undefined`. Why? Because `var` statements are defined that way in the spec. If you put `var a = 42;` into your console, you’ll see `undefined` reported back instead of 42.

Note: Technically, it’s a little more complex than that. In the ES5 spec, section 12.2 “Variable Statement,” the `VariableDeclaration` algorithm actually *does* return a value (a `string` containing the name of the variable declared – weird, huh!?), but that value is basically swallowed up (except for use by the `for..in` loop) by the `VariableStatement` algorithm, which forces an empty (aka `undefined`) completion value.

In fact, if you’ve done much code experimenting in your console (or in a JavaScript environment REPL – read/evaluate/print/loop tool), you’ve probably seen `undefined` reported after many different statements, and perhaps never realized why or what that was. Put simply, the console is just reporting the statement’s completion value.

But what the console prints out for the completion value isn’t something we can use inside our program. So how can we capture the completion value?

That’s a much more complicated task. Before we explain *how*, let’s explore *why* you would want to do that.

We need to consider other types of statement completion values. For example, any regular `{ .. }` block has a completion value of the completion value of its

last contained statement/expression.

Consider:

```
var b;

if (true) {
  b = 4 + 38;
}
```

If you typed that into your console/REPL, you'd probably see 42 reported, since 42 is the completion value of the `if` block, which took on the completion value of its last assignment expression statement `b = 4 + 38`.

In other words, the completion value of a block is like an *implicit return* of the last statement value in the block.

Note: This is conceptually familiar in languages like CoffeeScript, which have implicit `return` values from `functions` that are the same as the last statement value in the function.

But there's an obvious problem. This kind of code doesn't work:

```
var a, b;

a = if (true) {
  b = 4 + 38;
};
```

We can't capture the completion value of a statement and assign it into another variable in any easy syntactic/grammatical way (at least not yet!).

So, what can we do?

Warning: For demo purposes only – don't actually do the following in your real code!

We could use the much maligned `eval(..)` (sometimes pronounced “evil”) function to capture this completion value.

```
var a, b;

a = eval( "if (true) { b = 4 + 38; }" );

a; // 42
```

Yeeeeaaahhhh. That's terribly ugly. But it works! And it illustrates the point that statement completion values are a real thing that can be captured not just in our console but in our programs.

There's a proposal for ES7 called “do expression.” Here's how it might work:

```
var a, b;
```

```
a = do {
  if (true) {
    b = 4 + 38;
  }
};
```

```
a; // 42
```

The `do { ... }` expression executes a block (with one or many statements in it), and the final statement completion value inside the block becomes the completion value of the `do` expression, which can then be assigned to `a` as shown.

The general idea is to be able to treat statements as expressions – they can show up inside other statements – without needing to wrap them in an inline function expression and perform an explicit `return ...`.

For now, statement completion values are not much more than trivia. But they're probably going to take on more significance as JS evolves, and hopefully `do { ... }` expressions will reduce the temptation to use stuff like `eval(..)`.

Warning: Repeating my earlier admonition: avoid `eval(..)`. Seriously. See the *Scope & Closures* title of this series for more explanation.

Expression Side Effects

Most expressions don't have side effects. For example:

```
var a = 2;
var b = a + 3;
```

The expression `a + 3` did not *itself* have a side effect, like for instance changing `a`. It had a result, which is 5, and that result was assigned to `b` in the statement `b = a + 3`.

The most common example of an expression with (possible) side effects is a function call expression:

```
function foo() {
  a = a + 1;
}

var a = 1;
foo(); // result: `undefined`, side effect: changed `a`
```

There are other side-effecting expressions, though. For example:

```
var a = 42;
var b = a++;
```

The expression `a++` has two separate behaviors. *First*, it returns the current value of `a`, which is 42 (which then gets assigned to `b`). But *next*, it changes the value of `a` itself, incrementing it by one.

```
var a = 42;  
var b = a++;  
  
a; // 43  
b; // 42
```

Many developers would mistakenly believe that `b` has value 43 just like `a` does. But the confusion comes from not fully considering the *when* of the side effects of the `++` operator.

The `++` increment operator and the `--` decrement operator are both unary operators (see Chapter 4), which can be used in either a postfix (“after”) position or prefix (“before”) position.

```
var a = 42;  
  
a++; // 42  
a; // 43  
  
++a; // 44  
a; // 44
```

When `++` is used in the prefix position as `++a`, its side effect (incrementing `a`) happens *before* the value is returned from the expression, rather than *after* as with `a++`.

Note: Would you think `++a++` was legal syntax? If you try it, you’ll get a `ReferenceError` error, but why? Because side-effecting operators **require a variable reference** to target their side effects to. For `++a++`, the `a++` part is evaluated first (because of operator precedence – see below), which gives back the value of `a` *before* the increment. But then it tries to evaluate `++42`, which (if you try it) gives the same `ReferenceError` error, since `++` can’t have a side effect directly on a value like 42.

It is sometimes mistakenly thought that you can encapsulate the *after* side effect of `a++` by wrapping it in a `()` pair, like:

```
var a = 42;  
var b = (a++);  
  
a; // 43  
b; // 42
```

Unfortunately, `()` itself doesn’t define a new wrapped expression that would be evaluated *after* the *after side effect* of the `a++` expression, as we might have hoped. In fact, even if it did, `a++` returns 42 first, and unless you have another

expression that reevaluates `a` after the side effect of `++`, you're not going to get `43` from that expression, so `b` will not be assigned `43`.

There's an option, though: the `,` statement-series comma operator. This operator allows you to string together multiple standalone expression statements into a single statement:

```
var a = 42, b;  
b = ( a++, a );  
  
a; // 43  
b; // 43
```

Note: The `(...)` around `a++`, `a` is required here. The reason is operator precedence, which we'll cover later in this chapter.

The expression `a++`, `a` means that the second `a` statement expression gets evaluated *after* the *after side effects* of the first `a++` statement expression, which means it returns the `43` value for assignment to `b`.

Another example of a side-effecting operator is `delete`. As we showed in Chapter 2, `delete` is used to remove a property from an `object` or a slot from an `array`. But it's usually just called as a standalone statement:

```
var obj = {  
  a: 42  
};  
  
obj.a; // 42  
delete obj.a; // true  
obj.a; // undefined
```

The result value of the `delete` operator is `true` if the requested operation is valid/allowable, or `false` otherwise. But the side effect of the operator is that it removes the property (or array slot).

Note: What do we mean by valid/allowable? Nonexistent properties, or properties that exist and are configurable (see Chapter 3 of the *this & Object Prototypes* title of this series) will return `true` from the `delete` operator. Otherwise, the result will be `false` or an error.

One last example of a side-effecting operator, which may at once be both obvious and nonobvious, is the `=` assignment operator.

Consider:

```
var a;  
  
a = 42; // 42  
a; // 42
```

It may not seem like `=` in `a = 42` is a side-effecting operator for the expression. But if we examine the result value of the `a = 42` statement, it's the value that was just assigned (42), so the assignment of that same value into `a` is essentially a side effect.

Tip: The same reasoning about side effects goes for the compound-assignment operators like `+=`, `-=`, etc. For example, `a = b += 2` is processed first as `b += 2` (which is `b = b + 2`), and the result of *that* = assignment is then assigned to `a`.

This behavior that an assignment expression (or statement) results in the assigned value is primarily useful for chained assignments, such as:

```
var a, b, c;  
  
a = b = c = 42;
```

Here, `c = 42` is evaluated to 42 (with the side effect of assigning 42 to `c`), then `b = 42` is evaluated to 42 (with the side effect of assigning 42 to `b`), and finally `a = 42` is evaluated (with the side effect of assigning 42 to `a`).

Warning: A common mistake developers make with chained assignments is like `var a = b = 42`. While this looks like the same thing, it's not. If that statement were to happen without there also being a separate `var b` (somewhere in the scope) to formally declare `b`, then `var a = b = 42` would not declare `b` directly. Depending on `strict` mode, that would either throw an error or create an accidental global (see the *Scope & Closures* title of this series).

Another scenario to consider:

```
function vowels(str) {  
  var matches;  
  
  if (str) {  
    // pull out all the vowels  
    matches = str.match( /[aeiou]/g );  
  
    if (matches) {  
      return matches;  
    }  
  }  
  
  vowels( "Hello World" ); // ["e", "o", "o"]
```

This works, and many developers prefer such. But using an idiom where we take advantage of the assignment side effect, we can simplify by combining the two `if` statements into one:

```
function vowels(str) {
```

```

var matches;

// pull out all the vowels
if (str && (matches = str.match( /[aeiou]/g ))) {
  return matches;
}
}

vowels( "Hello World" ); // ["e", "o", "o"]

```

Note: The (...) around `matches = str.match..` is required. The reason is operator precedence, which we'll cover in the "Operator Precedence" section later in this chapter.

I prefer this shorter style, as I think it makes it clearer that the two conditionals are in fact related rather than separate. But as with most stylistic choices in JS, it's purely opinion which one is *better*.

Contextual Rules

There are quite a few places in the JavaScript grammar rules where the same syntax means different things depending on where/how it's used. This kind of thing can, in isolation, cause quite a bit of confusion.

We won't exhaustively list all such cases here, but just call out a few of the common ones.

{ .. } Curly Braces There's two main places (and more coming as JS evolves!) that a pair of { .. } curly braces will show up in your code. Let's take a look at each of them.

Object Literals First, as an object literal:

```

// assume there's a `bar()` function defined

var a = {
  foo: bar()
};

```

How do we know this is an object literal? Because the { .. } pair is a value that's getting assigned to `a`.

Note: The `a` reference is called an "l-value" (aka left-hand value) since it's the target of an assignment. The { .. } pair is an "r-value" (aka right-hand value) since it's used *just* as a value (in this case as the source of an assignment).

Labels What happens if we remove the `var a =` part of the above snippet?

```
// assume there's a `bar()` function defined

{
  foo: bar()
}
```

A lot of developers assume that the `{ .. }` pair is just a standalone `object` literal that doesn't get assigned anywhere. But it's actually entirely different.

Here, `{ .. }` is just a regular code block. It's not very idiomatic in JavaScript (much more so in other languages!) to have a standalone `{ .. }` block like that, but it's perfectly valid JS grammar. It can be especially helpful when combined with `let` block-scoping declarations (see the *Scope & Closures* title in this series).

The `{ .. }` code block here is functionally pretty much identical to the code block being attached to some statement, like a `for/while` loop, `if` conditional, etc.

But if it's a normal block of code, what's that bizarre looking `foo: bar()` syntax, and how is that legal?

It's because of a little known (and, frankly, discouraged) feature in JavaScript called "labeled statements." `foo` is a label for the statement `bar()` (which has omitted its trailing `;` – see "Automatic Semicolons" later in this chapter). But what's the point of a labeled statement?

If JavaScript had a `goto` statement, you'd theoretically be able to say `goto foo` and have execution jump to that location in code. `gotos` are usually considered terrible coding idioms as they make code much harder to understand (aka "spaghetti code"), so it's a *very good thing* that JavaScript doesn't have a general `goto`.

However, JS *does* support a limited, special form of `goto`: labeled jumps. Both the `continue` and `break` statements can optionally accept a specified label, in which case the program flow "jumps" kind of like a `goto`. Consider:

```
// `foo` labeled-loop
foo: for (var i=0; i<4; i++) {
  for (var j=0; j<4; j++) {
    // whenever the loops meet, continue outer loop
    if (j == i) {
      // jump to the next iteration of
      // the `foo` labeled-loop
      continue foo;
    }

    // skip odd multiples
    if ((j * i) % 2 == 1) {
      // normal (non-labeled) `continue` of inner loop
    }
}
```

```

        continue;
    }

    console.log( i, j );
}
}
// 1 0
// 2 0
// 2 1
// 3 0
// 3 2

```

Note: `continue foo` does not mean “go to the ‘foo’ labeled position to continue”, but rather, “continue the loop that is labeled ‘foo’ with its next iteration.” So, it’s not *really* an arbitrary `goto`.

As you can see, we skipped over the odd-multiple 3 1 iteration, but the labeled-loop jump also skipped iterations 1 1 and 2 2.

Perhaps a slightly more useful form of the labeled jump is with `break __` from inside an inner loop where you want to break out of the outer loop. Without a labeled `break`, this same logic could sometimes be rather awkward to write:

```

// `foo` labeled-loop
foo: for (var i=0; i<4; i++) {
    for (var j=0; j<4; j++) {
        if ((i * j) >= 3) {
            console.log( "stopping!", i, j );
            // break out of the `foo` labeled loop
            break foo;
        }

        console.log( i, j );
    }
}
// 0 0
// 0 1
// 0 2
// 0 3
// 1 0
// 1 1
// 1 2
// stopping! 1 3

```

Note: `break foo` does not mean “go to the ‘foo’ labeled position to continue,” but rather, “break out of the loop/block that is labeled ‘foo’ and continue *after* it.” Not exactly a `goto` in the traditional sense, huh?

The nonlabeled `break` alternative to the above would probably need to involve

one or more functions, shared scope variable access, etc. It would quite likely be more confusing than labeled `break`, so here using a labeled `break` is perhaps the better option.

A label can apply to a non-loop block, but only `break` can reference such a non-loop label. You can do a labeled `break` ___ out of any labeled block, but you cannot `continue` ___ a non-loop label, nor can you do a non-labeled `break` out of a block.

```
function foo() {
  // `bar` labeled-block
  bar: {
    console.log( "Hello" );
    break bar;
    console.log( "never runs" );
  }
  console.log( "World" );
}

foo();
// Hello
// World
```

Labeled loops/blocks are extremely uncommon, and often frowned upon. It's best to avoid them if possible; for example using function calls instead of the loop jumps. But there are perhaps some limited cases where they might be useful. If you're going to use a labeled jump, make sure to document what you're doing with plenty of comments!

It's a very common belief that JSON is a proper subset of JS, so a string of JSON (like `{"a":42}` – notice the quotes around the property name as JSON requires!) is thought to be a valid JavaScript program. **Not true!** Try putting `{"a":42}` into your JS console, and you'll get an error.

That's because statement labels cannot have quotes around them, so `"a"` is not a valid label, and thus `: can't come right after it.`

So, JSON is truly a subset of JS syntax, but JSON is not valid JS grammar by itself.

One extremely common misconception along these lines is that if you were to load a JS file into a `<script src=..>` tag that only has JSON content in it (like from an API call), the data would be read as valid JavaScript but just be inaccessible to the program. JSON-P (the practice of wrapping the JSON data in a function call, like `foo({"a":42})`) is usually said to solve this inaccessibility by sending the value to one of your program's functions.

Not true! The totally valid JSON value `{"a":42}` by itself would actually throw a JS error because it'd be interpreted as a statement block with an invalid label. But `foo({"a":42})` is valid JS because in it, `{"a":42}` is an `object`

literal value being passed to `foo(..)`. So, properly said, **JSON-P makes JSON into valid JS grammar!**

Blocks Another commonly cited JS gotcha (related to coercion – see Chapter 4) is:

```
[] + {}; // "[object Object]"  
{} + []; // 0
```

This seems to imply the `+` operator gives different results depending on whether the first operand is the `[]` or the `{}`. But that actually has nothing to do with it!

On the first line, `{}` appears in the `+` operator's expression, and is therefore interpreted as an actual value (an empty `object`). Chapter 4 explained that `[]` is coerced to `""` and thus `{}` is coerced to a `string` value as well: `"[object Object]"`.

But on the second line, `{}` is interpreted as a standalone `{}` empty block (which does nothing). Blocks don't need semicolons to terminate them, so the lack of one here isn't a problem. Finally, `+` `[]` is an expression that *explicitly coerces* (see Chapter 4) the `[]` to a `number`, which is the `0` value.

Object Destructuring Starting with ES6, another place that you'll see `{ .. }` pairs showing up is with “destructuring assignments” (see the *ES6 & Beyond* title of this series for more info), specifically `object` destructuring. Consider:

```
function getData() {  
  // ..  
  return {  
    a: 42,  
    b: "foo"  
  };  
  
  var { a, b } = getData();  
  
  console.log( a, b ); // 42 "foo"
```

As you can probably tell, `var { a, b } = ..` is a form of ES6 destructuring assignment, which is roughly equivalent to:

```
var res = getData();  
var a = res.a;  
var b = res.b;
```

Note: `{ a, b }` is actually ES6 destructuring shorthand for `{ a: a, b: b }`, so either will work, but it's expected that the shorter `{ a, b }` will become the preferred form.

Object destructuring with a `{ ... }` pair can also be used for named function arguments, which is sugar for this same sort of implicit object property assignment:

```
function foo({ a, b, c }) {
  // no need for:
  // var a = obj.a, b = obj.b, c = obj.c
  console.log( a, b, c );
}

foo( {
  c: [1,2,3],
  a: 42,
  b: "foo"
} );  // 42 "foo" [1, 2, 3]
```

So, the context we use `{ ... }` pairs in entirely determines what they mean, which illustrates the difference between syntax and grammar. It's very important to understand these nuances to avoid unexpected interpretations by the JS engine.

else if And Optional Blocks It's a common misconception that JavaScript has an `else if` clause, because you can do:

```
if (a) {
  // ...
}
else if (b) {
  // ...
}
else {
  // ...
}
```

But there's a hidden characteristic of the JS grammar here: there is no `else if`. But `if` and `else` statements are allowed to omit the `{ }` around their attached block if they only contain a single statement. You've seen this many times before, undoubtedly:

```
if (a) doSomething( a );
```

Many JS style guides will insist that you always use `{ }` around a single statement block, like:

```
if (a) { doSomething( a ); }
```

However, the exact same grammar rule applies to the `else` clause, so the `else if` form you've likely always coded is *actually* parsed as:

```

if (a) {
  // ...
}
else {
  if (b) {
    // ...
  }
  else {
    // ...
  }
}

```

The `if (b) { ... } else { ... }` is a single statement that follows the `else`, so you can either put the surrounding `{ }` in or not. In other words, when you use `else if`, you're technically breaking that common style guide rule and just defining your `else` with a single `if` statement.

Of course, the `else if` idiom is extremely common and results in one less level of indentation, so it's attractive. Whichever way you do it, just call out explicitly in your own style guide/rules and don't assume things like `else if` are direct grammar rules.

Operator Precedence

As we covered in Chapter 4, JavaScript's version of `&&` and `||` are interesting in that they select and return one of their operands, rather than just resulting in `true` or `false`. That's easy to reason about if there are only two operands and one operator.

```

var a = 42;
var b = "foo";

a && b; // "foo"
a || b; // 42

```

But what about when there's two operators involved, and three operands?

```

var a = 42;
var b = "foo";
var c = [1,2,3];

a && b || c; // ???
a || b && c; // ???

```

To understand what those expressions result in, we're going to need to understand what rules govern how the operators are processed when there's more than one present in an expression.

These rules are called “operator precedence.”

I bet most readers feel they have a decent grasp on operator precedence. But as with everything else we've covered in this book series, we're going to poke and prod at that understanding to see just how solid it really is, and hopefully learn a few new things along the way.

Recall the example from above:

```
var a = 42, b;  
b = ( a++, a );  
  
a; // 43  
b; // 43
```

But what would happen if we remove the ()?

```
var a = 42, b;  
b = a++, a;  
  
a; // 43  
b; // 42
```

Wait! Why did that change the value assigned to b?

Because the , operator has a lower precedence than the = operator. So, b = a++, a is interpreted as (b = a++), a. Because (as we explained earlier) a++ has *after side effects*, the assigned value to b is the value 42 before the ++ changes a.

This is just a simple matter of needing to understand operator precedence. If you're going to use , as a statement-series operator, it's important to know that it actually has the lowest precedence. Every other operator will more tightly bind than , will.

Now, recall this example from above:

```
if (str && (matches = str.match( /[aeiou]/g ))) {  
    // ...  
}
```

We said the () around the assignment is required, but why? Because && has higher precedence than =, so without the () to force the binding, the expression would instead be treated as (str && matches) = str.match... But this would be an error, because the result of (str && matches) isn't going to be a variable, but instead a value (in this case `undefined`), and so it can't be the left-hand side of an = assignment!

OK, so you probably think you've got this operator precedence thing down.

Let's move on to a more complex example (which we'll carry throughout the next several sections of this chapter) to *really* test your understanding:

```

var a = 42;
var b = "foo";
var c = false;

var d = a && b || c ? c || b ? a : c && b : a;

d;      // ??

```

OK, evil, I admit it. No one would write a string of expressions like that, right? *Probably* not, but we're going to use it to examine various issues around chaining multiple operators together, which *is* a very common task.

The result above is 42. But that's not nearly as interesting as how we can figure out that answer without just plugging it into a JS program to let JavaScript sort it out.

Let's dig in.

The first question – it may not have even occurred to you to ask – is, does the first part (a && b || c) behave like (a && b) || c or like a && (b || c)? Do you know for certain? Can you even convince yourself they are actually different?

```

(false && true) || true;    // true
false && (true || true);    // false

```

So, there's proof they're different. But still, how does **false** && **true** || **true** behave? The answer:

```

false && true || true;    // true
(false && true) || true;    // true

```

So we have our answer. The **&&** operator is evaluated first and the **||** operator is evaluated second.

But is that just because of left-to-right processing? Let's reverse the order of operators:

```

true || false && false;    // true

(true || false) && false;    // false -- nope
true || (false && false);    // true -- winner, winner!

```

Now we've proved that **&&** is evaluated first and then **||**, and in this case that was actually counter to generally expected left-to-right processing.

So what caused the behavior? **Operator precedence**.

Every language defines its own operator precedence list. It's dismaying, though, just how uncommon it is that JS developers have read JS's list.

If you knew it well, the above examples wouldn't have tripped you up in the slightest, because you'd already know that **&&** is more precedent than **||**. But I

bet a fair amount of readers had to think about it a little bit.

Note: Unfortunately, the JS spec doesn't really have its operator precedence list in a convenient, single location. You have to parse through and understand all the grammar rules. So we'll try to lay out the more common and useful bits here in a more convenient format. For a complete list of operator precedence, see "Operator Precedence" on the MDN site (* https://developer.mozilla.org/en-US/docs/Web/JavaScript/Reference/Operators/Operator_Precidence).

Short Circuited

In Chapter 4, we mentioned in a side note the "short circuiting" nature of operators like `&&` and `||`. Let's revisit that in more detail now.

For both `&&` and `||` operators, the right-hand operand will **not be evaluated** if the left-hand operand is sufficient to determine the outcome of the operation. Hence, the name "short circuited" (in that if possible, it will take an early shortcut out).

For example, with `a && b`, `b` is not evaluated if `a` is falsy, because the result of the `&&` operand is already certain, so there's no point in bothering to check `b`. Likewise, with `a || b`, if `a` is truthy, the result of the operand is already certain, so there's no reason to check `b`.

This short circuiting can be very helpful and is commonly used:

```
function doSomething(opts) {
  if (opts && opts.cool) {
    // ...
  }
}
```

The `opts` part of the `opts && opts.cool` test acts as sort of a guard, because if `opts` is unset (or is not an object), the expression `opts.cool` would throw an error. The `opts` test failing plus the short circuiting means that `opts.cool` won't even be evaluated, thus no error!

Similarly, you can use `||` short circuiting:

```
function doSomething(opts) {
  if (opts.cache || primeCache()) {
    // ...
  }
}
```

Here, we're checking for `opts.cache` first, and if it's present, we don't call the `primeCache()` function, thus avoiding potentially unnecessary work.

Tighter Binding

But let's turn our attention back to that earlier complex statement example with all the chained operators, specifically the `? :` ternary operator parts. Does the `? :` operator have more or less precedence than the `&&` and `||` operators?

```
a && b || c ? c || b ? a : c && b : a
```

Is that more like this:

```
a && b || (c ? c || (b ? a : c) && b : a)
```

or this?

```
(a && b || c) ? (c || b) ? a : (c && b) : a
```

The answer is the second one. But why?

Because `&&` is more precedent than `||`, and `||` is more precedent than `? :`.

So, the expression `(a && b || c)` is evaluated *first* before the `? :` it participates in. Another way this is commonly explained is that `&&` and `||` “bind more tightly” than `? :`. If the reverse was true, then `c ? c...` would bind more tightly, and it would behave (as the first choice) like `a && b || (c ? c...)`.

Associativity

So, the `&&` and `||` operators bind first, then the `? :` operator. But what about multiple operators of the same precedence? Do they always process left-to-right or right-to-left?

In general, operators are either left-associative or right-associative, referring to whether **grouping happens from the left or from the right**.

It's important to note that associativity is *not* the same thing as left-to-right or right-to-left processing.

But why does it matter whether processing is left-to-right or right-to-left? Because expressions can have side effects, like for instance with function calls:

```
var a = foo() && bar();
```

Here, `foo()` is evaluated first, and then possibly `bar()` depending on the result of the `foo()` expression. That definitely could result in different program behavior than if `bar()` was called before `foo()`.

But this behavior is *just* left-to-right processing (the default behavior in JavaScript!) – it has nothing to do with the associativity of `&&`. In that example, since there's only one `&&` and thus no relevant grouping here, associativity doesn't even come into play.

But with an expression like `a && b && c`, grouping *will* happen implicitly, meaning that either `a && b` or `b && c` will be evaluated first.

Technically, `a && b && c` will be handled as `(a && b) && c`, because `&&` is left-associative (so is `||`, by the way). However, the right-associative alternative `a && (b && c)` behaves observably the same way. For the same values, the same expressions are evaluated in the same order.

Note: If hypothetically `&&` was right-associative, it would be processed the same as if you manually used `()` to create grouping like `a && (b && c)`. But that still **doesn't mean** that `c` would be processed before `b`. Right-associativity does **not** mean right-to-left evaluation, it means right-to-left **grouping**. Either way, regardless of the grouping/associativity, the strict ordering of evaluation will be `a`, then `b`, then `c` (aka left-to-right).

So it doesn't really matter that much that `&&` and `||` are left-associative, other than to be accurate in how we discuss their definitions.

But that's not always the case. Some operators would behave very differently depending on left-associativity vs. right-associativity.

Consider the `? :` (“ternary” or “conditional”) operator:

`a ? b : c ? d : e;`

`? :` is right-associative, so which grouping represents how it will be processed?

- `a ? b : (c ? d : e)`
- `(a ? b : c) ? d : e`

The answer is `a ? b : (c ? d : e)`. Unlike with `&&` and `||` above, the right-associativity here actually matters, as `(a ? b : c) ? d : e` *will* behave differently for some (but not all!) combinations of values.

One such example:

```
true ? false : true ? true : true;      // false

true ? false : (true ? true : true);    // false
(true ? false : true) ? true : true;    // true
```

Even more nuanced differences lurk with other value combinations, even if the end result is the same. Consider:

```
true ? false : true ? true : false;      // false

true ? false : (true ? true : false);    // false
(true ? false : true) ? true : false;    // false
```

From that scenario, the same end result implies that the grouping is moot. However:

```
var a = true, b = false, c = true, d = true, e = false;

a ? b : (c ? d : e); // false, evaluates only `a` and `b`
(a ? b : c) ? d : e; // false, evaluates `a`, `b` AND `e`
```

So, we've clearly proved that `? :` is right-associative, and that it actually matters with respect to how the operator behaves if chained with itself.

Another example of right-associativity (grouping) is the `=` operator. Recall the chained assignment example from earlier in the chapter:

```
var a, b, c;
```

```
a = b = c = 42;
```

We asserted earlier that `a = b = c = 42` is processed by first evaluating the `c = 42` assignment, then `b = ...`, and finally `a = ...`. Why? Because of the right-associativity, which actually treats the statement like this: `a = (b = (c = 42))`.

Remember our running complex assignment expression example from earlier in the chapter?

```
var a = 42;
var b = "foo";
var c = false;
```

```
var d = a && b || c ? c || b ? a : c && b : a;
```

```
d; // 42
```

Armed with our knowledge of precedence and associativity, we should now be able to break down the code into its grouping behavior like this:

```
((a && b) || c) ? ((c || b) ? a : (c && b)) : a
```

Or, to present it indented if that's easier to understand:

```
(  
  (a && b)  
  ||  
  c  
)  
?  
(  
  (c || b)  
  ?  
  a  
  :  
  (c && b)  
)  
:  
a
```

Let's solve it now:

1. `(a && b)` is "foo".
2. `"foo" || c` is "foo".
3. For the first `?` test, `"foo"` is truthy.
4. `(c || b)` is "foo".
5. For the second `?` test, `"foo"` is truthy.
6. `a` is 42.

That's it, we're done! The answer is 42, just as we saw earlier. That actually wasn't so hard, was it?

Disambiguation

You should now have a much better grasp on operator precedence (and associativity) and feel much more comfortable understanding how code with multiple chained operators will behave.

But an important question remains: should we all write code understanding and perfectly relying on all the rules of operator precedence/associativity? Should we only use `()` manual grouping when it's necessary to force a different processing binding/order?

Or, on the other hand, should we recognize that even though such rules *are in fact* learnable, there's enough gotchas to warrant ignoring automatic precedence/associativity? If so, should we thus always use `()` manual grouping and remove all reliance on these automatic behaviors?

This debate is highly subjective, and heavily symmetrical to the debate in Chapter 4 over *implicit* coercion. Most developers feel the same way about both debates: either they accept both behaviors and code expecting them, or they discard both behaviors and stick to manual/explicit idioms.

Of course, I cannot answer this question definitively for the reader here anymore than I could in Chapter 4. But I've presented you the pros and cons, and hopefully encouraged enough deeper understanding that you can make informed rather than hype-driven decisions.

In my opinion, there's an important middle ground. We should mix both operator precedence/associativity *and* `()` manual grouping into our programs – I argue the same way in Chapter 4 for healthy/safe usage of *implicit* coercion, but certainly don't endorse it exclusively without bounds.

For example, `if (a && b && c) ..` is perfectly OK to me, and I wouldn't do `if ((a && b) && c) ..` just to explicitly call out the associativity, because I think it's overly verbose.

On the other hand, if I needed to chain two `?` : conditional operators together, I'd certainly use `()` manual grouping to make it absolutely clear what my intended logic is.

Thus, my advice here is similar to that of Chapter 4: **use operator precedence/associativity where it leads to shorter and cleaner code, but use () manual grouping in places where it helps create clarity and reduce confusion.**

Automatic Semicolons

ASI (Automatic Semicolon Insertion) is when JavaScript assumes a ; in certain places in your JS program even if you didn't put one there.

Why would it do that? Because if you omit even a single required ; your program would fail. Not very forgiving. ASI allows JS to be tolerant of certain places where ; aren't commonly thought to be necessary.

It's important to note that ASI will only take effect in the presence of a newline (aka line break). Semicolons are not inserted in the middle of a line.

Basically, if the JS parser parses a line where a parser error would occur (a missing expected ;), and it can reasonably insert one, it does so. What's reasonable for insertion? Only if there's nothing but whitespace and/or comments between the end of some statement and that line's newline/line break.

Consider:

```
var a = 42, b  
c;
```

Should JS treat the c on the next line as part of the var statement? It certainly would if a , had come anywhere (even another line) between b and c. But since there isn't one, JS assumes instead that there's an implied ; (at the newline) after b. Thus, c; is left as a standalone expression statement.

Similarly:

```
var a = 42, b = "foo";  
  
a  
b // "foo"
```

That's still a valid program without error, because expression statements also accept ASI.

There's certain places where ASI is helpful, like for instance:

```
var a = 42;  
  
do {  
    // ..  
} while (a) // <-- ; expected here!  
a;
```

The grammar requires a ; after a `do..while` loop, but not after `while` or `for` loops. But most developers don't remember that! So, ASI helpfully steps in and inserts one.

As we said earlier in the chapter, statement blocks do not require ; termination, so ASI isn't necessary:

```
var a = 42;

while (a) {
  // ..
} // <-- no ; expected here
a;
```

The other major case where ASI kicks in is with the `break`, `continue`, `return`, and (ES6) `yield` keywords:

```
function foo(a) {
  if (!a) return
  a *= 2;
  // ..
}
```

The `return` statement doesn't carry across the newline to the `a *= 2` expression, as ASI assumes the ; terminating the `return` statement. Of course, `return` statements *can* easily break across multiple lines, just not when there's nothing after `return` but the newline/line break.

```
function foo(a) {
  return (
    a * 2 + 3 / 12
  );
}
```

Identical reasoning applies to `break`, `continue`, and `yield`.

Error Correction

One of the most hotly contested *religious wars* in the JS community (besides tabs vs. spaces) is whether to rely heavily/exclusively on ASI or not.

Most, but not all, semicolons are optional, but the two ;s in the `for (...)` .. loop header are required.

On the pro side of this debate, many developers believe that ASI is a useful mechanism that allows them to write more terse (and more "beautiful") code by omitting all but the strictly required ;s (which are very few). It is often asserted that ASI makes many ;s optional, so a correctly written program *without them* is no different than a correctly written program *with them*.

On the con side of the debate, many other developers will assert that there are *too many* places that can be accidental gotchas, especially for newer, less experienced developers, where unintended ;s being magically inserted change the meaning. Similarly, some developers will argue that if they omit a semicolon, it's a flat-out mistake, and they want their tools (linters, etc.) to catch it before the JS engine *corrects* the mistake under the covers.

Let me just share my perspective. A strict reading of the spec implies that ASI is an “error correction” routine. What kind of error, you may ask? Specifically, a **parser error**. In other words, in an attempt to have the parser fail less, ASI lets it be more tolerant.

But tolerant of what? In my view, the only way a **parser error** occurs is if it's given an incorrect/errored program to parse. So, while ASI is strictly correcting parser errors, the only way it can get such errors is if there were first program authoring errors – omitting semicolons where the grammar rules require them.

So, to put it more bluntly, when I hear someone claim that they want to omit “optional semicolons,” my brain translates that claim to “I want to write the most parser-broken program I can that will still work.”

I find that to be a ludicrous position to take and the arguments of saving keystrokes and having more “beautiful code” to be weak at best.

Furthermore, I don't agree that this is the same thing as the spaces vs tabs debate – that it's purely cosmetic – but rather I believe it's a fundamental question of writing code that adheres to grammar requirements vs. code that relies on grammar exceptions to just barely skate through.

Another way of looking at it is that relying on ASI is essentially considering newlines to be significant “whitespace.” Other languages like Python have true significant whitespace. But is it really appropriate to think of JavaScript as having significant newlines as it stands today?

My take: **use semicolons wherever you know they are “required,” and limit your assumptions about ASI to a minimum.**

But don't just take my word for it. Back in 2012, creator of JavaScript Brendan Eich said (<http://brendaneich.com/2012/04/the-infernal-semicolon/>) the following:

The moral of this story: ASI is (formally speaking) a syntactic error correction procedure. If you start to code as if it were a universal significant-newline rule, you will get into trouble. ... I wish I had made newlines more significant in JS back in those ten days in May, 1995. ... Be careful not to use ASI as if it gave JS significant newlines.

Errors

Not only does JavaScript have different *subtypes* of errors (`TypeError`, `ReferenceError`, `SyntaxError`, etc.), but also the grammar defines certain errors to be enforced at compile time, as compared to all other errors that happen during runtime.

In particular, there have long been a number of specific conditions that should be caught and reported as “early errors” (during compilation). Any straight-up syntax error is an early error (e.g., `a = ,`), but also the grammar defines things that are syntactically valid but disallowed nonetheless.

Since execution of your code has not begun yet, these errors are not catchable with `try..catch`; they will just fail the parsing/compilation of your program.

Tip: There’s no requirement in the spec about exactly how browsers (and developer tools) should report errors. So you may see variations across browsers in the following error examples, in what specific subtype of error is reported or what the included error message text will be.

One simple example is with syntax inside a regular expression literal. There’s nothing wrong with the JS syntax here, but the invalid regex will throw an early error:

```
var a = /+foo/;      // Error!
```

The target of an assignment must be an identifier (or an ES6 destructuring expression that produces one or more identifiers), so a value like `42` in that position is illegal and can be reported right away:

```
var a;  
42 = a;      // Error!
```

ES5’s `strict` mode defines even more early errors. For example, in `strict` mode, function parameter names cannot be duplicated:

```
function foo(a,b,a) { }          // just fine  
  
function bar(a,b,a) { "use strict"; } // Error!
```

Another `strict` mode early error is an object literal having more than one property of the same name:

```
(function(){  
  "use strict";  
  
  var a = {  
    b: 42,  
    b: 43  
  };          // Error!  
}());
```

Note: Semantically speaking, such errors aren't technically *syntax* errors but more *grammar* errors – the above snippets are syntactically valid. But since there is no `GrammarError` type, some browsers use `SyntaxError` instead.

Using Variables Too Early

ES6 defines a (frankly confusingly named) new concept called the TDZ (“Temporal Dead Zone”).

The TDZ refers to places in code where a variable reference cannot yet be made, because it hasn't reached its required initialization.

The most clear example of this is with ES6 `let` block-scoping:

```
{  
  a = 2;      // ReferenceError!  
  let a;  
}
```

The assignment `a = 2` is accessing the `a` variable (which is indeed block-scoped to the `{ .. }` block) before it's been initialized by the `let a` declaration, so it's in the TDZ for `a` and throws an error.

Interestingly, while `typeof` has an exception to be safe for undeclared variables (see Chapter 1), no such safety exception is made for TDZ references:

```
{  
  typeof a;  // undefined  
  typeof b;  // ReferenceError! (TDZ)  
  let b;  
}
```

Function Arguments

Another example of a TDZ violation can be seen with ES6 default parameter values (see the *ES6 & Beyond* title of this series):

```
var b = 3;  
  
function foo( a = 42, b = a + b + 5 ) {  
  // ...  
}
```

The `b` reference in the assignment would happen in the TDZ for the parameter `b` (not pull in the outer `b` reference), so it will throw an error. However, the `a` in the assignment is fine since by that time it's past the TDZ for parameter `a`.

When using ES6's default parameter values, the default value is applied to the parameter if you either omit an argument, or you pass an `undefined` value in its place:

```

function foo( a = 42, b = a + 1 ) {
  console.log( a, b );
}

foo();           // 42 43
foo( undefined ); // 42 43
foo( 5 );        // 5 6
foo( void 0, 7 ); // 42 7
foo( null );     // null 1

```

Note: `null` is coerced to a `0` value in the `a + 1` expression. See Chapter 4 for more info.

From the ES6 default parameter values perspective, there's no difference between omitting an argument and passing an `undefined` value. However, there is a way to detect the difference in some cases:

```

function foo( a = 42, b = a + 1 ) {
  console.log(
    arguments.length, a, b,
    arguments[0], arguments[1]
  );
}

foo();           // 0 42 43 undefined undefined
foo( 10 );       // 1 10 11 10 undefined
foo( 10, undefined ); // 2 10 11 10 undefined
foo( 10, null ); // 2 10 null 10 null

```

Even though the default parameter values are applied to the `a` and `b` parameters, if no arguments were passed in those slots, the `arguments` array will not have entries.

Conversely, if you pass an `undefined` argument explicitly, an entry will exist in the `arguments` array for that argument, but it will be `undefined` and not (necessarily) the same as the default value that was applied to the named parameter for that same slot.

While ES6 default parameter values can create divergence between the `arguments` array slot and the corresponding named parameter variable, this same disjointedness can also occur in tricky ways in ES5:

```

function foo(a) {
  a = 42;
  console.log( arguments[0] );
}

foo( 2 ); // 42 (linked)
foo();    // undefined (not linked)

```

If you pass an argument, the `arguments` slot and the named parameter are linked to always have the same value. If you omit the argument, no such linkage occurs.

But in `strict` mode, the linkage doesn't exist regardless:

```
function foo(a) {
  "use strict";
  a = 42;
  console.log( arguments[0] );
}

foo( 2 );  // 2 (not linked)
foo();      // undefined (not linked)
```

It's almost certainly a bad idea to ever rely on any such linkage, and in fact the linkage itself is a leaky abstraction that's exposing an underlying implementation detail of the engine, rather than a properly designed feature.

Use of the `arguments` array has been deprecated (especially in favor of ES6 ... rest parameters – see the *ES6 & Beyond* title of this series), but that doesn't mean that it's all bad.

Prior to ES6, `arguments` is the only way to get an array of all passed arguments to pass along to other functions, which turns out to be quite useful. You can also mix named parameters with the `arguments` array and be safe, as long as you follow one simple rule: **never refer to a named parameter and its corresponding arguments slot at the same time**. If you avoid that bad practice, you'll never expose the leaky linkage behavior.

```
function foo(a) {
  console.log( a + arguments[1] ); // safe!
}

foo( 10, 32 ); // 42
```

try..finally

You're probably familiar with how the `try..catch` block works. But have you ever stopped to consider the `finally` clause that can be paired with it? In fact, were you aware that `try` only requires either `catch` or `finally`, though both can be present if needed.

The code in the `finally` clause *always* runs (no matter what), and it always runs right after the `try` (and `catch` if present) finish, before any other code runs. In one sense, you can kind of think of the code in a `finally` clause as being in a callback function that will always be called regardless of how the rest of the block behaves.

So what happens if there's a `return` statement inside a `try` clause? It obviously will return a value, right? But does the calling code that receives that value run

before or after the `finally`?

```
function foo() {
  try {
    return 42;
  }
  finally {
    console.log( "Hello" );
  }

  console.log( "never runs" );
}

console.log( foo() );
// Hello
// 42
```

The `return 42` runs right away, which sets up the completion value from the `foo()` call. This action completes the `try` clause and the `finally` clause immediately runs next. Only then is the `foo()` function complete, so that its completion value is returned back for the `console.log(...)` statement to use.

The exact same behavior is true of a `throw` inside `try`:

```
function foo() {
  try {
    throw 42;
  }
  finally {
    console.log( "Hello" );
  }

  console.log( "never runs" );
}

console.log( foo() );
// Hello
// Uncaught Exception: 42
```

Now, if an exception is thrown (accidentally or intentionally) inside a `finally` clause, it will override as the primary completion of that function. If a previous `return` in the `try` block had set a completion value for the function, that value will be abandoned.

```
function foo() {
  try {
    return 42;
  }
  finally {
```

```

        throw "Oops!";
    }

    console.log( "never runs" );
}

console.log( foo() );
// Uncaught Exception: Oops!

```

It shouldn't be surprising that other nonlinear control statements like `continue` and `break` exhibit similar behavior to `return` and `throw`:

```

for (var i=0; i<10; i++) {
    try {
        continue;
    }
    finally {
        console.log( i );
    }
}
// 0 1 2 3 4 5 6 7 8 9

```

The `console.log(i)` statement runs at the end of the loop iteration, which is caused by the `continue` statement. However, it still runs before the `i++` iteration update statement, which is why the values printed are 0..9 instead of 1..10.

Note: ES6 adds a `yield` statement, in generators (see the *Async & Performance* title of this series) which in some ways can be seen as an intermediate `return` statement. However, unlike a `return`, a `yield` isn't complete until the generator is resumed, which means a `try { ... yield ... }` has not completed. So an attached `finally` clause will not run right after the `yield` like it does with `return`.

A `return` inside a `finally` has the special ability to override a previous `return` from the `try` or `catch` clause, but only if `return` is explicitly called:

```

function foo() {
    try {
        return 42;
    }
    finally {
        // no `return ..` here, so no override
    }
}

function bar() {
    try {
        return 42;

```

```

    }
    finally {
        // override previous `return 42`
        return;
    }
}

function baz() {
    try {
        return 42;
    }
    finally {
        // override previous `return 42`
        return "Hello";
    }
}

foo(); // 42
bar(); // undefined
baz(); // "Hello"

```

Normally, the omission of **return** in a function is the same as **return**; or even **return undefined**;, but inside a **finally** block the omission of **return** does not act like an overriding **return undefined**; it just lets the previous **return** stand.

In fact, we can really up the craziness if we combine **finally** with labeled **break** (discussed earlier in the chapter):

```

function foo() {
    bar: {
        try {
            return 42;
        }
        finally {
            // break out of `bar` labeled block
            break bar;
        }
    }

    console.log( "Crazy" );

    return "Hello";
}

console.log( foo() );
// Crazy

```

```
// Hello
```

But... don't do this. Seriously. Using a `finally` + labeled `break` to effectively cancel a `return` is doing your best to create the most confusing code possible. I'd wager no amount of comments will redeem this code.

switch

Let's briefly explore the `switch` statement, a sort-of syntactic shorthand for an `if..else if..else..` statement chain.

```
switch (a) {
  case 2:
    // do something
    break;
  case 42:
    // do another thing
    break;
  default:
    // fallback to here
}
```

As you can see, it evaluates `a` once, then matches the resulting value to each `case` expression (just simple value expressions here). If a match is found, execution will begin in that matched `case`, and will either go until a `break` is encountered or until the end of the `switch` block is found.

That much may not surprise you, but there are several quirks about `switch` you may not have noticed before.

First, the matching that occurs between the `a` expression and each `case` expression is identical to the `==` algorithm (see Chapter 4). Often times `switches` are used with absolute values in `case` statements, as shown above, so strict matching is appropriate.

However, you may wish to allow coercive equality (aka `==`, see Chapter 4), and to do so you'll need to sort of "hack" the `switch` statement a bit:

```
var a = "42";

switch (true) {
  case a == 10:
    console.log( "10 or '10'" );
    break;
  case a == 42:
    console.log( "42 or '42'" );
    break;
  default:
    // never gets here
```

```
}

// 42 or '42'
```

This works because the `case` clause can have any expression (not just simple values), which means it will strictly match that expression's result to the test expression (`true`). Since `a == 42` results in `true` here, the match is made.

Despite `==`, the `switch` matching itself is still strict, between `true` and `true` here. If the `case` expression resulted in something that was truthy but not strictly `true` (see Chapter 4), it wouldn't work. This can bite you if you're for instance using a "logical operator" like `||` or `&&` in your expression:

```
var a = "hello world";
var b = 10;

switch (true) {
  case (a || b == 10):
    // never gets here
    break;
  default:
    console.log( "Oops" );
}
// Oops
```

Since the result of `(a || b == 10)` is "hello world" and not `true`, the strict match fails. In this case, the fix is to force the expression explicitly to be a `true` or `false`, such as `case !!(a || b == 10):` (see Chapter 4).

Lastly, the `default` clause is optional, and it doesn't necessarily have to come at the end (although that's the strong convention). Even in the `default` clause, the same rules apply about encountering a `break` or not:

```
var a = 10;

switch (a) {
  case 1:
  case 2:
    // never gets here
  default:
    console.log( "default" );
  case 3:
    console.log( "3" );
    break;
  case 4:
    console.log( "4" );
}
// default
// 3
```

Note: As discussed previously about labeled `breaks`, the `break` inside a `case` clause can also be labeled.

The way this snippet processes is that it passes through all the `case` clause matching first, finds no match, then goes back up to the `default` clause and starts executing. Since there's no `break` there, it continues executing in the already skipped over `case 3` block, before stopping once it hits that `break`.

While this sort of round-about logic is clearly possible in JavaScript, there's almost no chance that it's going to make for reasonable or understandable code. Be very skeptical if you find yourself wanting to create such circular logic flow, and if you really do, make sure you include plenty of code comments to explain what you're up to!

Review

JavaScript grammar has plenty of nuance that we as developers should spend a little more time paying closer attention to than we typically do. A little bit of effort goes a long way to solidifying your deeper knowledge of the language.

Statements and expressions have analogs in English language – statements are like sentences and expressions are like phrases. Expressions can be pure/self-contained, or they can have side effects.

The JavaScript grammar layers semantic usage rules (aka context) on top of the pure syntax. For example, `{ }` pairs used in various places in your program can mean statement blocks, `object` literals, (ES6) destructuring assignments, or (ES6) named function arguments.

JavaScript operators all have well-defined rules for precedence (which ones bind first before others) and associativity (how multiple operator expressions are implicitly grouped). Once you learn these rules, it's up to you to decide if precedence/associativity are *too implicit* for their own good, or if they will aid in writing shorter, clearer code.

ASI (Automatic Semicolon Insertion) is a parser-error-correction mechanism built into the JS engine, which allows it under certain circumstances to insert an assumed `;` in places where it is required, was omitted, *and* where insertion fixes the parser error. The debate rages over whether this behavior implies that most `;` are optional (and can/should be omitted for cleaner code) or whether it means that omitting them is making mistakes that the JS engine merely cleans up for you.

JavaScript has several types of errors, but it's less known that it has two classifications for errors: “early” (compiler thrown, uncatchable) and “runtime” (`try..catchable`). All syntax errors are obviously early errors that stop the program before it runs, but there are others, too.

Function arguments have an interesting relationship to their formal declared named parameters. Specifically, the `arguments` array has a number of gotchas

of leaky abstraction behavior if you're not careful. Avoid `arguments` if you can, but if you must use it, by all means avoid using the positional slot in `arguments` at the same time as using a named parameter for that same argument.

The `finally` clause attached to a `try` (or `try..catch`) offers some very interesting quirks in terms of execution processing order. Some of these quirks can be helpful, but it's possible to create lots of confusion, especially if combined with labeled blocks. As always, use `finally` to make code better and clearer, not more clever or confusing.

The `switch` offers some nice shorthand for `if..else if..` statements, but beware of many common simplifying assumptions about its behavior. There are several quirks that can trip you up if you're not careful, but there's also some neat hidden tricks that `switch` has up its sleeve!

You Don't Know JS: Types & Grammar

Appendix A: Mixed Environment JavaScript

Beyond the core language mechanics we've fully explored in this book, there are several ways that your JS code can behave differently when it runs in the real world. If JS was executing purely inside an engine, it'd be entirely predictable based on nothing but the black-and-white of the spec. But JS pretty much always runs in the context of a hosting environment, which exposes your code to some degree of unpredictability.

For example, when your code runs alongside code from other sources, or when your code runs in different types of JS engines (not just browsers), there are some things that may behave differently.

We'll briefly explore some of these concerns.

Annex B (ECMAScript)

It's a little known fact that the official name of the language is ECMAScript (referring to the ECMA standards body that manages it). What then is "JavaScript"? JavaScript is the common trademark of the language, of course, but more appropriately, JavaScript is basically the browser implementation of the spec.

The official ECMAScript specification includes "Annex B," which discusses specific deviations from the official spec for the purposes of JS compatibility in browsers.

The proper way to consider these deviations is that they are only reliably present/valid if your code is running in a browser. If your code always runs in browsers, you won't see any observable difference. If not (like if it can run in node.js, Rhino, etc.), or you're not sure, tread carefully.

The main compatibility differences:

- Octal number literals are allowed, such as 0123 (decimal 83) in non-`strict mode`.
- `window.escape(..)` and `window.unescape(..)` allow you to escape or unescape strings with %-delimited hexadecimal escape sequences. For example: `window.escape("?foo=97%&bar=3%")` produces `"%3Ffoo%3D97%25%26bar%3D3%25"`.
- `String.prototype.substr` is quite similar to `String.prototype.substring`, except that instead of the second parameter being the ending index (non-inclusive), the second parameter is the `length` (number of characters to include).

Web ECMAScript

The Web ECMAScript specification (<http://javascript.spec.whatwg.org/>) covers the differences between the official ECMAScript specification and the current JavaScript implementations in browsers.

In other words, these items are “required” of browsers (to be compatible with each other) but are not (as of the time of writing) listed in the “Annex B” section of the official spec:

- `<!--` and `-->` are valid single-line comment delimiters.
- `String.prototype` additions for returning HTML-formatted strings: `anchor(..)`, `big(..)`, `blink(..)`, `bold(..)`, `fixed(..)`, `fontcolor(..)`, `fontsize(..)`, `italics(..)`, `link(..)`, `small(..)`, `strike(..)`, and `sub(..)`. **Note:** These are very rarely used in practice, and are generally discouraged in favor of other built-in DOM APIs or user-defined utilities.
- `RegExp` extensions: `RegExp.$1 .. RegExp.$9` (match-groups) and `RegExp.lastMatch/RegExp["$&"]` (most recent match).
- `Function.prototype` additions: `Function.prototype.arguments` (aliases internal `arguments` object) and `Function.caller` (aliases internal `arguments.caller`). **Note:** `arguments` and thus `arguments.caller` are deprecated, so you should avoid using them if possible. That goes doubly so for these aliases – don’t use them!

Note: Some other minor and rarely used deviations are not included in our list here. See the external “Annex B” and “Web ECMAScript” documents for more detailed information as needed.

Generally speaking, all these differences are rarely used, so the deviations from the specification are not significant concerns. **Just be careful** if you rely on any of them.

Host Objects

The well-covered rules for how variables behave in JS have exceptions to them when it comes to variables that are auto-defined, or otherwise created and provided to JS by the environment that hosts your code (browser, etc.) – so called, “host objects” (which include both built-in `objects` and `functions`).

For example:

```
var a = document.createElement( "div" );

typeof a;                                // "object" -- as expected
Object.prototype.toString.call( a );        // "[object HTMLDivElement]"

a.tagName;                                // "DIV"
```

`a` is not just an `object`, but a special host object because it’s a DOM element. It has a different internal `[[Class]]` value (`"HTMLDivElement"`) and comes with

predefined (and often unchangeable) properties.

Another such quirk has already been covered, in the “Falsy Objects” section in Chapter 4: some objects can exist but when coerced to `boolean` they (confoundingly) will coerce to `false` instead of the expected `true`.

Other behavior variations with host objects to be aware of can include:

- not having access to normal `object` built-ins like `toString()`
- not being overwritable
- having certain predefined read-only properties
- having methods that cannot be `this`-overridden to other objects
- and more...

Host objects are critical to making our JS code work with its surrounding environment. But it’s important to note when you’re interacting with a host object and be careful assuming its behaviors, as they will quite often not conform to regular JS objects.

One notable example of a host object that you probably interact with regularly is the `console` object and its various functions (`log(..)`, `error(..)`, etc.). The `console` object is provided by the *hosting environment* specifically so your code can interact with it for various development-related output tasks.

In browsers, `console` hooks up to the developer tools’ console display, whereas in `node.js` and other server-side JS environments, `console` is generally connected to the standard-output (`stdout`) and standard-error (`stderr`) streams of the JavaScript environment system process.

Global DOM Variables

You’re probably aware that declaring a variable in the global scope (with or without `var`) creates not only a global variable, but also its mirror: a property of the same name on the `global` object (`window` in the browser).

But what may be less common knowledge is that (because of legacy browser behavior) creating DOM elements with `id` attributes creates global variables of those same names. For example:

```
<div id="foo"></div>
```

And:

```
if (typeof foo == "undefined") {
  foo = 42;           // will never run
}

console.log( foo ); // HTML element
```

You’re perhaps used to managing global variable tests (using `typeof` or `.. in window` checks) under the assumption that only JS code creates such variables,

but as you can see, the contents of your hosting HTML page can also create them, which can easily throw off your existence check logic if you're not careful.

This is yet one more reason why you should, if at all possible, avoid using global variables, and if you have to, use variables with unique names that won't likely collide. But you also need to make sure not to collide with the HTML content as well as any other code.

Native Prototypes

One of the most widely known and classic pieces of JavaScript *best practice* wisdom is: **never extend native prototypes**.

Whatever method or property name you come up with to add to `Array.prototype` that doesn't (yet) exist, if it's a useful addition and well-designed, and properly named, there's a strong chance it *could* eventually end up being added to the spec – in which case your extension is now in conflict.

Here's a real example that actually happened to me that illustrates this point well.

I was building an embeddable widget for other websites, and my widget relied on jQuery (though pretty much any framework would have suffered this gotcha). It worked on almost every site, but we ran across one where it was totally broken.

After almost a week of analysis/debugging, I found that the site in question had, buried deep in one of its legacy files, code that looked like this:

```
// Netscape 4 doesn't have Array.push
Array.prototype.push = function(item) {
  this[this.length] = item;
};
```

Aside from the crazy comment (who cares about Netscape 4 anymore!?), this looks reasonable, right?

The problem is, `Array.prototype.push` was added to the spec sometime subsequent to this Netscape 4 era coding, but what was added is not compatible with this code. The standard `push(..)` allows multiple items to be pushed at once. This hacked one ignores the subsequent items.

Basically all JS frameworks have code that relies on `push(..)` with multiple elements. In my case, it was code around the CSS selector engine that was completely busted. But there could conceivably be dozens of other places susceptible.

The developer who originally wrote that `push(..)` hack had the right instinct to call it `push`, but didn't foresee pushing multiple elements. They were certainly acting in good faith, but they created a landmine that didn't go off until almost 10 years later when I unwittingly came along.

There's multiple lessons to take away on all sides.

First, don't extend the natives unless you're absolutely sure your code is the only code that will ever run in that environment. If you can't say that 100%, then extending the natives is dangerous. You must weigh the risks.

Next, don't unconditionally define extensions (because you can overwrite natives accidentally). In this particular example, had the code said this:

```
if (!Array.prototype.push) {
  // Netscape 4 doesn't have Array.push
  Array.prototype.push = function(item) {
    this[this.length] = item;
  };
}
```

The `if` statement guard would have only defined this hacked `push()` for JS environments where it didn't exist. In my case, that probably would have been OK. But even this approach is not without risk:

1. If the site's code (for some crazy reason!) was relying on a `push(..)` that ignored multiple items, that code would have been broken years ago when the standard `push(..)` was rolled out.
2. If any other library had come in and hacked in a `push(..)` ahead of this `if` guard, and it did so in an incompatible way, that would have broken the site at that time.

What that highlights is an interesting question that, frankly, doesn't get enough attention from JS developers: **Should you EVER rely on native built-in behavior** if your code is running in any environment where it's not the only code present?

The strict answer is **no**, but that's awfully impractical. Your code usually can't redefine its own private untouchable versions of all built-in behavior relied on. Even if you *could*, that's pretty wasteful.

So, should you feature-test for the built-in behavior as well as compliance-testing that it does what you expect? And what if that test fails – should your code just refuse to run?

```
// don't trust Array.prototype.push
(function(){
  if (Array.prototype.push) {
    var a = [];
    a.push(1,2);
    if (a[0] === 1 && a[1] === 2) {
      // tests passed, safe to use!
      return;
    }
  }
})
```

```
        throw Error(
            "Array#push() is missing/broken!"
        );
    })();
}
```

In theory, that sounds plausible, but it's also pretty impractical to design tests for every single built-in method.

So, what should we do? Should we *trust but verify* (feature- and compliance-test) **everything**? Should we just assume existence is compliance and let breakage (caused by others) bubble up as it will?

There's no great answer. The only fact that can be observed is that extending native prototypes is the only way these things bite you.

If you don't do it, and no one else does in the code in your application, you're safe. Otherwise, you should build in at least a little bit of skepticism, pessimism, and expectation of possible breakage.

Having a full set of unit/regression tests of your code that runs in all known environments is one way to surface some of these issues earlier, but it doesn't do anything to actually protect you from these conflicts.

Shims/Polyfills

It's usually said that the only safe place to extend a native is in an older (non-spec-compliant) environment, since that's unlikely to ever change – new browsers with new spec features replace older browsers rather than amending them.

If you could see into the future, and know for sure what a future standard was going to be, like for `Array.prototype.foobar`, it'd be totally safe to make your own compatible version of it to use now, right?

```
if (!Array.prototype.foobar) {
    // silly, silly
    Array.prototype.foobar = function() {
        this.push( "foo", "bar" );
    };
}
```

If there's already a spec for `Array.prototype.foobar`, and the specified behavior is equal to this logic, you're pretty safe in defining such a snippet, and in that case it's generally called a “polyfill” (or “shim”).

Such code is **very** useful to include in your code base to “patch” older browser environments that aren't updated to the newest specs. Using polyfills is a great way to create predictable code across all your supported environments.

Tip: ES5-Shim (<https://github.com/es-shims/es5-shim>) is a comprehensive collection of shims/polyfills for bringing a project up to ES5 baseline, and similarly, ES6-Shim (<https://github.com/es-shims/es6-shim>) provides shims for new APIs added as of ES6. While APIs can be shimmed/polyfilled, new syntax generally cannot. To bridge the syntactic divide, you'll want to also use an ES6-to-ES5 transpiler like Traceur (<https://github.com/google/traceur-compiler/wiki/Getting-Started>).

If there's likely a coming standard, and most discussions agree what it's going to be called and how it will operate, creating the ahead-of-time polyfill for future-facing standards compliance is called "prollyfill" (probably-fill).

The real catch is if some new standard behavior can't be (fully) polyfilled/prollyfilled.

There's debate in the community if a partial-polyfill for the common cases is acceptable (documenting the parts that cannot be polyfilled), or if a polyfill should be avoided if it purely can't be 100% compliant to the spec.

Many developers at least accept some common partial polyfills (like for instance `Object.create(..)`), because the parts that aren't covered are not parts they intend to use anyway.

Some developers believe that the `if` guard around a polyfill/shim should include some form of conformance test, replacing the existing method either if it's absent or fails the tests. This extra layer of compliance testing is sometimes used to distinguish "shim" (compliance tested) from "polyfill" (existence checked).

The only absolute take-away is that there is no absolute *right* answer here. Extending natives, even when done "safely" in older environments, is not 100% safe. The same goes for relying upon (possibly extended) natives in the presence of others' code.

Either should always be done with caution, defensive code, and lots of obvious documentation about the risks.

`<script>`s

Most browser-viewed websites/applications have more than one file that contains their code, and it's common to have a few or several `<script src=..></script>` elements in the page that load these files separately, and even a few inline-code `<script> .. </script>` elements as well.

But do these separate files/code snippets constitute separate programs or are they collectively one JS program?

The (perhaps surprising) reality is they act more like independent JS programs in most, but not all, respects.

The one thing they *share* is the single `global` object (`window` in the browser), which means multiple files can append their code to that shared namespace and

they can all interact.

So, if one `script` element defines a global function `foo()`, when a second `script` later runs, it can access and call `foo()` just as if it had defined the function itself.

But global variable scope *hoisting* (see the *Scope & Closures* title of this series) does not occur across these boundaries, so the following code would not work (because `foo()`'s declaration isn't yet declared), regardless of if they are (as shown) inline `<script> .. </script>` elements or externally loaded `<script src=..></script>` files:

```
<script>foo();</script>

<script>
  function foo() { .. }
</script>
```

But either of these *would* work instead:

```
<script>
  foo();
  function foo() { .. }
</script>
```

Or:

```
<script>
  function foo() { .. }
</script>

<script>foo();</script>
```

Also, if an error occurs in a `script` element (inline or external), as a separate standalone JS program it will fail and stop, but any subsequent `scripts` will run (still with the shared `global`) unimpeded.

You can create `script` elements dynamically from your code, and inject them into the DOM of the page, and the code in them will behave basically as if loaded normally in a separate file:

```
var greeting = "Hello World";

var el = document.createElement( "script" );

el.text = "function foo(){ alert( greeting );\
} setTimeout( foo, 1000 );";

document.body.appendChild( el );
```

Note: Of course, if you tried the above snippet but set `el.src` to some file URL instead of setting `el.text` to the code contents, you'd be dynamically creating

an externally loaded `<script src=..></script>` element.

One difference between code in an inline code block and that same code in an external file is that in the inline code block, the sequence of characters `</script>` cannot appear together, as (regardless of where it appears) it would be interpreted as the end of the code block. So, beware of code like:

```
<script>
  var code = "<script>alert( 'Hello World' )</script>";
</script>
```

It looks harmless, but the `</script>` appearing inside the `string` literal will terminate the script block abnormally, causing an error. The most common workaround is:

```
"</sc" + "ript>";
```

Also, beware that code inside an external file will be interpreted in the character set (UTF-8, ISO-8859-8, etc.) the file is served with (or the default), but that same code in an inline `script` element in your HTML page will be interpreted by the character set of the page (or its default).

Warning: The `charset` attribute will not work on inline script elements.

Another deprecated practice with inline `script` elements is including HTML-style or X(HT)ML-style comments around inline code, like:

```
<script>
<!--
  alert( "Hello" );
//-->
</script>

<script>
<!--//--><![CDATA[//><!--
  alert( "World" );
//--><!]]>
</script>
```

Both of these are totally unnecessary now, so if you're still doing that, stop it!

Note: Both `<!--` and `-->` (HTML-style comments) are actually specified as valid single-line comment delimiters (`var x = 2; <!-- valid comment and --> another valid line comment`) in JavaScript (see the “Web ECMAScript” section earlier), purely because of this old technique. But never use them.

Reserved Words

The ES5 spec defines a set of “reserved words” in Section 7.6.1 that cannot be used as standalone variable names. Technically, there are four categories:

“keywords”, “future reserved words”, the `null` literal, and the `true` / `false` boolean literals.

Keywords are the obvious ones like `function` and `switch`. Future reserved words include things like `enum`, though many of the rest of them (`class`, `extends`, etc.) are all now actually used by ES6; there are other strict-mode only reserved words like `interface`.

StackOverflow user “art4theSould” creatively worked all these reserved words into a fun little poem (<http://stackoverflow.com/questions/26255/reserved-keywords-in-javascript/12114140#12114140>):

Let this long package float, Goto private class if short. While protected with debugger case, Continue volatile interface. Instanceof super synchronized throw, Extends final export throws.

Try import double enum? - False, boolean, abstract function, Implements typeof transient break! Void static, default do, Switch int native new. Else, delete null public var In return for const, true, char ...Finally catch byte.

Note: This poem includes words that were reserved in ES3 (`byte`, `long`, etc.) that are no longer reserved as of ES5.

Prior to ES5, the reserved words also could not be property names or keys in object literals, but that restriction no longer exists.

So, this is not allowed:

```
var import = "42";
```

But this is allowed:

```
var obj = { import: "42" };
console.log( obj.import );
```

You should be aware though that some older browser versions (mainly older IE) weren’t completely consistent on applying these rules, so there are places where using reserved words in object property name locations can still cause issues. Carefully test all supported browser environments.

Implementation Limits

The JavaScript spec does not place arbitrary limits on things such as the number of arguments to a function or the length of a string literal, but these limits exist nonetheless, because of implementation details in different engines.

For example:

```
function addAll() {
  var sum = 0;
  for (var i=0; i < arguments.length; i++) {
```

```

        sum += arguments[i];
    }
    return sum;
}

var nums = [];
for (var i=1; i < 100000; i++) {
    nums.push(i);
}

addAll( 2, 4, 6 );           // 12
addAll.apply( null, nums );  // should be: 499950000

```

In some JS engines, you'll get the correct 499950000 answer, but in others (like Safari 6.x), you'll get the error: "RangeError: Maximum call stack size exceeded."

Examples of other limits known to exist:

- maximum number of characters allowed in a string literal (not just a string value)
- size (bytes) of data that can be sent in arguments to a function call (aka stack size)
- number of parameters in a function declaration
- maximum depth of non-optimized call stack (i.e., with recursion): how long a chain of function calls from one to the other can be
- number of seconds a JS program can run continuously blocking the browser
- maximum length allowed for a variable name
- ...

It's not very common at all to run into these limits, but you should be aware that limits can and do exist, and importantly that they vary between engines.

Review

We know and can rely upon the fact that the JS language itself has one standard and is predictably implemented by all the modern browsers/engines. This is a very good thing!

But JavaScript rarely runs in isolation. It runs in an environment mixed in with code from third-party libraries, and sometimes it even runs in engines/environments that differ from those found in browsers.

Paying close attention to these issues improves the reliability and robustness of your code.

You Don't Know JS: Async & Performance

Foreword

Over the years, my employer has trusted me enough to conduct interviews. If we're looking for someone with skills in JavaScript, my first line of questioning... actually that's not true, I first check if the candidate needs the bathroom and/or a drink, because comfort is important, but once I'm past the bit about the candidate's fluid in/out-take, I set about determining if the candidate knows JavaScript, or just jQuery.

Not that there's anything wrong with jQuery. It lets you do a lot without really knowing JavaScript, and that's a feature not a bug. But if the job calls for advanced skills in JavaScript performance and maintainability, you need someone who knows how libraries such as jQuery are put together. You need to be able to harness the core of JavaScript the same way they do.

If I want to get a picture of someone's core JavaScript skill, I'm most interested in what they make of closures (you've read that book of this series already, right?) and how to get the most out of asynchronicity, which brings us to this book.

For starters, you'll be taken through callbacks, the bread and butter of asynchronous programming. Of course, bread and butter does not make for a particularly satisfying meal, but the next course is full of tasty tasty promises!

If you don't know promises, now is the time to learn. Promises are now the official way to provide async return values in both JavaScript and the DOM. All future async DOM APIs will use them, many already do, so be prepared! At the time of writing, Promises have shipped in most major browsers, with IE shipping soon. Once you've finished that, I hope you left room for the next course, Generators.

Generators snuck their way into stable versions of Chrome and Firefox without too much pomp and ceremony, because, frankly, they're more complicated than they are interesting. Or, that's what I thought until I saw them combined with promises. There, they become an important tool in readability and maintenance.

For dessert, well, I won't spoil the surprise, but prepare to gaze into the future of JavaScript! Features that give you more and more control over concurrency and asynchronicity.

Well, I won't block your enjoyment of the book any longer, on with the show! If you've already read part of the book before reading this Foreword, give yourself 10 asynchronous points! You deserve them!

Jake Archibald jakearchibald.com, [@jaffathecake](https://twitter.com/jaffathecake) Developer Advocate at Google Chrome

You Don't Know JS

Preface

I'm sure you noticed, but "JS" in the book series title is not an abbreviation for words used to curse about JavaScript, though cursing at the language's quirks is something we can probably all identify with!

From the earliest days of the web, JavaScript has been a foundational technology that drives interactive experience around the content we consume. While flickering mouse trails and annoying pop-up prompts may be where JavaScript started, nearly 2 decades later, the technology and capability of JavaScript has grown many orders of magnitude, and few doubt its importance at the heart of the world's most widely available software platform: the web.

But as a language, it has perpetually been a target for a great deal of criticism, owing partly to its heritage but even more to its design philosophy. Even the name evokes, as Brendan Eich once put it, "dumb kid brother" status next to its more mature older brother "Java". But the name is merely an accident of politics and marketing. The two languages are vastly different in many important ways. "JavaScript" is as related to "Java" as "Carnival" is to "Car".

Because JavaScript borrows concepts and syntax idioms from several languages, including proud C-style procedural roots as well as subtle, less obvious Scheme/Lisp-style functional roots, it is exceedingly approachable to a broad audience of developers, even those with just little to no programming experience. The "Hello World" of JavaScript is so simple that the language is inviting and easy to get comfortable with in early exposure.

While JavaScript is perhaps one of the easiest languages to get up and running with, its eccentricities make solid mastery of the language a vastly less common occurrence than in many other languages. Where it takes a pretty in-depth knowledge of a language like C or C++ to write a full-scale program, full-scale production JavaScript can, and often does, barely scratch the surface of what the language can do.

Sophisticated concepts which are deeply rooted into the language tend instead to surface themselves in *seemingly* simplistic ways, such as passing around functions as callbacks, which encourages the JavaScript developer to just use the language as-is and not worry too much about what's going on under the hood.

It is simultaneously a simple, easy-to-use language that has broad appeal, and a complex and nuanced collection of language mechanics which without careful study will elude *true understanding* even for the most seasoned of JavaScript developers.

Therein lies the paradox of JavaScript, the Achilles' Heel of the language, the challenge we are presently addressing. Because JavaScript *can* be used without understanding, the understanding of the language is often never attained.

Mission

If at every point that you encounter a surprise or frustration in JavaScript, your response is to add it to the blacklist, as some are accustomed to doing, you soon will be relegated to a hollow shell of the richness of JavaScript.

While this subset has been famously dubbed “The Good Parts”, I would implore you, dear reader, to instead consider it the “The Easy Parts”, “The Safe Parts”, or even “The Incomplete Parts”.

This *You Don’t Know JavaScript* book series offers a contrary challenge: learn and deeply understand *all* of JavaScript, even and especially “The Tough Parts”.

Here, we address head on the tendency of JS developers to learn “just enough” to get by, without ever forcing themselves to learn exactly how and why the language behaves the way it does. Furthermore, we eschew the common advice to *retreat* when the road gets rough.

I am not content, nor should you be, at stopping once something *just works*, and not really knowing *why*. I gently challenge you to journey down that bumpy “road less traveled” and embrace all that JavaScript is and can do. With that knowledge, no technique, no framework, no popular buzzword acronym of the week, will be beyond your understanding.

These books each take on specific core parts of the language which are most commonly misunderstood or under-understood, and dive very deep and exhaustively into them. You should come away from reading with a firm confidence in your understanding, not just of the theoretical, but the practical “what you need to know” bits.

The JavaScript you know *right now* is probably *parts* handed down to you by others who’ve been burned by incomplete understanding. *That* JavaScript is but a shadow of the true language. You don’t *really* know JavaScript, *yet*, but if you dig into this series, you *will*. Read on, my friends. JavaScript awaits you.

Summary

JavaScript is awesome. It’s easy to learn partially, and much harder to learn completely (or even *sufficiently*). When developers encounter confusion, they usually blame the language instead of their lack of understanding. These books aim to fix that, inspiring a strong appreciation for the language you can now, and *should*, deeply *know*.

Note: Many of the examples in this book assume modern (and future-reaching) JavaScript engine environments, such as ES6. Some code may not work as described if run in older (pre-ES6) engines.

You Don't Know JS: Async & Performance

Chapter 1: Asynchrony: Now & Later

One of the most important and yet often misunderstood parts of programming in a language like JavaScript is how to express and manipulate program behavior spread out over a period of time.

This is not just about what happens from the beginning of a `for` loop to the end of a `for` loop, which of course takes *some time* (microseconds to milliseconds) to complete. It's about what happens when part of your program runs *now*, and another part of your program runs *later* – there's a gap between *now* and *later* where your program isn't actively executing.

Practically all nontrivial programs ever written (especially in JS) have in some way or another had to manage this gap, whether that be in waiting for user input, requesting data from a database or file system, sending data across the network and waiting for a response, or performing a repeated task at a fixed interval of time (like animation). In all these various ways, your program has to manage state across the gap in time. As they famously say in London (of the chasm between the subway door and the platform): “mind the gap.”

In fact, the relationship between the *now* and *later* parts of your program is at the heart of asynchronous programming.

Asynchronous programming has been around since the beginning of JS, for sure. But most JS developers have never really carefully considered exactly how and why it crops up in their programs, or explored various *other* ways to handle it. The *good enough* approach has always been the humble callback function. Many to this day will insist that callbacks are more than sufficient.

But as JS continues to grow in both scope and complexity, to meet the ever-widening demands of a first-class programming language that runs in browsers and servers and every conceivable device in between, the pains by which we manage asynchrony are becoming increasingly crippling, and they cry out for approaches that are both more capable and more reason-able.

While this all may seem rather abstract right now, I assure you we'll tackle it more completely and concretely as we go on through this book. We'll explore a variety of emerging techniques for async JavaScript programming over the next several chapters.

But before we can get there, we're going to have to understand much more deeply what asynchrony is and how it operates in JS.

A Program in Chunks

You may write your JS program in one `.js` file, but your program is almost certainly comprised of several chunks, only one of which is going to execute

now, and the rest of which will execute *later*. The most common unit of *chunk* is the **function**.

The problem most developers new to JS seem to have is that *later* doesn't happen strictly and immediately after *now*. In other words, tasks that cannot complete *now* are, by definition, going to complete asynchronously, and thus we will not have blocking behavior as you might intuitively expect or want.

Consider:

```
// ajax(..) is some arbitrary Ajax function given by a library
var data = ajax( "http://some.url.1" );

console.log( data );
// Oops! `data` generally won't have the Ajax results
```

You're probably aware that standard Ajax requests don't complete synchronously, which means the `ajax(..)` function does not yet have any value to return back to be assigned to `data` variable. If `ajax(..)` *could* block until the response came back, then the `data = ..` assignment would work fine.

But that's not how we do Ajax. We make an asynchronous Ajax request *now*, and we won't get the results back until *later*.

The simplest (but definitely not only, or necessarily even best!) way of "waiting" from *now* until *later* is to use a function, commonly called a **callback function**:

```
// ajax(..) is some arbitrary Ajax function given by a library
ajax( "http://some.url.1", function myCallbackFunction(data){

    console.log( data ); // Yay, I gots me some `data`!

} );
```

Warning: You may have heard that it's possible to make synchronous Ajax requests. While that's technically true, you should never, ever do it, under any circumstances, because it locks the browser UI (buttons, menus, scrolling, etc.) and prevents any user interaction whatsoever. This is a terrible idea, and should always be avoided.

Before you protest in disagreement, no, your desire to avoid the mess of callbacks is *not* justification for blocking, synchronous Ajax.

For example, consider this code:

```
function now() {
    return 21;
}

function later() {
    answer = answer * 2;
```

```

        console.log( "Meaning of life:", answer );
    }

var answer = now();

setTimeout( later, 1000 ); // Meaning of life: 42

```

There are two chunks to this program: the stuff that will run *now*, and the stuff that will run *later*. It should be fairly obvious what those two chunks are, but let's be super explicit:

Now:

```

function now() {
    return 21;
}

function later() { .. }

var answer = now();

setTimeout( later, 1000 );

```

Later:

```

answer = answer * 2;
console.log( "Meaning of life:", answer );

```

The *now* chunk runs right away, as soon as you execute your program. But `setTimeout(..)` also sets up an event (a timeout) to happen *later*, so the contents of the `later()` function will be executed at a later time (1,000 milliseconds from now).

Any time you wrap a portion of code into a `function` and specify that it should be executed in response to some event (timer, mouse click, Ajax response, etc.), you are creating a *later* chunk of your code, and thus introducing asynchrony to your program.

Async Console

There is no specification or set of requirements around how the `console.*` methods work – they are not officially part of JavaScript, but are instead added to JS by the *hosting environment* (see the *Types & Grammar* title of this book series).

So, different browsers and JS environments do as they please, which can sometimes lead to confusing behavior.

In particular, there are some browsers and some conditions that `console.log(..)` does not actually immediately output what it's given. The main reason this

may happen is because I/O is a very slow and blocking part of many programs (not just JS). So, it may perform better (from the page/UI perspective) for a browser to handle `console` I/O asynchronously in the background, without you perhaps even knowing that occurred.

A not terribly common, but possible, scenario where this could be *observable* (not from code itself but from the outside):

```
var a = {  
  index: 1  
};  
  
// later  
console.log( a ); // ??  
  
// even later  
a.index++;
```

We'd normally expect to see the `a` object be snapshotted at the exact moment of the `console.log(..)` statement, printing something like `{ index: 1 }`, such that in the next statement when `a.index++` happens, it's modifying something different than, or just strictly after, the output of `a`.

Most of the time, the preceding code will probably produce an object representation in your developer tools' console that's what you'd expect. But it's possible this same code could run in a situation where the browser felt it needed to defer the `console` I/O to the background, in which case it's *possible* that by the time the object is represented in the browser console, the `a.index++` has already happened, and it shows `{ index: 2 }`.

It's a moving target under what conditions exactly `console` I/O will be deferred, or even whether it will be observable. Just be aware of this possible asynchronicity in I/O in case you ever run into issues in debugging where objects have been modified *after* a `console.log(..)` statement and yet you see the unexpected modifications show up.

Note: If you run into this rare scenario, the best option is to use breakpoints in your JS debugger instead of relying on `console` output. The next best option would be to force a "snapshot" of the object in question by serializing it to a `string`, like with `JSON.stringify(..)`.

Event Loop

Let's make a (perhaps shocking) claim: despite clearly allowing asynchronous JS code (like the timeout we just looked at), up until recently (ES6), JavaScript itself has actually never had any direct notion of asynchrony built into it.

What!? That seems like a crazy claim, right? In fact, it's quite true. The JS engine itself has never done anything more than execute a single chunk of your

program at any given moment, when asked to.

“Asked to.” By whom? That’s the important part!

The JS engine doesn’t run in isolation. It runs inside a *hosting environment*, which is for most developers the typical web browser. Over the last several years (but by no means exclusively), JS has expanded beyond the browser into other environments, such as servers, via things like Node.js. In fact, JavaScript gets embedded into all kinds of devices these days, from robots to lightbulbs.

But the one common “thread” (that’s a not-so-subtle asynchronous joke, for what it’s worth) of all these environments is that they have a mechanism in them that handles executing multiple chunks of your program *over time*, at each moment invoking the JS engine, called the “event loop.”

In other words, the JS engine has had no innate sense of *time*, but has instead been an on-demand execution environment for any arbitrary snippet of JS. It’s the surrounding environment that has always *scheduled* “events” (JS code executions).

So, for example, when your JS program makes an Ajax request to fetch some data from a server, you set up the “response” code in a function (commonly called a “callback”), and the JS engine tells the hosting environment, “Hey, I’m going to suspend execution for now, but whenever you finish with that network request, and you have some data, please *call* this function *back*.”

The browser is then set up to listen for the response from the network, and when it has something to give you, it schedules the callback function to be executed by inserting it into the *event loop*.

So what is the *event loop*?

Let’s conceptualize it first through some fake-ish code:

```
// `eventLoop` is an array that acts as a queue (first-in, first-out)
var eventLoop = [ ];
var event;

// keep going "forever"
while (true) {
    // perform a "tick"
    if (eventLoop.length > 0) {
        // get the next event in the queue
        event = eventLoop.shift();

        // now, execute the next event
        try {
            event();
        }
        catch (err) {
```

```
        reportError(err);
    }
}
}
```

This is, of course, vastly simplified pseudocode to illustrate the concepts. But it should be enough to help get a better understanding.

As you can see, there's a continuously running loop represented by the `while` loop, and each iteration of this loop is called a “tick.” For each tick, if an event is waiting on the queue, it's taken off and executed. These events are your function callbacks.

It's important to note that `setTimeout(..)` doesn't put your callback on the event loop queue. What it does is set up a timer; when the timer expires, the environment places your callback into the event loop, such that some future tick will pick it up and execute it.

What if there are already 20 items in the event loop at that moment? Your callback waits. It gets in line behind the others – there's not normally a path for preempting the queue and skipping ahead in line. This explains why `setTimeout(..)` timers may not fire with perfect temporal accuracy. You're guaranteed (roughly speaking) that your callback won't fire *before* the time interval you specify, but it can happen at or after that time, depending on the state of the event queue.

So, in other words, your program is generally broken up into lots of small chunks, which happen one after the other in the event loop queue. And technically, other events not related directly to your program can be interleaved within the queue as well.

Note: We mentioned “up until recently” in relation to ES6 changing the nature of where the event loop queue is managed. It's mostly a formal technicality, but ES6 now specifies how the event loop works, which means technically it's within the purview of the JS engine, rather than just the *hosting environment*. One main reason for this change is the introduction of ES6 Promises, which we'll discuss in Chapter 3, because they require the ability to have direct, fine-grained control over scheduling operations on the event loop queue (see the discussion of `setTimeout(..)` in the “Cooperation” section).

Parallel Threading

It's very common to conflate the terms “async” and “parallel,” but they are actually quite different. Remember, `async` is about the gap between *now* and *later*. But `parallel` is about things being able to occur simultaneously.

The most common tools for parallel computing are processes and threads. Processes and threads execute independently and may execute simultaneously: on

separate processors, or even separate computers, but multiple threads can share the memory of a single process.

An event loop, by contrast, breaks its work into tasks and executes them in serial, disallowing parallel access and changes to shared memory. Parallelism and “serialism” can coexist in the form of cooperating event loops in separate threads.

The interleaving of parallel threads of execution and the interleaving of asynchronous events occur at very different levels of granularity.

For example:

```
function later() {
    answer = answer * 2;
    console.log( "Meaning of life:", answer );
}
```

While the entire contents of `later()` would be regarded as a single event loop queue entry, when thinking about a thread this code would run on, there's actually perhaps a dozen different low-level operations. For example, `answer = answer * 2` requires first loading the current value of `answer`, then putting 2 somewhere, then performing the multiplication, then taking the result and storing it back into `answer`.

In a single-threaded environment, it really doesn't matter that the items in the thread queue are low-level operations, because nothing can interrupt the thread. But if you have a parallel system, where two different threads are operating in the same program, you could very likely have unpredictable behavior.

Consider:

```
var a = 20;

function foo() {
    a = a + 1;
}

function bar() {
    a = a * 2;
}

// ajax(..) is some arbitrary Ajax function given by a library
ajax( "http://some.url.1", foo );
ajax( "http://some.url.2", bar );
```

In JavaScript's single-threaded behavior, if `foo()` runs before `bar()`, the result is that `a` has 42, but if `bar()` runs before `foo()` the result in `a` will be 41.

If JS events sharing the same data executed in parallel, though, the problems would be much more subtle. Consider these two lists of pseudocode tasks as the

threads that could respectively run the code in `foo()` and `bar()`, and consider what happens if they are running at exactly the same time:

Thread 1 (X and Y are temporary memory locations):

```
foo():
  a. load value of `a` in `X`
  b. store `1` in `Y`
  c. add `X` and `Y`, store result in `X`
  d. store value of `X` in `a`
```

Thread 2 (X and Y are temporary memory locations):

```
bar():
  a. load value of `a` in `X`
  b. store `2` in `Y`
  c. multiply `X` and `Y`, store result in `X`
  d. store value of `X` in `a`
```

Now, let's say that the two threads are running truly in parallel. You can probably spot the problem, right? They use shared memory locations X and Y for their temporary steps.

What's the end result in `a` if the steps happen like this?

```
1a (load value of `a` in `X`    ==> `20`)
2a (load value of `a` in `X`    ==> `20`)
1b (store `1` in `Y`    ==> `1`)
2b (store `2` in `Y`    ==> `2`)
1c (add `X` and `Y`, store result in `X`    ==> `22`)
1d (store value of `X` in `a`    ==> `22`)
2c (multiply `X` and `Y`, store result in `X`    ==> `44`)
2d (store value of `X` in `a`    ==> `44`)
```

The result in `a` will be 44. But what about this ordering?

```
1a (load value of `a` in `X`    ==> `20`)
2a (load value of `a` in `X`    ==> `20`)
2b (store `2` in `Y`    ==> `2`)
1b (store `1` in `Y`    ==> `1`)
2c (multiply `X` and `Y`, store result in `X`    ==> `20`)
1c (add `X` and `Y`, store result in `X`    ==> `21`)
1d (store value of `X` in `a`    ==> `21`)
2d (store value of `X` in `a`    ==> `21`)
```

The result in `a` will be 21.

So, threaded programming is very tricky, because if you don't take special steps to prevent this kind of interruption/interleaving from happening, you can get very surprising, nondeterministic behavior that frequently leads to headaches.

JavaScript never shares data across threads, which means *that* level of nondeterminism isn't a concern. But that doesn't mean JS is always deterministic. Remember earlier, where the relative ordering of `foo()` and `bar()` produces two different results (41 or 42)?

Note: It may not be obvious yet, but not all nondeterminism is bad. Sometimes it's irrelevant, and sometimes it's intentional. We'll see more examples of that throughout this and the next few chapters.

Run-to-Completion

Because of JavaScript's single-threading, the code inside of `foo()` (and `bar()`) is atomic, which means that once `foo()` starts running, the entirety of its code will finish before any of the code in `bar()` can run, or vice versa. This is called "run-to-completion" behavior.

In fact, the run-to-completion semantics are more obvious when `foo()` and `bar()` have more code in them, such as:

```
var a = 1;
var b = 2;

function foo() {
    a++;
    b = b * a;
    a = b + 3;
}

function bar() {
    b--;
    a = 8 + b;
    b = a * 2;
}

// ajax(..) is some arbitrary Ajax function given by a library
ajax( "http://some.url.1", foo );
ajax( "http://some.url.2", bar );
```

Because `foo()` can't be interrupted by `bar()`, and `bar()` can't be interrupted by `foo()`, this program only has two possible outcomes depending on which starts running first – if threading were present, and the individual statements in `foo()` and `bar()` could be interleaved, the number of possible outcomes would be greatly increased!

Chunk 1 is synchronous (happens *now*), but chunks 2 and 3 are asynchronous (happen *later*), which means their execution will be separated by a gap of time.

Chunk 1:

```
var a = 1;  
var b = 2;
```

Chunk 2 (`foo()`):

```
a++;  
b = b * a;  
a = b + 3;
```

Chunk 3 (`bar()`):

```
b--;  
a = 8 + b;  
b = a * 2;
```

Chunks 2 and 3 may happen in either-first order, so there are two possible outcomes for this program, as illustrated here:

Outcome 1:

```
var a = 1;  
var b = 2;
```

```
// foo()  
a++;  
b = b * a;  
a = b + 3;
```

```
// bar()  
b--;  
a = 8 + b;  
b = a * 2;
```

```
a; // 11  
b; // 22
```

Outcome 2:

```
var a = 1;  
var b = 2;
```

```
// bar()  
b--;  
a = 8 + b;  
b = a * 2;
```

```
// foo()  
a++;  
b = b * a;  
a = b + 3;
```

```
a; // 183
b; // 180
```

Two outcomes from the same code means we still have nondeterminism! But it's at the function (event) ordering level, rather than at the statement ordering level (or, in fact, the expression operation ordering level) as it is with threads. In other words, it's *more deterministic* than threads would have been.

As applied to JavaScript's behavior, this function-ordering nondeterminism is the common term "race condition," as `foo()` and `bar()` are racing against each other to see which runs first. Specifically, it's a "race condition" because you cannot predict reliably how `a` and `b` will turn out.

Note: If there was a function in JS that somehow did not have run-to-completion behavior, we could have many more possible outcomes, right? It turns out ES6 introduces just such a thing (see Chapter 4 "Generators"), but don't worry right now, we'll come back to that!

Concurrency

Let's imagine a site that displays a list of status updates (like a social network news feed) that progressively loads as the user scrolls down the list. To make such a feature work correctly, (at least) two separate "processes" will need to be executing *simultaneously* (i.e., during the same window of time, but not necessarily at the same instant).

Note: We're using "process" in quotes here because they aren't true operating system-level processes in the computer science sense. They're virtual processes, or tasks, that represent a logically connected, sequential series of operations. We'll simply prefer "process" over "task" because terminology-wise, it will match the definitions of the concepts we're exploring.

The first "process" will respond to `onscroll` events (making Ajax requests for new content) as they fire when the user has scrolled the page further down. The second "process" will receive Ajax responses back (to render content onto the page).

Obviously, if a user scrolls fast enough, you may see two or more `onscroll` events fired during the time it takes to get the first response back and process, and thus you're going to have `onscroll` events and Ajax response events firing rapidly, interleaved with each other.

Concurrency is when two or more "processes" are executing simultaneously over the same period, regardless of whether their individual constituent operations happen *in parallel* (at the same instant on separate processors or cores) or not. You can think of concurrency then as "process"-level (or task-level) parallelism, as opposed to operation-level parallelism (separate-processor threads).

Note: Concurrency also introduces an optional notion of these “processes” interacting with each other. We’ll come back to that later.

For a given window of time (a few seconds worth of a user scrolling), let’s visualize each independent “process” as a series of events/operations:

“Process” 1 (onscroll events):

```
onscroll, request 1
onscroll, request 2
onscroll, request 3
onscroll, request 4
onscroll, request 5
onscroll, request 6
onscroll, request 7
```

“Process” 2 (Ajax response events):

```
response 1
response 2
response 3
response 4
response 5
response 6
response 7
```

It’s quite possible that an `onscroll` event and an Ajax response event could be ready to be processed at exactly the same *moment*. For example, let’s visualize these events in a timeline:

```
onscroll, request 1
onscroll, request 2      response 1
onscroll, request 3      response 2
response 3
onscroll, request 4
onscroll, request 5
onscroll, request 6      response 4
onscroll, request 7
response 6
response 5
response 7
```

But, going back to our notion of the event loop from earlier in the chapter, JS is only going to be able to handle one event at a time, so either `onscroll, request 2` is going to happen first or `response 1` is going to happen first, but they cannot happen at literally the same moment. Just like kids at a school cafeteria, no matter what crowd they form outside the doors, they’ll have to merge into a single line to get their lunch!

Let’s visualize the interleaving of all these events onto the event loop queue.

Event Loop Queue:

```
onscroll, request 1      <--- Process 1 starts
onscroll, request 2
response 1              <--- Process 2 starts
onscroll, request 3
response 2
response 3
onscroll, request 4
onscroll, request 5
onscroll, request 6
response 4
onscroll, request 7    <--- Process 1 finishes
response 6
response 5
response 7              <--- Process 2 finishes
```

“Process 1” and “Process 2” run concurrently (task-level parallel), but their individual events run sequentially on the event loop queue.

By the way, notice how `response 6` and `response 5` came back out of expected order?

The single-threaded event loop is one expression of concurrency (there are certainly others, which we'll come back to later).

Noninteracting

As two or more “processes” are interleaving their steps/events concurrently within the same program, they don't necessarily need to interact with each other if the tasks are unrelated. **If they don't interact, nondeterminism is perfectly acceptable.**

For example:

```
var res = {};  
  
function foo(results) {  
    res.foo = results;  
}  
  
function bar(results) {  
    res.bar = results;  
}  
  
// ajax(..) is some arbitrary Ajax function given by a library  
ajax( "http://some.url.1", foo );  
ajax( "http://some.url.2", bar );
```

`foo()` and `bar()` are two concurrent “processes,” and it’s nondeterminate which order they will be fired in. But we’ve constructed the program so it doesn’t matter what order they fire in, because they act independently and as such don’t need to interact.

This is not a “race condition” bug, as the code will always work correctly, regardless of the ordering.

Interaction

More commonly, concurrent “processes” will by necessity interact, indirectly through scope and/or the DOM. When such interaction will occur, you need to coordinate these interactions to prevent “race conditions,” as described earlier.

Here’s a simple example of two concurrent “processes” that interact because of implied ordering, which is only *sometimes broken*:

```
var res = [];  
  
function response(data) {  
    res.push( data );  
}  
  
// ajax(..) is some arbitrary Ajax function given by a library  
ajax( "http://some.url.1", response );  
ajax( "http://some.url.2", response );
```

The concurrent “processes” are the two `response()` calls that will be made to handle the Ajax responses. They can happen in either-first order.

Let’s assume the expected behavior is that `res[0]` has the results of the `"http://some.url.1"` call, and `res[1]` has the results of the `"http://some.url.2"` call. Sometimes that will be the case, but sometimes they’ll be flipped, depending on which call finishes first. There’s a pretty good likelihood that this nondeterminism is a “race condition” bug.

Note: Be extremely wary of assumptions you might tend to make in these situations. For example, it’s not uncommon for a developer to observe that `"http://some.url.2"` is “always” much slower to respond than `"http://some.url.1"`, perhaps by virtue of what tasks they’re doing (e.g., one performing a database task and the other just fetching a static file), so the observed ordering seems to always be as expected. Even if both requests go to the same server, and *it* intentionally responds in a certain order, there’s no *real* guarantee of what order the responses will arrive back in the browser.

So, to address such a race condition, you can coordinate ordering interaction:

```
var res = [];  
  
function response(data) {
```

```

    if (data.url == "http://some.url.1") {
        res[0] = data;
    }
    else if (data.url == "http://some.url.2") {
        res[1] = data;
    }
}

// ajax(..) is some arbitrary Ajax function given by a library
ajax( "http://some.url.1", response );
ajax( "http://some.url.2", response );

```

Regardless of which Ajax response comes back first, we inspect the `data.url` (assuming one is returned from the server, of course!) to figure out which position the response data should occupy in the `res` array. `res[0]` will always hold the `"http://some.url.1"` results and `res[1]` will always hold the `"http://some.url.2"` results. Through simple coordination, we eliminated the “race condition” nondeterminism.

The same reasoning from this scenario would apply if multiple concurrent function calls were interacting with each other through the shared DOM, like one updating the contents of a `<div>` and the other updating the style or attributes of the `<div>` (e.g., to make the DOM element visible once it has content). You probably wouldn’t want to show the DOM element before it had content, so the coordination must ensure proper ordering interaction.

Some concurrency scenarios are *always broken* (not just *sometimes*) without coordinated interaction. Consider:

```

var a, b;

function foo(x) {
    a = x * 2;
    baz();
}

function bar(y) {
    b = y * 2;
    baz();
}

function baz() {
    console.log(a + b);
}

// ajax(..) is some arbitrary Ajax function given by a library
ajax( "http://some.url.1", foo );
ajax( "http://some.url.2", bar );

```

In this example, whether `foo()` or `bar()` fires first, it will always cause `baz()` to run too early (either `a` or `b` will still be `undefined`), but the second invocation of `baz()` will work, as both `a` and `b` will be available.

There are different ways to address such a condition. Here's one simple way:

```
var a, b;

function foo(x) {
  a = x * 2;
  if (a && b) {
    baz();
  }
}

function bar(y) {
  b = y * 2;
  if (a && b) {
    baz();
  }
}

function baz() {
  console.log( a + b );
}

// ajax(..) is some arbitrary Ajax function given by a library
ajax( "http://some.url.1", foo );
ajax( "http://some.url.2", bar );
```

The `if (a && b)` conditional around the `baz()` call is traditionally called a “gate,” because we’re not sure what order `a` and `b` will arrive, but we wait for both of them to get there before we proceed to open the gate (call `baz()`).

Another concurrency interaction condition you may run into is sometimes called a “race,” but more correctly called a “latch.” It’s characterized by “only the first one wins” behavior. Here, nondeterminism is acceptable, in that you are explicitly saying it’s OK for the “race” to the finish line to have only one winner.

Consider this broken code:

```
var a;

function foo(x) {
  a = x * 2;
  baz();
}

function bar(x) {
```

```

    a = x / 2;
    baz();
}

function baz() {
    console.log( a );
}

// ajax(..) is some arbitrary Ajax function given by a library
ajax( "http://some.url.1", foo );
ajax( "http://some.url.2", bar );

```

Whichever one (`foo()` or `bar()`) fires last will not only overwrite the assigned `a` value from the other, but it will also duplicate the call to `baz()` (likely undesired).

So, we can coordinate the interaction with a simple latch, to let only the first one through:

```

var a;

function foo(x) {
    if (a == undefined) {
        a = x * 2;
        baz();
    }
}

function bar(x) {
    if (a == undefined) {
        a = x / 2;
        baz();
    }
}

function baz() {
    console.log( a );
}

// ajax(..) is some arbitrary Ajax function given by a library
ajax( "http://some.url.1", foo );
ajax( "http://some.url.2", bar );

```

The `if (a == undefined)` conditional allows only the first of `foo()` or `bar()` through, and the second (and indeed any subsequent) calls would just be ignored. There's just no virtue in coming in second place!

Note: In all these scenarios, we've been using global variables for simplistic

illustration purposes, but there's nothing about our reasoning here that requires it. As long as the functions in question can access the variables (via scope), they'll work as intended. Relying on lexically scoped variables (see the *Scope & Closures* title of this book series), and in fact global variables as in these examples, is one obvious downside to these forms of concurrency coordination. As we go through the next few chapters, we'll see other ways of coordination that are much cleaner in that respect.

Cooperation

Another expression of concurrency coordination is called “cooperative concurrency.” Here, the focus isn't so much on interacting via value sharing in scopes (though that's obviously still allowed!). The goal is to take a long-running “process” and break it up into steps or batches so that other concurrent “processes” have a chance to interleave their operations into the event loop queue.

For example, consider an Ajax response handler that needs to run through a long list of results to transform the values. We'll use `Array#map(..)` to keep the code shorter:

```
var res = [];  
  
// `response(..)` receives array of results from the Ajax call  
function response(data) {  
    // add onto existing `res` array  
    res = res.concat(  
        // make a new transformed array with all `data` values doubled  
        data.map( function(val){  
            return val * 2;  
        } )  
    );  
}  
  
// ajax(..) is some arbitrary Ajax function given by a library  
ajax( "http://some.url.1", response );  
ajax( "http://some.url.2", response );
```

If “`http://some.url.1`” gets its results back first, the entire list will be mapped into `res` all at once. If it's a few thousand or less records, this is not generally a big deal. But if it's say 10 million records, that can take a while to run (several seconds on a powerful laptop, much longer on a mobile device, etc.).

While such a “process” is running, nothing else in the page can happen, including no other `response(..)` calls, no UI updates, not even user events like scrolling, typing, button clicking, and the like. That's pretty painful.

So, to make a more cooperatively concurrent system, one that's friendlier and doesn't hog the event loop queue, you can process these results in asynchronous

batches, after each one “yielding” back to the event loop to let other waiting events happen.

Here’s a very simple approach:

```
var res = [];

// `response(..)` receives array of results from the Ajax call
function response(data) {
    // let's just do 1000 at a time
    var chunk = data.splice( 0, 1000 );

    // add onto existing `res` array
    res = res.concat(
        // make a new transformed array with all `chunk` values doubled
        chunk.map( function(val){
            return val * 2;
        })
    );

    // anything left to process?
    if (data.length > 0) {
        // async schedule next batch
        setTimeout( function(){
            response( data );
        }, 0 );
    }
}

// ajax(..) is some arbitrary Ajax function given by a library
ajax( "http://some.url.1", response );
ajax( "http://some.url.2", response );
```

We process the data set in maximum-sized chunks of 1,000 items. By doing so, we ensure a short-running “process,” even if that means many more subsequent “processes,” as the interleaving onto the event loop queue will give us a much more responsive (performant) site/app.

Of course, we’re not interaction-coordinating the ordering of any of these “processes,” so the order of results in `res` won’t be predictable. If ordering was required, you’d need to use interaction techniques like those we discussed earlier, or ones we will cover in later chapters of this book.

We use the `setTimeout(..0)` (hack) for async scheduling, which basically just means “stick this function at the end of the current event loop queue.”

Note: `setTimeout(..0)` is not technically inserting an item directly onto the event loop queue. The timer will insert the event at its next opportunity. For example, two subsequent `setTimeout(..0)` calls would not be strictly guaran-

teed to be processed in call order, so it *is* possible to see various conditions like timer drift where the ordering of such events isn't predictable. In Node.js, a similar approach is `process.nextTick(..)`. Despite how convenient (and usually more performant) it would be, there's not a single direct way (at least yet) across all environments to ensure async event ordering. We cover this topic in more detail in the next section.

Jobs

As of ES6, there's a new concept layered on top of the event loop queue, called the "Job queue." The most likely exposure you'll have to it is with the asynchronous behavior of Promises (see Chapter 3).

Unfortunately, at the moment it's a mechanism without an exposed API, and thus demonstrating it is a bit more convoluted. So we're going to have to just describe it conceptually, such that when we discuss async behavior with Promises in Chapter 3, you'll understand how those actions are being scheduled and processed.

So, the best way to think about this that I've found is that the "Job queue" is a queue hanging off the end of every tick in the event loop queue. Certain async-implied actions that may occur during a tick of the event loop will not cause a whole new event to be added to the event loop queue, but will instead add an item (aka Job) to the end of the current tick's Job queue.

It's kinda like saying, "oh, here's this other thing I need to do *later*, but make sure it happens right away before anything else can happen."

Or, to use a metaphor: the event loop queue is like an amusement park ride, where once you finish the ride, you have to go to the back of the line to ride again. But the Job queue is like finishing the ride, but then cutting in line and getting right back on.

A Job can also cause more Jobs to be added to the end of the same queue. So, it's theoretically possible that a Job "loop" (a Job that keeps adding another Job, etc.) could spin indefinitely, thus starving the program of the ability to move on to the next event loop tick. This would conceptually be almost the same as just expressing a long-running or infinite loop (like `while (true) ..`) in your code.

Jobs are kind of like the spirit of the `setTimeout(..0)` hack, but implemented in such a way as to have a much more well-defined and guaranteed ordering: **later, but as soon as possible**.

Let's imagine an API for scheduling Jobs (directly, without hacks), and call it `schedule(..)`. Consider:

```
console.log( "A" );
setTimeout( function(){
```

```

        console.log( "B" );
}, 0 );

// theoretical "Job API"
schedule( function(){
    console.log( "C" );

    schedule( function(){
        console.log( "D" );
    } );
} );

```

You might expect this to print out **A B C D**, but instead it would print out **A C D B**, because the Jobs happen at the end of the current event loop tick, and the timer fires to schedule for the *next* event loop tick (if available!).

In Chapter 3, we'll see that the asynchronous behavior of Promises is based on Jobs, so it's important to keep clear how that relates to event loop behavior.

Statement Ordering

The order in which we express statements in our code is not necessarily the same order as the JS engine will execute them. That may seem like quite a strange assertion to make, so we'll just briefly explore it.

But before we do, we should be crystal clear on something: the rules/grammar of the language (see the *Types & Grammar* title of this book series) dictate a very predictable and reliable behavior for statement ordering from the program point of view. So what we're about to discuss are **not things you should ever be able to observe** in your JS program.

Warning: If you are ever able to *observe* compiler statement reordering like we're about to illustrate, that'd be a clear violation of the specification, and it would unquestionably be due to a bug in the JS engine in question – one which should promptly be reported and fixed! But it's vastly more common that you *suspect* something crazy is happening in the JS engine, when in fact it's just a bug (probably a “race condition”!) in your own code – so look there first, and again and again. The JS debugger, using breakpoints and stepping through code line by line, will be your most powerful tool for sniffing out such bugs in *your code*.

Consider:

```

var a, b;

a = 10;
b = 30;

a = a + 1;

```

```
b = b + 1;  
  
console.log( a + b ); // 42
```

This code has no expressed asynchrony to it (other than the rare `console` async I/O discussed earlier!), so the most likely assumption is that it would process line by line in top-down fashion.

But it's *possible* that the JS engine, after compiling this code (yes, JS is compiled – see the *Scope & Closures* title of this book series!) might find opportunities to run your code faster by rearranging (safely) the order of these statements. Essentially, as long as you can't observe the reordering, anything's fair game.

For example, the engine might find it's faster to actually execute the code like this:

```
var a, b;  
  
a = 10;  
a++;  
  
b = 30;  
b++;  
  
console.log( a + b ); // 42
```

Or this:

```
var a, b;  
  
a = 11;  
b = 31;  
  
console.log( a + b ); // 42
```

Or even:

```
// because `a` and `b` aren't used anymore, we can  
// inline and don't even need them!  
console.log( 42 ); // 42
```

In all these cases, the JS engine is performing safe optimizations during its compilation, as the end *observable* result will be the same.

But here's a scenario where these specific optimizations would be unsafe and thus couldn't be allowed (of course, not to say that it's not optimized at all):

```
var a, b;  
  
a = 10;  
b = 30;
```

```
// we need `a` and `b` in their preincremented state!
console.log( a * b ); // 300

a = a + 1;
b = b + 1;

console.log( a + b ); // 42
```

Other examples where the compiler reordering could create observable side effects (and thus must be disallowed) would include things like any function call with side effects (even and especially getter functions), or ES6 Proxy objects (see the *ES6 & Beyond* title of this book series).

Consider:

```
function foo() {
  console.log( b );
  return 1;
}

var a, b, c;

// ES5.1 getter literal syntax
c = {
  get bar() {
    console.log( a );
    return 1;
  }
};

a = 10;
b = 30;

a += foo();           // 30
b += c.bar;          // 11

console.log( a + b ); // 42
```

If it weren't for the `console.log(..)` statements in this snippet (just used as a convenient form of observable side effect for the illustration), the JS engine would likely have been free, if it wanted to (who knows if it would!?), to reorder the code to:

```
// ...

a = 10 + foo();
b = 30 + c.bar;
```

// ...

While JS semantics thankfully protect us from the *observable* nightmares that compiler statement reordering would seem to be in danger of, it's still important to understand just how tenuous a link there is between the way source code is authored (in top-down fashion) and the way it runs after compilation.

Compiler statement reordering is almost a micro-metaphor for concurrency and interaction. As a general concept, such awareness can help you understand async JS code flow issues better.

Review

A JavaScript program is (practically) always broken up into two or more chunks, where the first chunk runs *now* and the next chunk runs *later*, in response to an event. Even though the program is executed chunk-by-chunk, all of them share the same access to the program scope and state, so each modification to state is made on top of the previous state.

Whenever there are events to run, the *event loop* runs until the queue is empty. Each iteration of the event loop is a “tick.” User interaction, IO, and timers enqueue events on the event queue.

At any given moment, only one event can be processed from the queue at a time. While an event is executing, it can directly or indirectly cause one or more subsequent events.

Concurrency is when two or more chains of events interleave over time, such that from a high-level perspective, they appear to be running *simultaneously* (even though at any given moment only one event is being processed).

It's often necessary to do some form of interaction coordination between these concurrent “processes” (as distinct from operating system processes), for instance to ensure ordering or to prevent “race conditions.” These “processes” can also *cooperate* by breaking themselves into smaller chunks and to allow other “process” interleaving.

You Don't Know JS: Async & Performance

Chapter 2: Callbacks

In Chapter 1, we explored the terminology and concepts around asynchronous programming in JavaScript. Our focus is on understanding the single-threaded (one-at-a-time) event loop queue that drives all “events” (async function invocations). We also explored various ways that concurrency patterns explain the relationships (if any!) between *simultaneously* running chains of events, or “processes” (tasks, function calls, etc.).

All our examples in Chapter 1 used the function as the individual, indivisible unit of operations, whereby inside the function, statements run in predictable order (above the compiler level!), but at the function-ordering level, events (aka async function invocations) can happen in a variety of orders.

In all these cases, the function is acting as a “callback,” because it serves as the target for the event loop to “call back into” the program, whenever that item in the queue is processed.

As you no doubt have observed, callbacks are by far the most common way that asynchrony in JS programs is expressed and managed. Indeed, the callback is the most fundamental async pattern in the language.

Countless JS programs, even very sophisticated and complex ones, have been written upon no other async foundation than the callback (with of course the concurrency interaction patterns we explored in Chapter 1). The callback function is the async work horse for JavaScript, and it does its job respectably.

Except... callbacks are not without their shortcomings. Many developers are excited by the *promise* (pun intended!) of better async patterns. But it's impossible to effectively use any abstraction if you don't understand what it's abstracting, and why.

In this chapter, we will explore a couple of those in depth, as motivation for why more sophisticated async patterns (explored in subsequent chapters of this book) are necessary and desired.

Continuations

Let's go back to the async callback example we started with in Chapter 1, but let me slightly modify it to illustrate a point:

```
// A
ajax( "...", function(..){
  // C
} );
// B
```

// A and // B represent the first half of the program (aka the *now*), and // C marks the second half of the program (aka the *later*). The first half executes right away, and then there's a “pause” of indeterminate length. At some future moment, if the Ajax call completes, then the program will pick up where it left off, and *continue* with the second half.

In other words, the callback function wraps or encapsulates the *continuation* of the program.

Let's make the code even simpler:

```
// A
setTimeout( function(){
  // C
}, 1000 );
// B
```

Stop for a moment and ask yourself how you'd describe (to someone else less informed about how JS works) the way that program behaves. Go ahead, try it out loud. It's a good exercise that will help my next points make more sense.

Most readers just now probably thought or said something to the effect of: “Do A, then set up a timeout to wait 1,000 milliseconds, then once that fires, do C.” How close was your rendition?

You might have caught yourself and self-edited to: “Do A, setup the timeout for 1,000 milliseconds, then do B, then after the timeout fires, do C.” That's more accurate than the first version. Can you spot the difference?

Even though the second version is more accurate, both versions are deficient in explaining this code in a way that matches our brains to the code, and the code to the JS engine. The disconnect is both subtle and monumental, and is at the very heart of understanding the shortcomings of callbacks as async expression and management.

As soon as we introduce a single continuation (or several dozen as many programs do!) in the form of a callback function, we have allowed a divergence to form between how our brains work and the way the code will operate. Any time these two diverge (and this is by far not the only place that happens, as I'm sure you know!), we run into the inevitable fact that our code becomes harder to understand, reason about, debug, and maintain.

Sequential Brain

I'm pretty sure most of you readers have heard someone say (even made the claim yourself), “I'm a multitasker.” The effects of trying to act as a multitasker range from humorous (e.g., the silly patting-head-rubbing-stomach kids' game) to mundane (chewing gum while walking) to downright dangerous (texting while driving).

But are we multitaskers? Can we really do two conscious, intentional actions at once and think/reason about both of them at exactly the same moment? Does our highest level of brain functionality have parallel multithreading going on?

The answer may surprise you: **probably not.**

That's just not really how our brains appear to be set up. We're much more single taskers than many of us (especially A-type personalities!) would like to admit. We can really only think about one thing at any given instant.

I'm not talking about all our involuntary, subconscious, automatic brain functions, such as heart beating, breathing, and eyelid blinking. Those are all vital tasks to our sustained life, but we don't intentionally allocate any brain power to them. Thankfully, while we obsess about checking social network feeds for the 15th time in three minutes, our brain carries on in the background (threads!) with all those important tasks.

We're instead talking about whatever task is at the forefront of our minds at the moment. For me, it's writing the text in this book right now. Am I doing any other higher level brain function at exactly this same moment? Nope, not really. I get distracted quickly and easily – a few dozen times in these last couple of paragraphs!

When we *fake* multitasking, such as trying to type something at the same time we're talking to a friend or family member on the phone, what we're actually most likely doing is acting as fast context switchers. In other words, we switch back and forth between two or more tasks in rapid succession, *simultaneously* progressing on each task in tiny, fast little chunks. We do it so fast that to the outside world it appears as if we're doing these things *in parallel*.

Does that sound suspiciously like async evented concurrency (like the sort that happens in JS) to you?! If not, go back and read Chapter 1 again!

In fact, one way of simplifying (i.e., abusing) the massively complex world of neurology into something I can remotely hope to discuss here is that our brains work kinda like the event loop queue.

If you think about every single letter (or word) I type as a single async event, in just this sentence alone there are several dozen opportunities for my brain to be interrupted by some other event, such as from my senses, or even just my random thoughts.

I don't get interrupted and pulled to another "process" at every opportunity that I could be (thankfully – or this book would never be written!). But it happens often enough that I feel my own brain is nearly constantly switching to various different contexts (aka "processes"). And that's an awful lot like how the JS engine would probably feel.

Doing Versus Planning

OK, so our brains can be thought of as operating in single-threaded event loop queue like ways, as can the JS engine. That sounds like a good match.

But we need to be more nuanced than that in our analysis. There's a big, observable difference between how we plan various tasks, and how our brains actually operate those tasks.

Again, back to the writing of this text as my metaphor. My rough mental outline plan here is to keep writing and writing, going sequentially through a set of points I have ordered in my thoughts. I don't plan to have any interruptions or nonlinear activity in this writing. But yet, my brain is nevertheless switching around all the time.

Even though at an operational level our brains are async evented, we seem to plan out tasks in a sequential, synchronous way. "I need to go to the store, then buy some milk, then drop off my dry cleaning."

You'll notice that this higher level thinking (planning) doesn't seem very async evented in its formulation. In fact, it's kind of rare for us to deliberately think solely in terms of events. Instead, we plan things out carefully, sequentially (A then B then C), and we assume to an extent a sort of temporal blocking that forces B to wait on A, and C to wait on B.

When a developer writes code, they are planning out a set of actions to occur. If they're any good at being a developer, they're **carefully planning** it out. "I need to set `z` to the value of `x`, and then `x` to the value of `y`," and so forth.

When we write out synchronous code, statement by statement, it works a lot like our errands to-do list:

```
// swap `x` and `y` (via temp variable `z`)
z = x;
x = y;
y = z;
```

These three assignment statements are synchronous, so `x = y` waits for `z = x` to finish, and `y = z` in turn waits for `x = y` to finish. Another way of saying it is that these three statements are temporally bound to execute in a certain order, one right after the other. Thankfully, we don't need to be bothered with any async evented details here. If we did, the code gets a lot more complex, quickly!

So if synchronous brain planning maps well to synchronous code statements, how well do our brains do at planning out asynchronous code?

It turns out that how we express asynchrony (with callbacks) in our code doesn't map very well at all to that synchronous brain planning behavior.

Can you actually imagine having a line of thinking that plans out your to-do errands like this?

“I need to go to the store, but on the way I’m sure I’ll get a phone call, so ‘Hi, Mom’, and while she starts talking, I’ll be looking up the store address on GPS, but that’ll take a second to load, so I’ll turn down the radio so I can hear Mom better, then I’ll realize I forgot to put on a jacket and it’s cold outside, but no matter, keep driving and talking to Mom, and then the seatbelt ding reminds me to buckle up, so ‘Yes, Mom, I am wearing my seatbelt, I always do!’. Ah, finally the GPS got the directions, now...”

As ridiculous as that sounds as a formulation for how we plan our day out and think about what to do and in what order, nonetheless it’s exactly how our brains operate at a functional level. Remember, that’s not multitasking, it’s just fast context switching.

The reason it’s difficult for us as developers to write async evented code, especially when all we have is the callback to do it, is that stream of consciousness thinking/planning is unnatural for most of us.

We think in step-by-step terms, but the tools (callbacks) available to us in code are not expressed in a step-by-step fashion once we move from synchronous to asynchronous.

And **that** is why it’s so hard to accurately author and reason about async JS code with callbacks: because it’s not how our brain planning works.

Note: The only thing worse than not knowing why some code breaks is not knowing why it worked in the first place! It’s the classic “house of cards” mentality: “it works, but not sure why, so nobody touch it!” You may have heard, “Hell is other people” (Sartre), and the programmer meme twist, “Hell is other people’s code.” I believe truly: “Hell is not understanding my own code.” And callbacks are one main culprit.

Nested/Chained Callbacks

Consider:

```
listen( "click", function handler(evt){  
    setTimeout( function request(){  
        ajax( "http://some.url.1", function response(text){  
            if (text == "hello") {  
                handler();  
            }  
            else if (text == "world") {  
                request();  
            }  
        } );  
    } );
```

```
    }, 500) ;  
} );
```

There's a good chance code like that is recognizable to you. We've got a chain of three functions nested together, each one representing a step in an asynchronous series (task, "process").

This kind of code is often called "callback hell," and sometimes also referred to as the "pyramid of doom" (for its sideways-facing triangular shape due to the nested indentation).

But "callback hell" actually has almost nothing to do with the nesting/indentation. It's a far deeper problem than that. We'll see how and why as we continue through the rest of this chapter.

First, we're waiting for the "click" event, then we're waiting for the timer to fire, then we're waiting for the Ajax response to come back, at which point it might do it all again.

At first glance, this code may seem to map its asynchrony naturally to sequential brain planning.

First (*now*), we:

```
listen( "...", function handler(...){  
    // ..  
} );
```

Then *later*, we:

```
setTimeout( function request(...){  
    // ..  
}, 500) ;
```

Then still *later*, we:

```
ajax( "...", function response(...){  
    // ..  
} );
```

And finally (most *later*), we:

```
if ( .. ) {  
    // ..  
}  
else ..
```

But there's several problems with reasoning about this code linearly in such a fashion.

First, it's an accident of the example that our steps are on subsequent lines (1, 2, 3, and 4...). In real async JS programs, there's often a lot more noise cluttering things up, noise that we have to deftly maneuver past in our brains

as we jump from one function to the next. Understanding the async flow in such callback-laden code is not impossible, but it's certainly not natural or easy, even with lots of practice.

But also, there's something deeper wrong, which isn't evident just in that code example. Let me make up another scenario (pseudocode-ish) to illustrate it:

```
doA( function(){  
    doB();  
  
    doC( function(){  
        doD();  
    } )  
  
    doE();  
} );  
  
doF();
```

While the experienced among you will correctly identify the true order of operations here, I'm betting it is more than a little confusing at first glance, and takes some concerted mental cycles to arrive at. The operations will happen in this order:

- doA()
- doF()
- doB()
- doC()
- doE()
- doD()

Did you get that right the very first time you glanced at the code?

OK, some of you are thinking I was unfair in my function naming, to intentionally lead you astray. I swear I was just naming in top-down appearance order. But let me try again:

```
doA( function(){  
    doC();  
  
    doD( function(){  
        doF();  
    } )  
  
    doE();  
} );  
  
doB();
```

Now, I've named them alphabetically in order of actual execution. But I still

bet, even with experience now in this scenario, tracing through the A → B → C → D → E → F order doesn't come natural to many if any of you readers. Certainly, your eyes do an awful lot of jumping up and down the code snippet, right?

But even if that all comes natural to you, there's still one more hazard that could wreak havoc. Can you spot what it is?

What if `doA(..)` or `doD(..)` aren't actually async, the way we obviously assumed them to be? Uh oh, now the order is different. If they're both sync (and maybe only sometimes, depending on the conditions of the program at the time), the order is now A → C → D → F → E → B.

That sound you just heard faintly in the background is the sighs of thousands of JS developers who just had a face-in-hands moment.

Is nesting the problem? Is that what makes it so hard to trace the async flow? That's part of it, certainly.

But let me rewrite the previous nested event/timeout/Ajax example without using nesting:

```
listen( "click", handler );

function handler() {
    setTimeout( request, 500 );
}

function request(){
    ajax( "http://some.url.1", response );
}

function response(text){
    if (text == "hello") {
        handler();
    }
    else if (text == "world") {
        request();
    }
}
```

This formulation of the code is not hardly as recognizable as having the nesting/indentation woes of its previous form, and yet it's every bit as susceptible to "callback hell." Why?

As we go to linearly (sequentially) reason about this code, we have to skip from one function, to the next, to the next, and bounce all around the code base to "see" the sequence flow. And remember, this is simplified code in sort of best-case fashion. We all know that real async JS program code bases are often

fantastically more jumbled, which makes such reasoning orders of magnitude more difficult.

Another thing to notice: to get steps 2, 3, and 4 linked together so they happen in succession, the only affordance callbacks alone gives us is to hardcode step 2 into step 1, step 3 into step 2, step 4 into step 3, and so on. The hardcoding isn't necessarily a bad thing, if it really is a fixed condition that step 2 should always lead to step 3.

But the hardcoding definitely makes the code a bit more brittle, as it doesn't account for anything going wrong that might cause a deviation in the progression of steps. For example, if step 2 fails, step 3 never gets reached, nor does step 2 retry, or move to an alternate error handling flow, and so on.

All of these issues are things you *can* manually hardcode into each step, but that code is often very repetitive and not reusable in other steps or in other async flows in your program.

Even though our brains might plan out a series of tasks in a sequential type of way (this, then this, then this), the evented nature of our brain operation makes recovery/retry/forking of flow control almost effortless. If you're out running errands, and you realize you left a shopping list at home, it doesn't end the day because you didn't plan that ahead of time. Your brain routes around this hiccup easily: you go home, get the list, then head right back out to the store.

But the brittle nature of manually hardcoded callbacks (even with hardcoded error handling) is often far less graceful. Once you end up specifying (aka pre-planning) all the various eventualities/paths, the code becomes so convoluted that it's hard to ever maintain or update it.

That is what “callback hell” is all about! The nesting/indentation are basically a side show, a red herring.

And as if all that's not enough, we haven't even touched what happens when two or more chains of these callback continuations are happening *simultaneously*, or when the third step branches out into “parallel” callbacks with gates or latches, or... OMG, my brain hurts, how about yours!?

Are you catching the notion here that our sequential, blocking brain planning behaviors just don't map well onto callback-oriented async code? That's the first major deficiency to articulate about callbacks: they express asynchrony in code in ways our brains have to fight just to keep in sync with (pun intended!).

Trust Issues

The mismatch between sequential brain planning and callback-driven async JS code is only part of the problem with callbacks. There's something much deeper to be concerned about.

Let's once again revisit the notion of a callback function as the continuation (aka the second half) of our program:

```
// A
ajax( "...", function(..){
  // C
} );
// B
```

// A and // B happen *now*, under the direct control of the main JS program. But // C gets deferred to happen *later*, and under the control of another party – in this case, the `ajax(..)` function. In a basic sense, that sort of hand-off of control doesn't regularly cause lots of problems for programs.

But don't be fooled by its infrequency that this control switch isn't a big deal. In fact, it's one of the worst (and yet most subtle) problems about callback-driven design. It revolves around the idea that sometimes `ajax(..)` (i.e., the “party” you hand your callback continuation to) is not a function that you wrote, or that you directly control. Many times, it's a utility provided by some third party.

We call this “inversion of control,” when you take part of your program and give over control of its execution to another third party. There's an unspoken “contract” that exists between your code and the third-party utility – a set of things you expect to be maintained.

Tale of Five Callbacks

It might not be terribly obvious why this is such a big deal. Let me construct an exaggerated scenario to illustrate the hazards of trust at play.

Imagine you're a developer tasked with building out an ecommerce checkout system for a site that sells expensive TVs. You already have all the various pages of the checkout system built out just fine. On the last page, when the user clicks “confirm” to buy the TV, you need to call a third-party function (provided say by some analytics tracking company) so that the sale can be tracked.

You notice that they've provided what looks like an async tracking utility, probably for the sake of performance best practices, which means you need to pass in a callback function. In this continuation that you pass in, you will have the final code that charges the customer's credit card and displays the thank you page.

This code might look like:

```
analytics.trackPurchase( purchaseData, function(){
  chargeCreditCard();
  displayThankyouPage();
} );
```

Easy enough, right? You write the code, test it, everything works, and you deploy to production. Everyone's happy!

Six months go by and no issues. You've almost forgotten you even wrote that code. One morning, you're at a coffee shop before work, casually enjoying your latte, when you get a panicked call from your boss insisting you drop the coffee and rush into work right away.

When you arrive, you find out that a high-profile customer has had his credit card charged five times for the same TV, and he's understandably upset. Customer service has already issued an apology and processed a refund. But your boss demands to know how this could possibly have happened. "Don't we have tests for stuff like this!?"

You don't even remember the code you wrote. But you dig back in and start trying to find out what could have gone awry.

After digging through some logs, you come to the conclusion that the only explanation is that the analytics utility somehow, for some reason, called your callback five times instead of once. Nothing in their documentation mentions anything about this.

Frustrated, you contact customer support, who of course is as astonished as you are. They agree to escalate it to their developers, and promise to get back to you. The next day, you receive a lengthy email explaining what they found, which you promptly forward to your boss.

Apparently, the developers at the analytics company had been working on some experimental code that, under certain conditions, would retry the provided callback once per second, for five seconds, before failing with a timeout. They had never intended to push that into production, but somehow they did, and they're totally embarrassed and apologetic. They go into plenty of detail about how they've identified the breakdown and what they'll do to ensure it never happens again. Yadda, yadda.

What's next?

You talk it over with your boss, but he's not feeling particularly comfortable with the state of things. He insists, and you reluctantly agree, that you can't trust *them* anymore (that's what bit you), and that you'll need to figure out how to protect the checkout code from such a vulnerability again.

After some tinkering, you implement some simple ad hoc code like the following, which the team seems happy with:

```
var tracked = false;

analytics.trackPurchase( purchaseData, function(){
    if (!tracked) {
        tracked = true;
        chargeCreditCard();
    }
})
```

```
        displayThankyouPage();
    }
} );
```

Note: This should look familiar to you from Chapter 1, because we're essentially creating a latch to handle if there happen to be multiple concurrent invocations of our callback.

But then one of your QA engineers asks, “what happens if they never call the callback?” Oops. Neither of you had thought about that.

You begin to chase down the rabbit hole, and think of all the possible things that could go wrong with them calling your callback. Here's roughly the list you come up with of ways the analytics utility could misbehave:

- Call the callback too early (before it's been tracked)
- Call the callback too late (or never)
- Call the callback too few or too many times (like the problem you encountered!)
- Fail to pass along any necessary environment/parameters to your callback
- Swallow any errors/exceptions that may happen
- ...

That should feel like a troubling list, because it is. You're probably slowly starting to realize that you're going to have to invent an awful lot of ad hoc logic **in each and every single callback** that's passed to a utility you're not positive you can trust.

Now you realize a bit more completely just how hellish “callback hell” is.

Not Just Others' Code

Some of you may be skeptical at this point whether this is as big a deal as I'm making it out to be. Perhaps you don't interact with truly third-party utilities much if at all. Perhaps you use versioned APIs or self-host such libraries, so that its behavior can't be changed out from underneath you.

So, contemplate this: can you even *really* trust utilities that you do theoretically control (in your own code base)?

Think of it this way: most of us agree that at least to some extent we should build our own internal functions with some defensive checks on the input parameters, to reduce/prevent unexpected issues.

Overly trusting of input:

```
function addNumbers(x,y) {
    // + is overloaded with coercion to also be
    // string concatenation, so this operation
    // isn't strictly safe depending on what's
    // passed in.
```

```

        return x + y;
    }

addNumbers( 21, 21 ); // 42
addNumbers( 21, "21" ); // "2121"

```

Defensive against untrusted input:

```

function addNumbers(x,y) {
    // ensure numerical input
    if (typeof x != "number" || typeof y != "number") {
        throw Error( "Bad parameters" );
    }

    // if we get here, + will safely do numeric addition
    return x + y;
}

addNumbers( 21, 21 ); // 42
addNumbers( 21, "21" ); // Error: "Bad parameters"

```

Or perhaps still safe but friendlier:

```

function addNumbers(x,y) {
    // ensure numerical input
    x = Number( x );
    y = Number( y );

    // + will safely do numeric addition
    return x + y;
}

addNumbers( 21, 21 ); // 42
addNumbers( 21, "21" ); // 42

```

However you go about it, these sorts of checks/normalizations are fairly common on function inputs, even with code we theoretically entirely trust. In a crude sort of way, it's like the programming equivalent of the geopolitical principle of "Trust But Verify."

So, doesn't it stand to reason that we should do the same thing about composition of async function callbacks, not just with truly external code but even with code we know is generally "under our own control"? **Of course we should.**

But callbacks don't really offer anything to assist us. We have to construct all that machinery ourselves, and it often ends up being a lot of boilerplate/overhead that we repeat for every single async callback.

The most troublesome problem with callbacks is *inversion of control* leading to a complete breakdown along all those trust lines.

If you have code that uses callbacks, especially but not exclusively with third-party utilities, and you’re not already applying some sort of mitigation logic for all these *inversion of control* trust issues, your code *has* bugs in it right now even though they may not have bitten you yet. Latent bugs are still bugs.

Hell indeed.

Trying to Save Callbacks

There are several variations of callback design that have attempted to address some (not all!) of the trust issues we’ve just looked at. It’s a valiant, but doomed, effort to save the callback pattern from imploding on itself.

For example, regarding more graceful error handling, some API designs provide for split callbacks (one for the success notification, one for the error notification):

```
function success(data) {
  console.log( data );
}

function failure(err) {
  console.error( err );
}

ajax( "http://some.url.1", success, failure );
```

In APIs of this design, often the `failure()` error handler is optional, and if not provided it will be assumed you want the errors swallowed. Ugh.

Note: This split-callback design is what the ES6 Promise API uses. We’ll cover ES6 Promises in much more detail in the next chapter.

Another common callback pattern is called “error-first style” (sometimes called “Node style,” as it’s also the convention used across nearly all Node.js APIs), where the first argument of a single callback is reserved for an error object (if any). If success, this argument will be empty/falsy (and any subsequent arguments will be the success data), but if an error result is being signaled, the first argument is set/truthy (and usually nothing else is passed):

```
function response(err,data) {
  // error?
  if (err) {
    console.error( err );
  }
  // otherwise, assume success
  else {
    console.log( data );
  }
}
```

```
ajax( "http://some.url.1", response );
```

In both of these cases, several things should be observed.

First, it has not really resolved the majority of trust issues like it may appear. There's nothing about either callback that prevents or filters unwanted repeated invocations. Moreover, things are worse now, because you may get both success and error signals, or neither, and you still have to code around either of those conditions.

Also, don't miss the fact that while it's a standard pattern you can employ, it's definitely more verbose and boilerplate-ish without much reuse, so you're going to get weary of typing all that out for every single callback in your application.

What about the trust issue of never being called? If this is a concern (and it probably should be!), you likely will need to set up a timeout that cancels the event. You could make a utility (proof-of-concept only shown) to help you with that:

```
function timeoutify(fn,delay) {
  var intv = setTimeout( function(){
    intv = null;
    fn( new Error( "Timeout!" ) );
  }, delay )
;

  return function() {
    // timeout hasn't happened yet?
    if (intv) {
      clearTimeout( intv );
      fn.apply( this, [ null ].concat( [].slice.call( arguments ) ) );
    }
  };
}
```

Here's how you use it:

```
// using "error-first style" callback design
function foo(err,data) {
  if (err) {
    console.error( err );
  }
  else {
    console.log( data );
  }
}

ajax( "http://some.url.1", timeoutify( foo, 500 ) );
```

Another trust issue is being called “too early.” In application-specific terms, this may actually involve being called before some critical task is complete. But more generally, the problem is evident in utilities that can either invoke the callback you provide *now* (synchronously), or *later* (asynchronously).

This nondeterminism around the sync-or-async behavior is almost always going to lead to very difficult to track down bugs. In some circles, the fictional insanity-inducing monster named Zalgo is used to describe the sync/async nightmares. “Don’t release Zalgo!” is a common cry, and it leads to very sound advice: always invoke callbacks asynchronously, even if that’s “right away” on the next turn of the event loop, so that all callbacks are predictably async.

Note: For more information on Zalgo, see Oren Golan’s “Don’t Release Zalgo!” (<https://github.com/oren/oren.github.io/blob/master/posts/zalgo.md>) and Isaac Z. Schlueter’s “Designing APIs for Asynchrony” (<http://blog.izs.me/post/59142742143/designing-apis-for-asynchrony>).

Consider:

```
function result(data) {
  console.log( a );
}

var a = 0;

ajax( "..pre-cached-url..", result );
a++;
```

Will this code print 0 (sync callback invocation) or 1 (async callback invocation)? Depends... on the conditions.

You can see just how quickly the unpredictability of Zalgo can threaten any JS program. So the silly-sounding “never release Zalgo” is actually incredibly common and solid advice. Always be asyncing.

What if you don’t know whether the API in question will always execute async? You could invent a utility like this `asyncify(..)` proof-of-concept:

```
function asyncify(fn) {
  var orig_fn = fn,
    intv = setTimeout( function(){
      intv = null;
      if (fn) fn();
    }, 0 )
  ;

  fn = null;

  return function() {
    // firing too quickly, before `intv` timer has fired to
```

```

// indicate async turn has passed?
if (intv) {
  fn = orig_fn.bind.apply(
    orig_fn,
    // add the wrapper's `this` to the `bind(..)`
    // call parameters, as well as currying any
    // passed in parameters
    [this].concat( [].slice.call( arguments ) )
  );
}
// already async
else {
  // invoke original function
  orig_fn.apply( this, arguments );
}
};

}

```

You use `asyncify(..)` like this:

```

function result(data) {
  console.log( a );
}

var a = 0;

ajax( "..pre-cached-url..", asyncify( result ) );
a++;

```

Whether the Ajax request is in the cache and resolves to try to call the callback right away, or must be fetched over the wire and thus complete later asynchronously, this code will always output 1 instead of 0 – `result(..)` cannot help but be invoked asynchronously, which means the `a++` has a chance to run before `result(..)` does.

Yay, another trust issued “solved”! But it’s inefficient, and yet again more bloated boilerplate to weigh your project down.

That’s just the story, over and over again, with callbacks. They can do pretty much anything you want, but you have to be willing to work hard to get it, and oftentimes this effort is much more than you can or should spend on such code reasoning.

You might find yourself wishing for built-in APIs or other language mechanics to address these issues. Finally ES6 has arrived on the scene with some great answers, so keep reading!

Review

Callbacks are the fundamental unit of asynchrony in JS. But they're not enough for the evolving landscape of async programming as JS matures.

First, our brains plan things out in sequential, blocking, single-threaded semantic ways, but callbacks express asynchronous flow in a rather nonlinear, non-sequential way, which makes reasoning properly about such code much harder. Bad to reason about code is bad code that leads to bad bugs.

We need a way to express asynchrony in a more synchronous, sequential, blocking manner, just like our brains do.

Second, and more importantly, callbacks suffer from *inversion of control* in that they implicitly give control over to another party (often a third-party utility not in your control!) to invoke the *continuation* of your program. This control transfer leads us to a troubling list of trust issues, such as whether the callback is called more times than we expect.

Inventing ad hoc logic to solve these trust issues is possible, but it's more difficult than it should be, and it produces clunkier and harder to maintain code, as well as code that is likely insufficiently protected from these hazards until you get visibly bitten by the bugs.

We need a generalized solution to **all of the trust issues**, one that can be reused for as many callbacks as we create without all the extra boilerplate overhead.

We need something better than callbacks. They've served us well to this point, but the *future* of JavaScript demands more sophisticated and capable async patterns. The subsequent chapters in this book will dive into those emerging evolutions.

You Don't Know JS: Async & Performance

Chapter 3: Promises

In Chapter 2, we identified two major categories of deficiencies with using callbacks to express program asynchrony and manage concurrency: lack of sequentiality and lack of trustability. Now that we understand the problems more intimately, it's time we turn our attention to patterns that can address them.

The issue we want to address first is the *inversion of control*, the trust that is so fragilely held and so easily lost.

Recall that we wrap up the *continuation* of our program in a callback function, and hand that callback over to another party (potentially even external code) and just cross our fingers that it will do the right thing with the invocation of the callback.

We do this because we want to say, “here's what happens *later*, after the current step finishes.”

But what if we could uninvert that *inversion of control*? What if instead of handing the continuation of our program to another party, we could expect it to return us a capability to know when its task finishes, and then our code could decide what to do next?

This paradigm is called **Promises**.

Promises are starting to take the JS world by storm, as developers and specification writers alike desperately seek to untangle the insanity of callback hell in their code/design. In fact, most new async APIs being added to JS/DOM platform are being built on Promises. So it's probably a good idea to dig in and learn them, don't you think!?

Note: The word “immediately” will be used frequently in this chapter, generally to refer to some Promise resolution action. However, in essentially all cases, “immediately” means in terms of the Job queue behavior (see Chapter 1), not in the strictly synchronous *now* sense.

What Is a Promise?

When developers decide to learn a new technology or pattern, usually their first step is “Show me the code!” It's quite natural for us to just jump in feet first and learn as we go.

But it turns out that some abstractions get lost on the APIs alone. Promises are one of those tools where it can be painfully obvious from how someone uses it whether they understand what it's for and about versus just learning and using the API.

So before I show the Promise code, I want to fully explain what a Promise really is conceptually. I hope this will then guide you better as you explore integrating Promise theory into your own async flow.

With that in mind, let's look at two different analogies for what a Promise *is*.

Future Value

Imagine this scenario: I walk up to the counter at a fast-food restaurant, and place an order for a cheeseburger. I hand the cashier \$1.47. By placing my order and paying for it, I've made a request for a *value* back (the cheeseburger). I've started a transaction.

But often, the cheeseburger is not immediately available for me. The cashier hands me something in place of my cheeseburger: a receipt with an order number on it. This order number is an IOU ("I owe you") *promise* that ensures that eventually, I should receive my cheeseburger.

So I hold onto my receipt and order number. I know it represents my *future cheeseburger*, so I don't need to worry about it anymore – aside from being hungry!

While I wait, I can do other things, like send a text message to a friend that says, "Hey, can you come join me for lunch? I'm going to eat a cheeseburger."

I am reasoning about my *future cheeseburger* already, even though I don't have it in my hands yet. My brain is able to do this because it's treating the order number as a placeholder for the cheeseburger. The placeholder essentially makes the value *time independent*. It's a **future value**.

Eventually, I hear, "Order 113!" and I gleefully walk back up to the counter with receipt in hand. I hand my receipt to the cashier, and I take my cheeseburger in return.

In other words, once my *future value* was ready, I exchanged my value-promise for the value itself.

But there's another possible outcome. They call my order number, but when I go to retrieve my cheeseburger, the cashier regretfully informs me, "I'm sorry, but we appear to be all out of cheeseburgers." Setting aside the customer frustration of this scenario for a moment, we can see an important characteristic of *future values*: they can either indicate a success or failure.

Every time I order a cheeseburger, I know that I'll either get a cheeseburger eventually, or I'll get the sad news of the cheeseburger shortage, and I'll have to figure out something else to eat for lunch.

Note: In code, things are not quite as simple, because metaphorically the order number may never be called, in which case we're left indefinitely in an unresolved state. We'll come back to dealing with that case later.

Values Now and Later This all might sound too mentally abstract to apply to your code. So let's be more concrete.

However, before we can introduce how Promises work in this fashion, we're going to derive in code that we already understand – callbacks! – how to handle these *future values*.

When you write code to reason about a value, such as performing math on a **number**, whether you realize it or not, you've been assuming something very fundamental about that value, which is that it's a concrete *now* value already:

```
var x, y = 2;

console.log( x + y ); // NaN <-- because `x` isn't set yet
```

The `x + y` operation assumes both `x` and `y` are already set. In terms we'll expound on shortly, we assume the `x` and `y` values are already *resolved*.

It would be nonsense to expect that the `+` operator by itself would somehow be magically capable of detecting and waiting around until both `x` and `y` are resolved (aka *ready*), only then to do the operation. That would cause chaos in the program if different statements finished *now* and others finished *later*, right?

How could you possibly reason about the relationships between two statements if either one (or both) of them might not be finished yet? If statement 2 relies on statement 1 being finished, there are just two outcomes: either statement 1 finished right *now* and everything proceeds fine, or statement 1 didn't finish yet, and thus statement 2 is going to fail.

If this sort of thing sounds familiar from Chapter 1, good!

Let's go back to our `x + y` math operation. Imagine if there was a way to say, “Add `x` and `y`, but if either of them isn't ready yet, just wait until they are. Add them as soon as you can.”

Your brain might have just jumped to callbacks. OK, so...

```
function add(getX,getY,cb) {
  var x, y;
  getX( function(xVal){
    x = xVal;
    // both are ready?
    if (y != undefined) {
      cb( x + y ); // send along sum
    }
  });
  getY( function(yVal){
    y = yVal;
    // both are ready?
    if (x != undefined) {
      cb( x + y ); // send along sum
    }
  });
}
```

```

        }
    } );
}

// `fetchX()` and `fetchY()` are sync or async
// functions
add( fetchX, fetchY, function(sum){
    console.log( sum ); // that was easy, huh?
} );

```

Take just a moment to let the beauty (or lack thereof) of that snippet sink in (whistles patiently).

While the ugliness is undeniable, there's something very important to notice about this `async` pattern.

In that snippet, we treated `x` and `y` as future values, and we express an operation `add(...)` that (from the outside) does not care whether `x` or `y` or both are available right away or not. In other words, it normalizes the *now* and *later*, such that we can rely on a predictable outcome of the `add(...)` operation.

By using an `add(...)` that is temporally consistent – it behaves the same across *now* and *later* times – the `async` code is much easier to reason about.

To put it more plainly: to consistently handle both *now* and *later*, we make both of them *later*: all operations become `async`.

Of course, this rough callbacks-based approach leaves much to be desired. It's just a first tiny step toward realizing the benefits of reasoning about *future values* without worrying about the time aspect of when it's available or not.

Promise Value We'll definitely go into a lot more detail about Promises later in the chapter – so don't worry if some of this is confusing – but let's just briefly glimpse at how we can express the `x + y` example via Promises:

```

function add(xPromise,yPromise) {
    // `Promise.all([ .. ])` takes an array of promises,
    // and returns a new promise that waits on them
    // all to finish
    return Promise.all( [xPromise, yPromise] )

    // when that promise is resolved, let's take the
    // received `X` and `Y` values and add them together.
    .then( function(values){
        // `values` is an array of the messages from the
        // previously resolved promises
        return values[0] + values[1];
    } );
}

```

```

// `fetchX()` and `fetchY()` return promises for
// their respective values, which may be ready
// *now* or *later*.
add( fetchX(), fetchY() )

// we get a promise back for the sum of those
// two numbers.
// now we chain-call `then(..)` to wait for the
// resolution of that returned promise.
.then( function(sum){
  console.log( sum ); // that was easier!
} );

```

There are two layers of Promises in this snippet.

`fetchX()` and `fetchY()` are called directly, and the values they return (promises!) are passed into `add(..)`. The underlying values those promises represent may be ready *now* or *later*, but each promise normalizes the behavior to be the same regardless. We reason about X and Y values in a time-independent way. They are *future values*.

The second layer is the promise that `add(..)` creates (via `Promise.all([..])`) and returns, which we wait on by calling `then(..)`. When the `add(..)` operation completes, our `sum` *future value* is ready and we can print it out. We hide inside of `add(..)` the logic for waiting on the X and Y *future values*.

Note: Inside `add(..)`, the `Promise.all([..])` call creates a promise (which is waiting on `promiseX` and `promiseY` to resolve). The chained call to `.then(..)` creates another promise, which the `return values[0] + values[1]` line immediately resolves (with the result of the addition). Thus, the `then(..)` call we chain off the end of the `add(..)` call – at the end of the snippet – is actually operating on that second promise returned, rather than the first one created by `Promise.all([..])`. Also, though we are not chaining off the end of that second `then(..)`, it too has created another promise, had we chosen to observe/use it. This Promise chaining stuff will be explained in much greater detail later in this chapter.

Just like with cheeseburger orders, it's possible that the resolution of a Promise is rejection instead of fulfillment. Unlike a fulfilled Promise, where the value is always programmatic, a rejection value – commonly called a “rejection reason” – can either be set directly by the program logic, or it can result implicitly from a runtime exception.

With Promises, the `then(..)` call can actually take two functions, the first for fulfillment (as shown earlier), and the second for rejection:

```

add( fetchX(), fetchY() )
.then(

```

```

    // fulfillment handler
    function(sum) {
      console.log( sum );
    },
    // rejection handler
    function(err) {
      console.error( err ); // bummer!
    }
);

```

If something went wrong getting X or Y, or something somehow failed during the addition, the promise that `add(..)` returns is rejected, and the second callback error handler passed to `then(..)` will receive the rejection value from the promise.

Because Promises encapsulate the time-dependent state – waiting on the fulfillment or rejection of the underlying value – from the outside, the Promise itself is time-independent, and thus Promises can be composed (combined) in predictable ways regardless of the timing or outcome underneath.

Moreover, once a Promise is resolved, it stays that way forever – it becomes an *immutable value* at that point – and can then be *observed* as many times as necessary.

Note: Because a Promise is externally immutable once resolved, it's now safe to pass that value around to any party and know that it cannot be modified accidentally or maliciously. This is especially true in relation to multiple parties observing the resolution of a Promise. It is not possible for one party to affect another party's ability to observe Promise resolution. Immutability may sound like an academic topic, but it's actually one of the most fundamental and important aspects of Promise design, and shouldn't be casually passed over.

That's one of the most powerful and important concepts to understand about Promises. With a fair amount of work, you could ad hoc create the same effects with nothing but ugly callback composition, but that's not really an effective strategy, especially because you have to do it over and over again.

Promises are an easily repeatable mechanism for encapsulating and composing *future values*.

Completion Event

As we just saw, an individual Promise behaves as a *future value*. But there's another way to think of the resolution of a Promise: as a flow-control mechanism – a temporal this-then-that – for two or more steps in an asynchronous task.

Let's imagine calling a function `foo(..)` to perform some task. We don't know about any of its details, nor do we care. It may complete the task right away, or it may take a while.

We just simply need to know when `foo(..)` finishes so that we can move on to our next task. In other words, we'd like a way to be notified of `foo(..)`'s completion so that we can *continue*.

In typical JavaScript fashion, if you need to listen for a notification, you'd likely think of that in terms of events. So we could reframe our need for notification as a need to listen for a *completion* (or *continuation*) event emitted by `foo(..)`.

Note: Whether you call it a “completion event” or a “continuation event” depends on your perspective. Is the focus more on what happens with `foo(..)`, or what happens *after* `foo(..)` finishes? Both perspectives are accurate and useful. The event notification tells us that `foo(..)` has *completed*, but also that it's OK to *continue* with the next step. Indeed, the callback you pass to be called for the event notification is itself what we've previously called a *continuation*. Because *completion event* is a bit more focused on the `foo(..)`, which more has our attention at present, we slightly favor *completion event* for the rest of this text.

With callbacks, the “notification” would be our callback invoked by the task (`foo(..)`). But with Promises, we turn the relationship around, and expect that we can listen for an event from `foo(..)`, and when notified, proceed accordingly.

First, consider some pseudocode:

```
foo(x) {  
    // start doing something that could take a while  
}  
  
foo( 42 )  
  
on (foo "completion") {  
    // now we can do the next step!  
}  
  
on (foo "error") {  
    // oops, something went wrong in `foo(..)`  
}
```

We call `foo(..)` and then we set up two event listeners, one for “completion” and one for “error” – the two possible *final* outcomes of the `foo(..)` call. In essence, `foo(..)` doesn't even appear to be aware that the calling code has subscribed to these events, which makes for a very nice *separation of concerns*.

Unfortunately, such code would require some “magic” of the JS environment that doesn't exist (and would likely be a bit impractical). Here's the more natural way we could express that in JS:

```
function foo(x) {  
    // start doing something that could take a while
```

```

    // make a 'listener' event notification
    // capability to return

    return listener;
}

var evt = foo( 42 );

evt.on( "completion", function(){
    // now we can do the next step!
} );

evt.on( "failure", function(err){
    // oops, something went wrong in `foo(..)`
} );

```

`foo(..)` expressly creates an event subscription capability to return back, and the calling code receives and registers the two event handlers against it.

The inversion from normal callback-oriented code should be obvious, and it's intentional. Instead of passing the callbacks to `foo(..)`, it returns an event capability we call `evt`, which receives the callbacks.

But if you recall from Chapter 2, callbacks themselves represent an *inversion of control*. So inverting the callback pattern is actually an *inversion of inversion*, or an *uninversion of control* – restoring control back to the calling code where we wanted it to be in the first place.

One important benefit is that multiple separate parts of the code can be given the event listening capability, and they can all independently be notified of when `foo(..)` completes to perform subsequent steps after its completion:

```

var evt = foo( 42 );

// let `bar(..)` listen to `foo(..)`'s completion
bar( evt );

// also, let `baz(..)` listen to `foo(..)`'s completion
baz( evt );

```

Uninversion of control enables a nicer *separation of concerns*, where `bar(..)` and `baz(..)` don't need to be involved in how `foo(..)` is called. Similarly, `foo(..)` doesn't need to know or care that `bar(..)` and `baz(..)` exist or are waiting to be notified when `foo(..)` completes.

Essentially, this `evt` object is a neutral third-party negotiation between the separate concerns.

Promise “Events” As you may have guessed by now, the `evt` event listening capability is an analogy for a Promise.

In a Promise-based approach, the previous snippet would have `foo(..)` creating and returning a `Promise` instance, and that promise would then be passed to `bar(..)` and `baz(..)`.

Note: The Promise resolution “events” we listen for aren’t strictly events (though they certainly behave like events for these purposes), and they’re not typically called “completion” or “error”. Instead, we use `then(..)` to register a “then” event. Or perhaps more precisely, `then(..)` registers “fulfillment” and/or “rejection” event(s), though we don’t see those terms used explicitly in the code.

Consider:

```
function foo(x) {
  // start doing something that could take a while

  // construct and return a promise
  return new Promise( function(resolve,reject){
    // eventually, call `resolve(..)` or `reject(..)`,
    // which are the resolution callbacks for
    // the promise.
  } );
}

var p = foo( 42 );

bar( p );

baz( p );
```

Note: The pattern shown with `new Promise(function(..){ .. })` is generally called the “revealing constructor”. The function passed in is executed immediately (not async deferred, as callbacks to `then(..)` are), and it’s provided two parameters, which in this case we’ve named `resolve` and `reject`. These are the resolution functions for the promise. `resolve(..)` generally signals fulfillment, and `reject(..)` signals rejection.

You can probably guess what the internals of `bar(..)` and `baz(..)` might look like:

```
function bar(fooPromise) {
  // listen for `foo(..)` to complete
  fooPromise.then(
    function(){
      // `foo(..)` has now finished, so
      // do `bar(..)`'s task
```

```

    },
    function(){
        // oops, something went wrong in `foo(..)`
    }
);

}

// ditto for `baz(..)`

```

Promise resolution doesn't necessarily need to involve sending along a message, as it did when we were examining Promises as *future values*. It can just be a flow-control signal, as used in the previous snippet.

Another way to approach this is:

```

function bar() {
    // `foo(..)` has definitely finished, so
    // do `bar(..)`'s task
}

function oopsBar() {
    // oops, something went wrong in `foo(..)`,
    // so `bar(..)` didn't run
}

// ditto for `baz()` and `oopsBaz()`

var p = foo( 42 );

```

```

p.then( bar, oopsBar );

```

```

p.then( baz, oopsBaz );

```

Note: If you've seen Promise-based coding before, you might be tempted to believe that the last two lines of that code could be written as `p.then(..).then(..)`, using chaining, rather than `p.then(..); p.then(..)`. That would have an entirely different behavior, so be careful! The difference might not be clear right now, but it's actually a different async pattern than we've seen thus far: splitting/forking. Don't worry! We'll come back to this point later in this chapter.

Instead of passing the `p` promise to `bar(..)` and `baz(..)`, we use the promise to control when `bar(..)` and `baz(..)` will get executed, if ever. The primary difference is in the error handling.

In the first snippet's approach, `bar(..)` is called regardless of whether `foo(..)` succeeds or fails, and it handles its own fallback logic if it's notified that `foo(..)` failed. The same is true for `baz(..)`, obviously.

In the second snippet, `bar(..)` only gets called if `foo(..)` succeeds, and otherwise `oopsBar(..)` gets called. Ditto for `baz(..)`.

Neither approach is *correct* per se. There will be cases where one is preferred over the other.

In either case, the promise `p` that comes back from `foo(..)` is used to control what happens next.

Moreover, the fact that both snippets end up calling `then(..)` twice against the same promise `p` illustrates the point made earlier, which is that Promises (once resolved) retain their same resolution (fulfillment or rejection) forever, and can subsequently be observed as many times as necessary.

Whenever `p` is resolved, the next step will always be the same, both *now* and *later*.

Thenable Duck Typing

In Promises-land, an important detail is how to know for sure if some value is a genuine Promise or not. Or more directly, is it a value that will behave like a Promise?

Given that Promises are constructed by the `new Promise(..)` syntax, you might think that `p instanceof Promise` would be an acceptable check. But unfortunately, there are a number of reasons that's not totally sufficient.

Mainly, you can receive a Promise value from another browser window (iframe, etc.), which would have its own Promise different from the one in the current window/frame, and that check would fail to identify the Promise instance.

Moreover, a library or framework may choose to vend its own Promises and not use the native ES6 `Promise` implementation to do so. In fact, you may very well be using Promises with libraries in older browsers that have no `Promise` at all.

When we discuss Promise resolution processes later in this chapter, it will become more obvious why a non-genuine-but-Promise-like value would still be very important to be able to recognize and assimilate. But for now, just take my word for it that it's a critical piece of the puzzle.

As such, it was decided that the way to recognize a Promise (or something that behaves like a Promise) would be to define something called a “thenable” as any object or function which has a `then(..)` method on it. It is assumed that any such value is a Promise-conforming thenable.

The general term for “type checks” that make assumptions about a value’s “type” based on its shape (what properties are present) is called “duck typing” – “If it looks like a duck, and quacks like a duck, it must be a duck” (see the *Types & Grammar* title of this book series). So the duck typing check for a thenable would roughly be:

```

if (
  p !== null &&
  (
    typeof p === "object" ||
    typeof p === "function"
  ) &&
  typeof p.then === "function"
) {
  // assume it's a thenable!
}
else {
  // not a thenable
}

```

Yuck! Setting aside the fact that this logic is a bit ugly to implement in various places, there's something deeper and more troubling going on.

If you try to fulfill a Promise with any object/function value that happens to have a `then(..)` function on it, but you weren't intending it to be treated as a Promise/thenable, you're out of luck, because it will automatically be recognized as thenable and treated with special rules (see later in the chapter).

This is even true if you didn't realize the value has a `then(..)` on it. For example:

```

var o = { then: function(){ } };

// make `v` be `[[Prototype]]`-linked to `o`
var v = Object.create( o );

v.someStuff = "cool";
v.otherStuff = "not so cool";

v.hasOwnProperty( "then" );      // false

```

`v` doesn't look like a Promise or thenable at all. It's just a plain object with some properties on it. You're probably just intending to send that value around like any other object.

But unknown to you, `v` is also `[[Prototype]]`-linked (see the *this & Object Prototypes* title of this book series) to another object `o`, which happens to have a `then(..)` on it. So the thenable duck typing checks will think and assume `v` is a thenable. Uh oh.

It doesn't even need to be something as directly intentional as that:

```

Object.prototype.then = function(){};
Array.prototype.then = function(){};

```

```
var v1 = { hello: "world" };
var v2 = [ "Hello", "World" ];
```

Both `v1` and `v2` will be assumed to be thenables. You can't control or predict if any other code accidentally or maliciously adds `then(..)` to `Object.prototype`, `Array.prototype`, or any of the other native prototypes. And if what's specified is a function that doesn't call either of its parameters as callbacks, then any Promise resolved with such a value will just silently hang forever! Crazy.

Sound implausible or unlikely? Perhaps.

But keep in mind that there were several well-known non-Promise libraries pre-existing in the community prior to ES6 that happened to already have a method on them called `then(..)`. Some of those libraries chose to rename their own methods to avoid collision (that sucks!). Others have simply been relegated to the unfortunate status of “incompatible with Promise-based coding” in reward for their inability to change to get out of the way.

The standards decision to hijack the previously nonreserved – and completely general-purpose sounding – `then` property name means that no value (or any of its delegates), either past, present, or future, can have a `then(..)` function present, either on purpose or by accident, or that value will be confused for a thenable in Promises systems, which will probably create bugs that are really hard to track down.

Warning: I do not like how we ended up with duck typing of thenables for Promise recognition. There were other options, such as “branding” or even “anti-branding”; what we got seems like a worst-case compromise. But it's not all doom and gloom. Thenable duck typing can be helpful, as we'll see later. Just beware that thenable duck typing can be hazardous if it incorrectly identifies something as a Promise that isn't.

Promise Trust

We've now seen two strong analogies that explain different aspects of what Promises can do for our async code. But if we stop there, we've missed perhaps the single most important characteristic that the Promise pattern establishes: trust.

Whereas the *future values* and *completion events* analogies play out explicitly in the code patterns we've explored, it won't be entirely obvious why or how Promises are designed to solve all of the *inversion of control* trust issues we laid out in the “Trust Issues” section of Chapter 2. But with a little digging, we can uncover some important guarantees that restore the confidence in async coding that Chapter 2 tore down!

Let's start by reviewing the trust issues with callbacks-only coding. When you pass a callback to a utility `foo(..)`, it might:

- Call the callback too early
- Call the callback too late (or never)
- Call the callback too few or too many times
- Fail to pass along any necessary environment/parameters
- Swallow any errors/exceptions that may happen

The characteristics of Promises are intentionally designed to provide useful, repeatable answers to all these concerns.

Calling Too Early

Primarily, this is a concern of whether code can introduce Zalgo-like effects (see Chapter 2), where sometimes a task finishes synchronously and sometimes asynchronously, which can lead to race conditions.

Promises by definition cannot be susceptible to this concern, because even an immediately fulfilled Promise (like `new Promise(function(resolve){ resolve(42); })`) cannot be *observed* synchronously.

That is, when you call `then(..)` on a Promise, even if that Promise was already resolved, the callback you provide to `then(..)` will **always** be called asynchronously (for more on this, refer back to “Jobs” in Chapter 1).

No more need to insert your own `setTimeout(..,0)` hacks. Promises prevent Zalgo automatically.

Calling Too Late

Similar to the previous point, a Promise’s `then(..)` registered observation callbacks are automatically scheduled when either `resolve(..)` or `reject(..)` are called by the Promise creation capability. Those scheduled callbacks will predictably be fired at the next asynchronous moment (see “Jobs” in Chapter 1).

It’s not possible for synchronous observation, so it’s not possible for a synchronous chain of tasks to run in such a way to in effect “delay” another callback from happening as expected. That is, when a Promise is resolved, all `then(..)` registered callbacks on it will be called, in order, immediately at the next asynchronous opportunity (again, see “Jobs” in Chapter 1), and nothing that happens inside of one of those callbacks can affect/delay the calling of the other callbacks.

For example:

```
p.then( function(){
  p.then( function(){
    console.log( "C" );
  } );
  console.log( "A" );
} );
```

```

p.then( function(){
  console.log( "B" );
} );
// A B C

```

Here, "C" cannot interrupt and precede "B", by virtue of how Promises are defined to operate.

Promise Scheduling Quirks It's important to note, though, that there are lots of nuances of scheduling where the relative ordering between callbacks chained off two separate Promises is not reliably predictable.

If two promises p1 and p2 are both already resolved, it should be true that p1.then(..); p2.then(..) would end up calling the callback(s) for p1 before the ones for p2. But there are subtle cases where that might not be true, such as the following:

```

var p3 = new Promise( function(resolve,reject){
  resolve( "B" );
} );

var p1 = new Promise( function(resolve,reject){
  resolve( p3 );
} );

var p2 = new Promise( function(resolve,reject){
  resolve( "A" );
} );

p1.then( function(v){
  console.log( v );
} );

p2.then( function(v){
  console.log( v );
} );

// A B <-- not B A as you might expect

```

We'll cover this more later, but as you can see, p1 is resolved not with an immediate value, but with another promise p3 which is itself resolved with the value "B". The specified behavior is to *unwrap* p3 into p1, but asynchronously, so p1's callback(s) are *behind* p2's callback(s) in the asynchronous Job queue (see Chapter 1).

To avoid such nuanced nightmares, you should never rely on anything about the ordering/scheduling of callbacks across Promises. In fact, a good practice is not

to code in such a way where the ordering of multiple callbacks matters at all. Avoid that if you can.

Never Calling the Callback

This is a very common concern. It's addressable in several ways with Promises.

First, nothing (not even a JS error) can prevent a Promise from notifying you of its resolution (if it's resolved). If you register both fulfillment and rejection callbacks for a Promise, and the Promise gets resolved, one of the two callbacks will always be called.

Of course, if your callbacks themselves have JS errors, you may not see the outcome you expect, but the callback will in fact have been called. We'll cover later how to be notified of an error in your callback, because even those don't get swallowed.

But what if the Promise itself never gets resolved either way? Even that is a condition that Promises provide an answer for, using a higher level abstraction called a "race":

```
// a utility for timing out a Promise
function timeoutPromise(delay) {
  return new Promise( function(resolve,reject){
    setTimeout( function(){
      reject( "Timeout!" );
    }, delay );
  } );
}

// setup a timeout for `foo()`
Promise.race( [
  foo(), // attempt `foo()`
  timeoutPromise( 3000 ) // give it 3 seconds
] )
.then(
  function(){
    // `foo(..)` fulfilled in time!
  },
  function(err){
    // either `foo()` rejected, or it just
    // didn't finish in time, so inspect
    // `err` to know which
  }
);

```

There are more details to consider with this Promise timeout pattern, but we'll come back to it later.

Importantly, we can ensure a signal as to the outcome of `foo()`, to prevent it from hanging our program indefinitely.

Calling Too Few or Too Many Times

By definition, *one* is the appropriate number of times for the callback to be called. The “too few” case would be zero calls, which is the same as the “never” case we just examined.

The “too many” case is easy to explain. Promises are defined so that they can only be resolved once. If for some reason the Promise creation code tries to call `resolve(..)` or `reject(..)` multiple times, or tries to call both, the Promise will accept only the first resolution, and will silently ignore any subsequent attempts.

Because a Promise can only be resolved once, any `then(..)` registered callbacks will only ever be called once (each).

Of course, if you register the same callback more than once, (e.g., `p.then(f); p.then(f);`), it’ll be called as many times as it was registered. The guarantee that a response function is called only once does not prevent you from shooting yourself in the foot.

Failing to Pass Along Any Parameters/Environment

Promises can have, at most, one resolution value (fulfillment or rejection).

If you don’t explicitly resolve with a value either way, the value is `undefined`, as is typical in JS. But whatever the value, it will always be passed to all registered (and appropriate: fulfillment or rejection) callbacks, either *now* or in the future.

Something to be aware of: If you call `resolve(..)` or `reject(..)` with multiple parameters, all subsequent parameters beyond the first will be silently ignored. Although that might seem a violation of the guarantee we just described, it’s not exactly, because it constitutes an invalid usage of the Promise mechanism. Other invalid usages of the API (such as calling `resolve(..)` multiple times) are similarly *protected*, so the Promise behavior here is consistent (if not a tiny bit frustrating).

If you want to pass along multiple values, you must wrap them in another single value that you pass, such as an `array` or an `object`.

As for environment, functions in JS always retain their closure of the scope in which they’re defined (see the *Scope & Closures* title of this series), so they of course would continue to have access to whatever surrounding state you provide. Of course, the same is true of callbacks-only design, so this isn’t a specific augmentation of benefit from Promises – but it’s a guarantee we can rely on nonetheless.

Swallowing Any Errors/Exceptions

In the base sense, this is a restatement of the previous point. If you reject a Promise with a *reason* (aka error message), that value is passed to the rejection callback(s).

But there's something much bigger at play here. If at any point in the creation of a Promise, or in the observation of its resolution, a JS exception error occurs, such as a `TypeError` or `ReferenceError`, that exception will be caught, and it will force the Promise in question to become rejected.

For example:

```
var p = new Promise( function(resolve,reject){
  foo.bar(); // `foo` is not defined, so error!
  resolve( 42 ); // never gets here :(
} );

p.then(
  function fulfilled(){
    // never gets here :(
  },
  function rejected(err){
    // `err` will be a `TypeError` exception object
    // from the `foo.bar()` line.
  }
);
```

The JS exception that occurs from `foo.bar()` becomes a Promise rejection that you can catch and respond to.

This is an important detail, because it effectively solves another potential Zalgo moment, which is that errors could create a synchronous reaction whereas non-errors would be asynchronous. Promises turn even JS exceptions into asynchronous behavior, thereby reducing the race condition chances greatly.

But what happens if a Promise is fulfilled, but there's a JS exception error during the observation (in a `then(..)` registered callback)? Even those aren't lost, but you may find how they're handled a bit surprising, until you dig in a little deeper:

```
var p = new Promise( function(resolve,reject){
  resolve( 42 );
} );

p.then(
  function fulfilled(msg){
    foo.bar();
    console.log( msg ); // never gets here :(
  }
);
```

```

},
function rejected(err){
  // never gets here either :(
}
);

```

Wait, that makes it seem like the exception from `foo.bar()` really did get swallowed. Never fear, it didn't. But something deeper is wrong, which is that we've failed to listen for it. The `p.then(..)` call itself returns another promise, and it's *that* promise that will be rejected with the `TypeError` exception.

Why couldn't it just call the error handler we have defined there? Seems like a logical behavior on the surface. But it would violate the fundamental principle that Promises are **immutable** once resolved. `p` was already fulfilled to the value `42`, so it can't later be changed to a rejection just because there's an error in observing `p`'s resolution.

Besides the principle violation, such behavior could wreak havoc, if say there were multiple `then(..)` registered callbacks on the promise `p`, because some would get called and others wouldn't, and it would be very opaque as to why.

Trustable Promise?

There's one last detail to examine to establish trust based on the Promise pattern.

You've no doubt noticed that Promises don't get rid of callbacks at all. They just change where the callback is passed to. Instead of passing a callback to `foo(..)`, we get *something* (ostensibly a genuine Promise) back from `foo(..)`, and we pass the callback to that *something* instead.

But why would this be any more trustable than just callbacks alone? How can we be sure the *something* we get back is in fact a trustable Promise? Isn't it basically all just a house of cards where we can trust only because we already trusted?

One of the most important, but often overlooked, details of Promises is that they have a solution to this issue as well. Included with the native ES6 `Promise` implementation is `Promise.resolve(..)`.

If you pass an immediate, non-Promise, non-thenable value to `Promise.resolve(..)`, you get a promise that's fulfilled with that value. In other words, these two promises `p1` and `p2` will behave basically identically:

```

var p1 = new Promise( function(resolve,reject){
  resolve( 42 );
}
);

var p2 = Promise.resolve( 42 );

```

But if you pass a genuine Promise to `Promise.resolve(..)`, you just get the same promise back:

```
var p1 = Promise.resolve( 42 );

var p2 = Promise.resolve( p1 );

p1 === p2; // true
```

Even more importantly, if you pass a non-Promise thenable value to `Promise.resolve(..)`, it will attempt to unwrap that value, and the unwrapping will keep going until a concrete final non-Promise-like value is extracted.

Recall our previous discussion of thenables?

Consider:

```
var p = {
  then: function(cb) {
    cb( 42 );
  }
};

// this works OK, but only by good fortune
p
.then(
  function fulfilled(val){
    console.log( val ); // 42
  },
  function rejected(err){
    // never gets here
  }
);
```

This `p` is a thenable, but it's not a genuine Promise. Luckily, it's reasonable, as most will be. But what if you got back instead something that looked like:

```
var p = {
  then: function(cb,errcb) {
    cb( 42 );
    errcb( "evil laugh" );
  }
};

p
.then(
  function fulfilled(val){
    console.log( val ); // 42
```

```

},
function rejected(err){
  // oops, shouldn't have run
  console.log( err ); // evil laugh
}
);

```

This `p` is a thenable but it's not so well behaved of a promise. Is it malicious? Or is it just ignorant of how Promises should work? It doesn't really matter, to be honest. In either case, it's not trustable as is.

Nonetheless, we can pass either of these versions of `p` to `Promise.resolve(..)`, and we'll get the normalized, safe result we'd expect:

```

Promise.resolve( p )
.then(
  function fulfilled(val){
    console.log( val ); // 42
  },
  function rejected(err){
    // never gets here
  }
);

```

`Promise.resolve(..)` will accept any thenable, and will unwrap it to its non-thenable value. But you get back from `Promise.resolve(..)` a real, genuine Promise in its place, **one that you can trust**. If what you passed in is already a genuine Promise, you just get it right back, so there's no downside at all to filtering through `Promise.resolve(..)` to gain trust.

So let's say we're calling a `foo(..)` utility and we're not sure we can trust its return value to be a well-behaving Promise, but we know it's at least a thenable. `Promise.resolve(..)` will give us a trustable Promise wrapper to chain off of:

```

// don't just do this:
foo( 42 )
.then( function(v){
  console.log( v );
} );

// instead, do this:
Promise.resolve( foo( 42 ) )
.then( function(v){
  console.log( v );
} );

```

Note: Another beneficial side effect of wrapping `Promise.resolve(..)` around any function's return value (thenable or not) is that it's an easy way to normalize that function call into a well-behaving async task. If `foo(42)` returns an imme-

diate value sometimes, or a Promise other times, `Promise.resolve(foo(42))` makes sure it's always a Promise result. And avoiding Zalgo makes for much better code.

Trust Built

Hopefully the previous discussion now fully “resolves” (pun intended) in your mind why the Promise is trustable, and more importantly, why that trust is so critical in building robust, maintainable software.

Can you write async code in JS without trust? Of course you can. We JS developers have been coding async with nothing but callbacks for nearly two decades.

But once you start questioning just how much you can trust the mechanisms you build upon to actually be predictable and reliable, you start to realize callbacks have a pretty shaky trust foundation.

Promises are a pattern that augments callbacks with trustable semantics, so that the behavior is more reason-able and more reliable. By uninverting the *inversion of control* of callbacks, we place the control with a trustable system (Promises) that was designed specifically to bring sanity to our async.

Chain Flow

We've hinted at this a couple of times already, but Promises are not just a mechanism for a single-step *this-then-that* sort of operation. That's the building block, of course, but it turns out we can string multiple Promises together to represent a sequence of async steps.

The key to making this work is built on two behaviors intrinsic to Promises:

- Every time you call `then(..)` on a Promise, it creates and returns a new Promise, which we can *chain* with.
- Whatever value you return from the `then(..)` call's fulfillment callback (the first parameter) is automatically set as the fulfillment of the *chained* Promise (from the first point).

Let's first illustrate what that means, and *then* we'll derive how that helps us create async sequences of flow control. Consider the following:

```
var p = Promise.resolve( 21 );

var p2 = p.then( function(v){
  console.log( v ); // 21

  // fulfill `p2` with value `42`
  return v * 2;
} );
```

```
// chain off `p2`  
p2.then( function(v){  
  console.log( v );  // 42  
} );
```

By returning `v * 2` (i.e., 42), we fulfill the `p2` promise that the first `then(..)` call created and returned. When `p2`'s `then(..)` call runs, it's receiving the fulfillment from the `return v * 2` statement. Of course, `p2.then(..)` creates yet another promise, which we could have stored in a `p3` variable.

But it's a little annoying to have to create an intermediate variable `p2` (or `p3`, etc.). Thankfully, we can easily just chain these together:

```
var p = Promise.resolve( 21 );  
  
p  
.then( function(v){  
  console.log( v );  // 21  
  
  // fulfill the chained promise with value `42`  
  return v * 2;  
} )  
// here's the chained promise  
.then( function(v){  
  console.log( v );  // 42  
} );
```

So now the first `then(..)` is the first step in an async sequence, and the second `then(..)` is the second step. This could keep going for as long as you needed it to extend. Just keep chaining off a previous `then(..)` with each automatically created Promise.

But there's something missing here. What if we want step 2 to wait for step 1 to do something asynchronous? We're using an immediate `return` statement, which immediately fulfills the chained promise.

The key to making a Promise sequence truly async capable at every step is to recall how `Promise.resolve(..)` operates when what you pass to it is a Promise or thenable instead of a final value. `Promise.resolve(..)` directly returns a received genuine Promise, or it unwraps the value of a received thenable – and keeps going recursively while it keeps unwrapping thenables.

The same sort of unwrapping happens if you `return` a thenable or Promise from the fulfillment (or rejection) handler. Consider:

```
var p = Promise.resolve( 21 );  
  
p.then( function(v){  
  console.log( v );  // 21
```

```

    // create a promise and return it
    return new Promise( function(resolve,reject){
        // fulfill with value `42`
        resolve( v * 2 );
    } );
}
.then( function(v){
    console.log( v );    // 42
} );

```

Even though we wrapped 42 up in a promise that we returned, it still got unwrapped and ended up as the resolution of the chained promise, such that the second `then(..)` still received 42. If we introduce asynchrony to that wrapping promise, everything still nicely works the same:

```

var p = Promise.resolve( 21 );

p.then( function(v){
    console.log( v );    // 21

    // create a promise to return
    return new Promise( function(resolve,reject){
        // introduce asynchrony!
        setTimeout( function(){
            // fulfill with value `42`
            resolve( v * 2 );
        }, 100 );
    } );
}
.then( function(v){
    // runs after the 100ms delay in the previous step
    console.log( v );    // 42
} );

```

That's incredibly powerful! Now we can construct a sequence of however many async steps we want, and each step can delay the next step (or not!), as necessary.

Of course, the value passing from step to step in these examples is optional. If you don't return an explicit value, an implicit `undefined` is assumed, and the promises still chain together the same way. Each Promise resolution is thus just a signal to proceed to the next step.

To further the chain illustration, let's generalize a delay-Promise creation (without resolution messages) into a utility we can reuse for multiple steps:

```

function delay(time) {
    return new Promise( function(resolve,reject){

```

```

        setTimeout( resolve, time );
    } );
}

delay( 100 ) // step 1
.then( function STEP2(){
    console.log( "step 2 (after 100ms)" );
    return delay( 200 );
} )
.then( function STEP3(){
    console.log( "step 3 (after another 200ms)" );
} )
.then( function STEP4(){
    console.log( "step 4 (next Job)" );
    return delay( 50 );
} )
.then( function STEP5(){
    console.log( "step 5 (after another 50ms)" );
} )
...

```

Calling `delay(200)` creates a promise that will fulfill in 200ms, and then we return that from the first `then(..)` fulfillment callback, which causes the second `then(..)`'s promise to wait on that 200ms promise.

Note: As described, technically there are two promises in that interchange: the 200ms-delay promise and the chained promise that the second `then(..)` chains from. But you may find it easier to mentally combine these two promises together, because the Promise mechanism automatically merges their states for you. In that respect, you could think of `return delay(200)` as creating a promise that replaces the earlier-returned chained promise.

To be honest, though, sequences of delays with no message passing isn't a terribly useful example of Promise flow control. Let's look at a scenario that's a little more practical.

Instead of timers, let's consider making Ajax requests:

```

// assume an `ajax( {url}, {callback} )` utility

// Promise-aware ajax
function request(url) {
    return new Promise( function(resolve,reject){
        // the `ajax(..)` callback should be our
        // promise's `resolve(..)` function
        ajax( url, resolve );
    } );
}

```

We first define a `request(..)` utility that constructs a promise to represent the completion of the `ajax(..)` call:

```
request( "http://some.url.1/" )
  .then( function(response1){
    return request( "http://some.url.2/?v=" + response1 );
  })
  .then( function(response2){
    console.log( response2 );
  });
}
```

Note: Developers commonly encounter situations in which they want to do Promise-aware async flow control with utilities that are not themselves Promise-enabled (like `ajax(..)` here, which expects a callback). Although the native ES6 `Promise` mechanism doesn't automatically solve this pattern for us, practically all Promise libraries *do*. They usually call this process "lifting" or "promisifying" or some variation thereof. We'll come back to this technique later.

Using the Promise-returning `request(..)`, we create the first step in our chain implicitly by calling it with the first URL, and chain off that returned promise with the first `then(..)`.

Once `response1` comes back, we use that value to construct a second URL, and make a second `request(..)` call. That second `request(..)` promise is `returned` so that the third step in our async flow control waits for that Ajax call to complete. Finally, we print `response2` once it returns.

The Promise chain we construct is not only a flow control that expresses a multistep async sequence, but it also acts as a message channel to propagate messages from step to step.

What if something went wrong in one of the steps of the Promise chain? An error/exception is on a per-Promise basis, which means it's possible to catch such an error at any point in the chain, and that catching acts to sort of "reset" the chain back to normal operation at that point:

```
// step 1:
request( "http://some.url.1/" )

// step 2:
.then( function(response1){
  foo.bar(); // undefined, error!

  // never gets here
  return request( "http://some.url.2/?v=" + response1 );
})
}

// step 3:
.then(
```

```

        function fulfilled(response2){
            // never gets here
        },
        // rejection handler to catch the error
        function rejected(err){
            console.log( err ); // `TypeError` from `foo.bar()` error
            return 42;
        }
    )

    // step 4:
    .then( function(msg){
        console.log( msg ); // 42
    } );
}

```

When the error occurs in step 2, the rejection handler in step 3 catches it. The return value (42 in this snippet), if any, from that rejection handler fulfills the promise for the next step (4), such that the chain is now back in a fulfillment state.

Note: As we discussed earlier, when returning a promise from a fulfillment handler, it's unwrapped and can delay the next step. That's also true for returning promises from rejection handlers, such that if the `return 42` in step 3 instead returned a promise, that promise could delay step 4. A thrown exception inside either the fulfillment or rejection handler of a `then(..)` call causes the next (chained) promise to be immediately rejected with that exception.

If you call `then(..)` on a promise, and you only pass a fulfillment handler to it, an assumed rejection handler is substituted:

```

var p = new Promise( function(resolve,reject){
    reject( "Oops" );
} );

var p2 = p.then(
    function fulfilled(){
        // never gets here
    }
    // assumed rejection handler, if omitted or
    // any other non-function value passed
    // function(err) {
    //     throw err;
    // }
);

```

As you can see, the assumed rejection handler simply rethrows the error, which ends up forcing `p2` (the chained promise) to reject with the same error reason. In essence, this allows the error to continue propagating along a Promise chain

until an explicitly defined rejection handler is encountered.

Note: We'll cover more details of error handling with Promises a little later, because there are other nuanced details to be concerned about.

If a proper valid function is not passed as the fulfillment handler parameter to `then(..)`, there's also a default handler substituted:

```
var p = Promise.resolve( 42 );

p.then(
  // assumed fulfillment handler, if omitted or
  // any other non-function value passed
  // function(v) {
  //   return v;
  // }
  null,
  function rejected(err){
    // never gets here
  }
);
```

As you can see, the default fulfillment handler simply passes whatever value it receives along to the next step (Promise).

Note: The `then(null, function(err){ .. })` pattern – only handling rejections (if any) but letting fulfillments pass through – has a shortcut in the API: `catch(function(err){ .. })`. We'll cover `catch(..)` more fully in the next section.

Let's review briefly the intrinsic behaviors of Promises that enable chaining flow control:

- A `then(..)` call against one Promise automatically produces a new Promise to return from the call.
- Inside the fulfillment/rejection handlers, if you return a value or an exception is thrown, the new returned (chainable) Promise is resolved accordingly.
- If the fulfillment or rejection handler returns a Promise, it is unwrapped, so that whatever its resolution is will become the resolution of the chained Promise returned from the current `then(..)`.

While chaining flow control is helpful, it's probably most accurate to think of it as a side benefit of how Promises compose (combine) together, rather than the main intent. As we've discussed in detail several times already, Promises normalize asynchrony and encapsulate time-dependent value state, and *that* is what lets us chain them together in this useful way.

Certainly, the sequential expressiveness of the chain (this-then-this-then-this...) is a big improvement over the tangled mess of callbacks as we identified in Chap-

ter 2. But there's still a fair amount of boilerplate (`then(..)` and `function(){ .. }`) to wade through. In the next chapter, we'll see a significantly nicer pattern for sequential flow control expressivity, with generators.

Terminology: Resolve, Fulfill, and Reject

There's some slight confusion around the terms "resolve," "fulfill," and "reject" that we need to clear up, before you get too much deeper into learning about Promises. Let's first consider the `Promise(..)` constructor:

```
var p = new Promise( function(X,Y){  
  // X() for fulfillment  
  // Y() for rejection  
} );
```

As you can see, two callbacks (here labeled `X` and `Y`) are provided. The first is *usually* used to mark the Promise as fulfilled, and the second *always* marks the Promise as rejected. But what's the "usually" about, and what does that imply about accurately naming those parameters?

Ultimately, it's just your user code and the identifier names aren't interpreted by the engine to mean anything, so it doesn't *technically* matter; `foo(..)` and `bar(..)` are equally functional. But the words you use can affect not only how you are thinking about the code, but how other developers on your team will think about it. Thinking wrongly about carefully orchestrated async code is almost surely going to be worse than the spaghetti-callback alternatives.

So it actually does kind of matter what you call them.

The second parameter is easy to decide. Almost all literature uses `reject(..)` as its name, and because that's exactly (and only!) what it does, that's a very good choice for the name. I'd strongly recommend you always use `reject(..)`.

But there's a little more ambiguity around the first parameter, which in Promise literature is often labeled `resolve(..)`. That word is obviously related to "resolution," which is what's used across the literature (including this book) to describe setting a final value/state to a Promise. We've already used "resolve the Promise" several times to mean either fulfilling or rejecting the Promise.

But if this parameter seems to be used to specifically fulfill the Promise, why shouldn't we call it `fulfill(..)` instead of `resolve(..)` to be more accurate? To answer that question, let's also take a look at two of the `Promise` API methods:

```
var fulfilledPr = Promise.resolve( 42 );  
  
var rejectedPr = Promise.reject( "Oops" );
```

`Promise.resolve(..)` creates a Promise that's resolved to the value given to it. In this example, 42 is a normal, non-Promise, non-thenable value, so the fulfilled

promise `fulfilledPr` is created for the value 42. `Promise.reject("Oops")` creates the rejected promise `rejectedPr` for the reason "Oops".

Let's now illustrate why the word "resolve" (such as in `Promise.resolve(..)`) is unambiguous and indeed more accurate, if used explicitly in a context that could result in either fulfillment or rejection:

```
var rejectedTh = {
  then: function(resolved,rejected) {
    rejected( "Oops" );
  }
};

var rejectedPr = Promise.resolve( rejectedTh );
```

As we discussed earlier in this chapter, `Promise.resolve(..)` will return a received genuine Promise directly, or unwrap a received thenable. If that thenable unwrapping reveals a rejected state, the Promise returned from `Promise.resolve(..)` is in fact in that same rejected state.

So `Promise.resolve(..)` is a good, accurate name for the API method, because it can actually result in either fulfillment or rejection.

The first callback parameter of the `Promise(..)` constructor will unwrap either a thenable (identically to `Promise.resolve(..)`) or a genuine Promise:

```
var rejectedPr = new Promise( function(resolve,reject){
  // resolve this promise with a rejected promise
  resolve( Promise.reject( "Oops" ) );
} );

rejectedPr.then(
  function fulfilled(){}
  // never gets here
),
  function rejected(err){
    console.log( err ); // "Oops"
}
);
```

It should be clear now that `resolve(..)` is the appropriate name for the first callback parameter of the `Promise(..)` constructor.

Warning: The previously mentioned `reject(..)` does **not** do the unwrapping that `resolve(..)` does. If you pass a Promise/thenable value to `reject(..)`, that untouched value will be set as the rejection reason. A subsequent rejection handler would receive the actual Promise/thenable you passed to `reject(..)`, not its underlying immediate value.

But now let's turn our attention to the callbacks provided to `then(..)`. What

should they be called (both in literature and in code)? I would suggest `fulfilled(..)` and `rejected(..)`:

```
function fulfilled(msg) {
  console.log( msg );
}

function rejected(err) {
  console.error( err );
}

p.then(
  fulfilled,
  rejected
);
```

In the case of the first parameter to `then(..)`, it's unambiguously always the fulfillment case, so there's no need for the duality of "resolve" terminology. As a side note, the ES6 specification uses `onFulfilled(..)` and `onRejected(..)` to label these two callbacks, so they are accurate terms.

Error Handling

We've already seen several examples of how Promise rejection – either intentional through calling `reject(..)` or accidental through JS exceptions – allows saner error handling in asynchronous programming. Let's circle back though and be explicit about some of the details that we glossed over.

The most natural form of error handling for most developers is the synchronous `try..catch` construct. Unfortunately, it's synchronous-only, so it fails to help in async code patterns:

```
function foo() {
  setTimeout( function(){
    baz.bar();
  }, 100 );
}

try {
  foo();
  // later throws global error from `baz.bar()`
}
catch (err) {
  // never gets here
}
```

`try..catch` would certainly be nice to have, but it doesn't work across async operations. That is, unless there's some additional environmental support, which

we'll come back to with generators in Chapter 4.

In callbacks, some standards have emerged for patterned error handling, most notably the “error-first callback” style:

```
function foo(cb) {
  setTimeout( function(){
    try {
      var x = baz.bar();
      cb( null, x ); // success!
    }
    catch (err) {
      cb( err );
    }
  }, 100 );
}

foo( function(err,val){
  if (err) {
    console.error( err ); // bummer :(
  }
  else {
    console.log( val );
  }
} );
```

Note: The `try..catch` here works only from the perspective that the `baz.bar()` call will either succeed or fail immediately, synchronously. If `baz.bar()` was itself its own async completing function, any async errors inside it would not be catchable.

The callback we pass to `foo(..)` expects to receive a signal of an error by the reserved first parameter `err`. If present, error is assumed. If not, success is assumed.

This sort of error handling is technically *async capable*, but it doesn't compose well at all. Multiple levels of error-first callbacks woven together with these ubiquitous `if` statement checks inevitably will lead you to the perils of callback hell (see Chapter 2).

So we come back to error handling in Promises, with the rejection handler passed to `then(..)`. Promises don't use the popular “error-first callback” design style, but instead use “split callbacks” style; there's one callback for fulfillment and one for rejection:

```
var p = Promise.reject( "Oops" );

p.then(
  function fulfilled(){
```

```

        // never gets here
    },
    function rejected(err){
        console.log( err ); // "Oops"
    }
);

```

While this pattern of error handling makes fine sense on the surface, the nuances of Promise error handling are often a fair bit more difficult to fully grasp.

Consider:

```

var p = Promise.resolve( 42 );

p.then(
    function fulfilled(msg){
        // numbers don't have string functions,
        // so will throw an error
        console.log( msg.toLowerCase() );
    },
    function rejected(err){
        // never gets here
    }
);

```

If the `msg.toLowerCase()` legitimately throws an error (it does!), why doesn't our error handler get notified? As we explained earlier, it's because *that* error handler is for the `p` promise, which has already been fulfilled with value `42`. The `p` promise is immutable, so the only promise that can be notified of the error is the one returned from `p.then(...)`, which in this case we don't capture.

That should paint a clear picture of why error handling with Promises is error-prone (pun intended). It's far too easy to have errors swallowed, as this is very rarely what you'd intend.

Warning: If you use the Promise API in an invalid way and an error occurs that prevents proper Promise construction, the result will be an immediately thrown exception, **not a rejected Promise**. Some examples of incorrect usage that fail Promise construction: `new Promise(null)`, `Promise.all()`, `Promise.race(42)`, and so on. You can't get a rejected Promise if you don't use the Promise API validly enough to actually construct a Promise in the first place!

Pit of Despair

Jeff Atwood noted years ago: programming languages are often set up in such a way that by default, developers fall into the “pit of despair” (<http://blog.codinghorror.com/falling-into-the-pit-of-success/>) – where accidents are punished – and that you have to try harder to get it right. He

implored us to instead create a “pit of success,” where by default you fall into expected (successful) action, and thus would have to try hard to fail.

Promise error handling is unquestionably “pit of despair” design. By default, it assumes that you want any error to be swallowed by the Promise state, and if you forget to observe that state, the error silently languishes/dies in obscurity – usually despair.

To avoid losing an error to the silence of a forgotten/discarded Promise, some developers have claimed that a “best practice” for Promise chains is to always end your chain with a final `catch(..)`, like:

```
var p = Promise.resolve( 42 );

p.then(
  function fulfilled(msg){
    // numbers don't have string functions,
    // so will throw an error
    console.log( msg.toLowerCase() );
  }
)
.catch( handleErrors );
```

Because we didn’t pass a rejection handler to the `then(..)`, the default handler was substituted, which simply propagates the error to the next promise in the chain. As such, both errors that come into `p`, and errors that come *after* `p` in its resolution (like the `msg.toLowerCase()` one) will filter down to the final `handleErrors(..)`.

Problem solved, right? Not so fast!

What happens if `handleErrors(..)` itself also has an error in it? Who catches that? There’s still yet another unattended promise: the one `catch(..)` returns, which we don’t capture and don’t register a rejection handler for.

You can’t just stick another `catch(..)` on the end of that chain, because it too could fail. The last step in any Promise chain, whatever it is, always has the possibility, even decreasingly so, of dangling with an uncaught error stuck inside an unobserved Promise.

Sound like an impossible conundrum yet?

Uncaught Handling

It’s not exactly an easy problem to solve completely. There are other ways to approach it which many would say are *better*.

Some Promise libraries have added methods for registering something like a “global unhandled rejection” handler, which would be called instead of a globally thrown error. But their solution for how to identify an error as “uncaught” is

to have an arbitrary-length timer, say 3 seconds, running from time of rejection. If a Promise is rejected but no error handler is registered before the timer fires, then it's assumed that you won't ever be registering a handler, so it's "uncaught."

In practice, this has worked well for many libraries, as most usage patterns don't typically call for significant delay between Promise rejection and observation of that rejection. But this pattern is troublesome because 3 seconds is so arbitrary (even if empirical), and also because there are indeed some cases where you want a Promise to hold on to its rejectedness for some indefinite period of time, and you don't really want to have your "uncaught" handler called for all those false positives (not-yet-handled "uncaught errors").

Another more common suggestion is that Promises should have a `done(..)` added to them, which essentially marks the Promise chain as "done." `done(..)` doesn't create and return a Promise, so the callbacks passed to `done(..)` are obviously not wired up to report problems to a chained Promise that doesn't exist.

So what happens instead? It's treated as you might usually expect in uncaught error conditions: any exception inside a `done(..)` rejection handler would be thrown as a global uncaught error (in the developer console, basically):

```
var p = Promise.resolve( 42 );

p.then(
  function fulfilled(msg){
    // numbers don't have string functions,
    // so will throw an error
    console.log( msg.toLowerCase() );
  }
)
.done( null, handleErrors );

// if `handleErrors(..)` caused its own exception, it would
// be thrown globally here
```

This might sound more attractive than the never-ending chain or the arbitrary timeouts. But the biggest problem is that it's not part of the ES6 standard, so no matter how good it sounds, at best it's a lot longer way off from being a reliable and ubiquitous solution.

Are we just stuck, then? Not entirely.

Browsers have a unique capability that our code does not have: they can track and know for sure when any object gets thrown away and garbage collected. So, browsers can track Promise objects, and whenever they get garbage collected, if they have a rejection in them, the browser knows for sure this was a legitimate "uncaught error," and can thus confidently know it should report it to the developer console.

Note: At the time of this writing, both Chrome and Firefox have early attempts at that sort of “uncaught rejection” capability, though support is incomplete at best.

However, if a Promise doesn’t get garbage collected – it’s exceedingly easy for that to accidentally happen through lots of different coding patterns – the browser’s garbage collection sniffing won’t help you know and diagnose that you have a silently rejected Promise laying around.

Is there any other alternative? Yes.

Pit of Success

The following is just theoretical, how Promises *could* be someday changed to behave. I believe it would be far superior to what we currently have. And I think this change would be possible even post-ES6 because I don’t think it would break web compatibility with ES6 Promises. Moreover, it can be polyfilled/prollyfilled in, if you’re careful. Let’s take a look:

- Promises could default to reporting (to the developer console) any rejection, on the next Job or event loop tick, if at that exact moment no error handler has been registered for the Promise.
- For the cases where you want a rejected Promise to hold onto its rejected state for an indefinite amount of time before observing, you could call `defer()`, which suppresses automatic error reporting on that Promise.

If a Promise is rejected, it defaults to noisily reporting that fact to the developer console (instead of defaulting to silence). You can opt out of that reporting either implicitly (by registering an error handler before rejection), or explicitly (with `defer()`). In either case, *you* control the false positives.

Consider:

```
var p = Promise.reject( "Oops" ).defer();

// `foo(..)` is Promise-aware
foo( 42 )
.then(
  function fulfilled(){
    return p;
  },
  function rejected(err){
    // handle `foo(..)` error
  }
);
...

```

When we create `p`, we know we’re going to wait a while to use/observe its rejection, so we call `defer()` – thus no global reporting. `defer()` simply returns

the same promise, for chaining purposes.

The promise returned from `foo(..)` gets an error handler attached *right away*, so it's implicitly opted out and no global reporting for it occurs either.

But the promise returned from the `then(..)` call has no `defer()` or error handler attached, so if it rejects (from inside either resolution handler), then *it* will be reported to the developer console as an uncaught error.

This design is a pit of success. By default, all errors are either handled or reported – what almost all developers in almost all cases would expect. You either have to register a handler or you have to intentionally opt out, and indicate you intend to defer error handling until *later*; you're opting for the extra responsibility in just that specific case.

The only real danger in this approach is if you `defer()` a Promise but then fail to actually ever observe/handle its rejection.

But you had to intentionally call `defer()` to opt into that pit of despair – the default was the pit of success – so there's not much else we could do to save you from your own mistakes.

I think there's still hope for Promise error handling (post-ES6). I hope the powers that be will rethink the situation and consider this alternative. In the meantime, you can implement this yourself (a challenging exercise for the reader!), or use a *smarter* Promise library that does so for you!

Note: This exact model for error handling/reporting is implemented in my `asynquence` Promise abstraction library, which will be discussed in Appendix A of this book.

Promise Patterns

We've already implicitly seen the sequence pattern with Promise chains (this-then-this-then-that flow control) but there are lots of variations on asynchronous patterns that we can build as abstractions on top of Promises. These patterns serve to simplify the expression of async flow control – which helps make our code more reasonable and more maintainable – even in the most complex parts of our programs.

Two such patterns are codified directly into the native ES6 `Promise` implementation, so we get them for free, to use as building blocks for other patterns.

`Promise.all([..])`

In an async sequence (Promise chain), only one async task is being coordinated at any given moment – step 2 strictly follows step 1, and step 3 strictly follows step 2. But what about doing two or more steps concurrently (aka “in parallel”)?

In classic programming terminology, a “gate” is a mechanism that waits on two or more parallel/concurrent tasks to complete before continuing. It doesn't

matter what order they finish in, just that all of them have to complete for the gate to open and let the flow control through.

In the Promise API, we call this pattern `all([...])`.

Say you wanted to make two Ajax requests at the same time, and wait for both to finish, regardless of their order, before making a third Ajax request. Consider:

```
// `request(..)` is a Promise-aware Ajax utility,
// like we defined earlier in the chapter

var p1 = request( "http://some.url.1/" );
var p2 = request( "http://some.url.2/" );

Promise.all( [p1,p2] )
.then( function(msgs){
    // both `p1` and `p2` fulfill and pass in
    // their messages here
    return request(
        "http://some.url.3/?v=" + msgs.join(",")
    );
} )
.then( function(msg){
    console.log( msg );
} );

Promise.all([ ... ]) expects a single argument, an array, consisting generally of Promise instances. The promise returned from the Promise.all([ ... ]) call will receive a fulfillment message (msgs in this snippet) that is an array of all the fulfillment messages from the passed in promises, in the same order as specified (regardless of fulfillment order).
```

Note: Technically, the `array` of values passed into `Promise.all([...])` can include Promises, thenables, or even immediate values. Each value in the list is essentially passed through `Promise.resolve(..)` to make sure it's a genuine Promise to be waited on, so an immediate value will just be normalized into a Promise for that value. If the `array` is empty, the main Promise is immediately fulfilled.

The main promise returned from `Promise.all([...])` will only be fulfilled if and when all its constituent promises are fulfilled. If any one of those promises instead is rejected, the main `Promise.all([...])` promise is immediately rejected, discarding all results from any other promises.

Remember to always attach a rejection/error handler to every promise, even and especially the one that comes back from `Promise.all([...])`.

Promise.race([..])

While `Promise.all([..])` coordinates multiple Promises concurrently and assumes all are needed for fulfillment, sometimes you only want to respond to the “first Promise to cross the finish line,” letting the other Promises fall away.

This pattern is classically called a “latch,” but in Promises it’s called a “race.”

Warning: While the metaphor of “only the first across the finish line wins” fits the behavior well, unfortunately “race” is kind of a loaded term, because “race conditions” are generally taken as bugs in programs (see Chapter 1). Don’t confuse `Promise.race([..])` with “race condition.”

`Promise.race([..])` also expects a single `array` argument, containing one or more Promises, thenables, or immediate values. It doesn’t make much practical sense to have a race with immediate values, because the first one listed will obviously win – like a foot race where one runner starts at the finish line!

Similar to `Promise.all([..])`, `Promise.race([..])` will fulfill if and when any Promise resolution is a fulfillment, and it will reject if and when any Promise resolution is a rejection.

Warning: A “race” requires at least one “runner,” so if you pass an empty `array`, instead of immediately resolving, the main `race([..])` Promise will never resolve. This is a footgun! ES6 should have specified that it either fulfills, rejects, or just throws some sort of synchronous error. Unfortunately, because of precedence in Promise libraries predating ES6 `Promise`, they had to leave this gotcha in there, so be careful never to send in an empty `array`.

Let’s revisit our previous concurrent Ajax example, but in the context of a race between `p1` and `p2`:

```
// `request(..)` is a Promise-aware Ajax utility,
// like we defined earlier in the chapter

var p1 = request( "http://some.url.1/" );
var p2 = request( "http://some.url.2/" );

Promise.race( [p1,p2] )
.then( function(msg){
    // either `p1` or `p2` will win the race
    return request(
        "http://some.url.3/?v=" + msg
    );
} )
.then( function(msg){
    console.log( msg );
} );
```

Because only one promise wins, the fulfillment value is a single message, not an

array as it was for `Promise.all([...])`.

Timeout Race We saw this example earlier, illustrating how `Promise.race([...])` can be used to express the “promise timeout” pattern:

```
// `foo()` is a Promise-aware function

// `timeoutPromise(..)`, defined earlier, returns
// a Promise that rejects after a specified delay

// setup a timeout for `foo()`
Promise.race([
  foo(),                      // attempt `foo()`
  timeoutPromise( 3000 )      // give it 3 seconds
])
.then(
  function(){
    // `foo(..)` fulfilled in time!
  },
  function(err){
    // either `foo()` rejected, or it just
    // didn't finish in time, so inspect
    // `err` to know which
  }
);

```

This timeout pattern works well in most cases. But there are some nuances to consider, and frankly they apply to both `Promise.race([...])` and `Promise.all([...])` equally.

“Finally” The key question to ask is, “What happens to the promises that get discarded/ignored?” We’re not asking that question from the performance perspective – they would typically end up garbage collection eligible – but from the behavioral perspective (side effects, etc.). Promises cannot be canceled – and shouldn’t be as that would destroy the external immutability trust discussed in the “Promise Uncancelable” section later in this chapter – so they can only be silently ignored.

But what if `foo()` in the previous example is reserving some sort of resource for usage, but the timeout fires first and causes that promise to be ignored? Is there anything in this pattern that proactively frees the reserved resource after the timeout, or otherwise cancels any side effects it may have had? What if all you wanted was to log the fact that `foo()` timed out?

Some developers have proposed that Promises need a `finally(..)` callback registration, which is always called when a Promise resolves, and allows you to specify any cleanup that may be necessary. This doesn’t exist in the specification

at the moment, but it may come in ES7+. We'll have to wait and see.

It might look like:

```
var p = Promise.resolve( 42 );

p.then( something )
  .finally( cleanup )
  .then( another )
  .finally( cleanup );
```

Note: In various Promise libraries, `finally(..)` still creates and returns a new Promise (to keep the chain going). If the `cleanup(..)` function were to return a Promise, it would be linked into the chain, which means you could still have the unhandled rejection issues we discussed earlier.

In the meantime, we could make a static helper utility that lets us observe (without interfering) the resolution of a Promise:

```
// polyfill-safe guard check
if (!Promise.observe) {
  Promise.observe = function(pr,cb) {
    // side-observe `pr`'s resolution
    pr.then(
      function fulfilled(msg){
        // schedule callback async (as Job)
        Promise.resolve( msg ).then( cb );
      },
      function rejected(err){
        // schedule callback async (as Job)
        Promise.resolve( err ).then( cb );
      }
    );

    // return original promise
    return pr;
  };
}
```

Here's how we'd use it in the timeout example from before:

```
Promise.race( [
  Promise.observe(
    foo(), // attempt `foo()`
    function cleanup(msg){
      // clean up after `foo()`, even if it
      // didn't finish before the timeout
    }
  ),
  ...
]);
```

```
    timeoutPromise( 3000 ) // give it 3 seconds
] )
```

This `Promise.observe(..)` helper is just an illustration of how you could observe the completions of Promises without interfering with them. Other Promise libraries have their own solutions. Regardless of how you do it, you'll likely have places where you want to make sure your Promises aren't *just* silently ignored by accident.

Variations on `all([..])` and `race([..])`

While native ES6 Promises come with built-in `Promise.all([..])` and `Promise.race([..])`, there are several other commonly used patterns with variations on those semantics:

- `none([..])` is like `all([..])`, but fulfillments and rejections are transposed. All Promises need to be rejected – rejections become the fulfillment values and vice versa.
- `any([..])` is like `all([..])`, but it ignores any rejections, so only one needs to fulfill instead of *all* of them.
- `first([..])` is like a race with `any([..])`, which is that it ignores any rejections and fulfills as soon as the first Promise fulfills.
- `last([..])` is like `first([..])`, but only the latest fulfillment wins.

Some Promise abstraction libraries provide these, but you could also define them yourself using the mechanics of Promises, `race([..])` and `all([..])`.

For example, here's how we could define `first([..])`:

```
// polyfill-safe guard check
if (!Promise.first) {
  Promise.first = function(prs) {
    return new Promise( function(resolve,reject){
      // loop through all promises
      prs.forEach( function(pr){
        // normalize the value
        Promise.resolve( pr )
        // whichever one fulfills first wins, and
        // gets to resolve the main promise
        .then( resolve );
      } );
    } );
  };
}
```

Note: This implementation of `first(..)` does not reject if all its promises reject; it simply hangs, much like a `Promise.race([])` does. If desired, you could add additional logic to track each promise rejection and if all reject, call `reject()` on the main promise. We'll leave that as an exercise for the reader.

Concurrent Iterations

Sometimes you want to iterate over a list of Promises and perform some task against all of them, much like you can do with synchronous `arrays` (e.g., `forEach(..)`, `map(..)`, `some(..)`, and `every(..)`). If the task to perform against each Promise is fundamentally synchronous, these work fine, just as we used `forEach(..)` in the previous snippet.

But if the tasks are fundamentally asynchronous, or can/should otherwise be performed concurrently, you can use `async` versions of these utilities as provided by many libraries.

For example, let's consider an asynchronous `map(..)` utility that takes an `array` of values (could be Promises or anything else), plus a function (task) to perform against each. `map(..)` itself returns a promise whose fulfillment value is an `array` that holds (in the same mapping order) the `async` fulfillment value from each task:

```
if (!Promise.map) {
  Promise.map = function(vals,cb) {
    // new promise that waits for all mapped promises
    return Promise.all(
      // note: regular array `map(..)`, turns
      // the array of values into an array of
      // promises
      vals.map( function(val){
        // replace `val` with a new promise that
        // resolves after `val` is async mapped
        return new Promise( function(resolve){
          cb( val, resolve );
        } );
      } )
    );
  };
}
```

Note: In this implementation of `map(..)`, you can't signal `async` rejection, but if a synchronous exception/error occurs inside of the mapping callback (`cb(..)`), the main `Promise.map(..)` returned promise would reject.

Let's illustrate using `map(..)` with a list of Promises (instead of simple values):

```
var p1 = Promise.resolve( 21 );
var p2 = Promise.resolve( 42 );
var p3 = Promise.reject( "Oops" );

// double values in list even if they're in
// Promises
Promise.map( [p1,p2,p3], function(pr,done){
```

```

// make sure the item itself is a Promise
Promise.resolve( pr )
  .then(
    // extract value as `v`
    function(v){
      // map fulfillment `v` to new value
      done( v * 2 );
    },
    // or, map to promise rejection message
    done
  );
}
.then( function(vals){
  console.log( vals );    // [42,84, "Oops"]
} );

```

Promise API Recap

Let's review the ES6 `Promise` API that we've already seen unfold in bits and pieces throughout this chapter.

Note: The following API is native only as of ES6, but there are specification-compliant polyfills (not just extended Promise libraries) which can define `Promise` and all its associated behavior so that you can use native Promises even in pre-ES6 browsers. One such polyfill is "Native Promise Only" (<http://github.com/getify/native-promise-only>), which I wrote!

`new Promise(..)` Constructor

The *revealing constructor* `Promise(..)` must be used with `new`, and must be provided a function callback that is synchronously/immediately called. This function is passed two function callbacks that act as resolution capabilities for the promise. We commonly label these `resolve(..)` and `reject(..)`:

```

var p = new Promise( function(resolve,reject){
  // `resolve(..)` to resolve/fulfill the promise
  // `reject(..)` to reject the promise
} );

```

`reject(..)` simply rejects the promise, but `resolve(..)` can either fulfill the promise or reject it, depending on what it's passed. If `resolve(..)` is passed an immediate, non-Promise, non-thenable value, then the promise is fulfilled with that value.

But if `resolve(..)` is passed a genuine Promise or thenable value, that value is unwrapped recursively, and whatever its final resolution/state is will be adopted by the promise.

Promise.resolve(..) and Promise.reject(..)

A shortcut for creating an already-rejected Promise is `Promise.reject(..)`, so these two promises are equivalent:

```
var p1 = new Promise( function(resolve,reject){  
    reject( "Oops" );  
} );  
  
var p2 = Promise.reject( "Oops" );
```

`Promise.resolve(..)` is usually used to create an already-fulfilled Promise in a similar way to `Promise.reject(..)`. However, `Promise.resolve(..)` also unwraps thenable values (as discussed several times already). In that case, the Promise returned adopts the final resolution of the thenable you passed in, which could either be fulfillment or rejection:

```
var fulfilledTh = {  
    then: function(cb) { cb( 42 ); }  
};  
var rejectedTh = {  
    then: function(cb,errCb) {  
        errCb( "Oops" );  
    }  
};  
  
var p1 = Promise.resolve( fulfilledTh );  
var p2 = Promise.resolve( rejectedTh );  
  
// `p1` will be a fulfilled promise  
// `p2` will be a rejected promise
```

And remember, `Promise.resolve(..)` doesn't do anything if what you pass is already a genuine Promise; it just returns the value directly. So there's no overhead to calling `Promise.resolve(..)` on values that you don't know the nature of, if one happens to already be a genuine Promise.

then(..) and catch(..)

Each Promise instance (**not** the `Promise` API namespace) has `then(..)` and `catch(..)` methods, which allow registering of fulfillment and rejection handlers for the Promise. Once the Promise is resolved, one or the other of these handlers will be called, but not both, and it will always be called asynchronously (see "Jobs" in Chapter 1).

`then(..)` takes one or two parameters, the first for the fulfillment callback, and the second for the rejection callback. If either is omitted or is otherwise passed as a non-function value, a default callback is substituted respectively. The default

fulfillment callback simply passes the message along, while the default rejection callback simply rethrows (propagates) the error reason it receives.

`catch(..)` takes only the rejection callback as a parameter, and automatically substitutes the default fulfillment callback, as just discussed. In other words, it's equivalent to `then(null,..)`:

```
p.then( fulfilled );
p.then( fulfilled, rejected );
p.catch( rejected ); // or `p.then( null, rejected )`
```

`then(..)` and `catch(..)` also create and return a new promise, which can be used to express Promise chain flow control. If the fulfillment or rejection callbacks have an exception thrown, the returned promise is rejected. If either callback returns an immediate, non-Promise, non-thenable value, that value is set as the fulfillment for the returned promise. If the fulfillment handler specifically returns a promise or thenable value, that value is unwrapped and becomes the resolution of the returned promise.

Promise.all([..]) and Promise.race([..])

The static helpers `Promise.all([..])` and `Promise.race([..])` on the ES6 `Promise` API both create a `Promise` as their return value. The resolution of that promise is controlled entirely by the array of promises that you pass in.

For `Promise.all([..])`, all the promises you pass in must fulfill for the returned promise to fulfill. If any promise is rejected, the main returned promise is immediately rejected, too (discarding the results of any of the other promises). For fulfillment, you receive an `array` of all the passed in promises' fulfillment values. For rejection, you receive just the first promise rejection reason value. This pattern is classically called a “gate”: all must arrive before the gate opens.

For `Promise.race([..])`, only the first promise to resolve (fulfillment or rejection) “wins,” and whatever that resolution is becomes the resolution of the returned promise. This pattern is classically called a “latch”: first one to open the latch gets through. Consider:

```
var p1 = Promise.resolve( 42 );
var p2 = Promise.resolve( "Hello World" );
var p3 = Promise.reject( "Oops" );

Promise.race( [p1,p2,p3] )
.then( function(msg){
    console.log( msg );      // 42
} );

Promise.all( [p1,p2,p3] )
```

```

    .catch( function(err){
      console.error( err );    // "Oops"
    } );
  }

Promise.all( [p1,p2] )
.then( function(msgs){
  console.log( msgs );    // [42, "Hello World"]
} );

```

Warning: Be careful! If an empty array is passed to `Promise.all([..])`, it will fulfill immediately, but `Promise.race([..])` will hang forever and never resolve.

The ES6 `Promise` API is pretty simple and straightforward. It's at least good enough to serve the most basic of async cases, and is a good place to start when rearranging your code from callback hell to something better.

But there's a whole lot of async sophistication that apps often demand which Promises themselves will be limited in addressing. In the next section, we'll dive into those limitations as motivations for the benefit of Promise libraries.

Promise Limitations

Many of the details we'll discuss in this section have already been alluded to in this chapter, but we'll just make sure to review these limitations specifically.

Sequence Error Handling

We covered Promise-flavored error handling in detail earlier in this chapter. The limitations of how Promises are designed – how they chain, specifically – creates a very easy pitfall where an error in a Promise chain can be silently ignored accidentally.

But there's something else to consider with Promise errors. Because a Promise chain is nothing more than its constituent Promises wired together, there's no entity to refer to the entire chain as a single *thing*, which means there's no external way to observe any errors that may occur.

If you construct a Promise chain that has no error handling in it, any error anywhere in the chain will propagate indefinitely down the chain, until observed (by registering a rejection handler at some step). So, in that specific case, having a reference to the *last* promise in the chain is enough (p in the following snippet), because you can register a rejection handler there, and it will be notified of any propagated errors:

```

// `foo(..)`, `STEP2(..)` and `STEP3(..)` are
// all promise-aware utilities

var p = foo( 42 )

```

```
.then( STEP2 )
.then( STEP3 );
```

Although it may seem sneakily confusing, `p` here doesn't point to the first promise in the chain (the one from the `foo(42)` call), but instead from the last promise, the one that comes from the `then(STEP3)` call.

Also, no step in the promise chain is observably doing its own error handling. That means that you could then register a rejection error handler on `p`, and it would be notified if any errors occur anywhere in the chain:

```
p.catch( handleErrors );
```

But if any step of the chain in fact does its own error handling (perhaps hidden/abstracted away from what you can see), your `handleErrors(..)` won't be notified. This may be what you want – it was, after all, a "handled rejection" – but it also may *not* be what you want. The complete lack of ability to be notified (of "already handled" rejection errors) is a limitation that restricts capabilities in some use cases.

It's basically the same limitation that exists with a `try..catch` that can catch an exception and simply swallow it. So this isn't a limitation **unique to Promises**, but it *is* something we might wish to have a workaround for.

Unfortunately, many times there is no reference kept for the intermediate steps in a Promise-chain sequence, so without such references, you cannot attach error handlers to reliably observe the errors.

Single Value

Promises by definition only have a single fulfillment value or a single rejection reason. In simple examples, this isn't that big of a deal, but in more sophisticated scenarios, you may find this limiting.

The typical advice is to construct a values wrapper (such as an `object` or `array`) to contain these multiple messages. This solution works, but it can be quite awkward and tedious to wrap and unwrap your messages with every single step of your Promise chain.

Splitting Values Sometimes you can take this as a signal that you could/should decompose the problem into two or more Promises.

Imagine you have a utility `foo(..)` that produces two values (`x` and `y`) asynchronously:

```
function getY(x) {
  return new Promise( function(resolve,reject){
    setTimeout( function(){
      resolve( (3 * x) - 1 );
    }, 100 );
  });
}
```

```

        } );
    }

    function foo(bar,baz) {
        var x = bar * baz;

        return getY( x )
            .then( function(y){
                // wrap both values into container
                return [x,y];
            } );
    }

    foo( 10, 20 )
        .then( function(msgs){
            var x = msgs[0];
            var y = msgs[1];

            console.log( x, y );    // 200 599
        } );
}

```

First, let's rearrange what `foo(..)` returns so that we don't have to wrap `x` and `y` into a single `array` value to transport through one `Promise`. Instead, we can wrap each value into its own `promise`:

```

function foo(bar,baz) {
    var x = bar * baz;

    // return both promises
    return [
        Promise.resolve( x ),
        getY( x )
    ];
}

Promise.all(
    foo( 10, 20 )
)
    .then( function(msgs){
        var x = msgs[0];
        var y = msgs[1];

        console.log( x, y );
    } );
}

```

Is an `array` of `promises` really better than an `array` of `values` passed through a single `promise`? Syntactically, it's not much of an improvement.

But this approach more closely embraces the Promise design theory. It's now easier in the future to refactor to split the calculation of `x` and `y` into separate functions. It's cleaner and more flexible to let the calling code decide how to orchestrate the two promises – using `Promise.all([...])` here, but certainly not the only option – rather than to abstract such details away inside of `foo(..)`.

Unwrap/Spread Arguments The `var x = ..` and `var y = ..` assignments are still awkward overhead. We can employ some functional trickery (hat tip to Reginald Braithwaite, @raganwald on Twitter) in a helper utility:

```
function spread(fn) {
  return Function.prototype.bind(fn, null);
}

Promise.all(
  foo( 10, 20 )
)
.then(
  spread(function(x,y){
    console.log( x, y );    // 200 599
  })
)

```

That's a bit nicer! Of course, you could inline the functional magic to avoid the extra helper:

```
Promise.all(
  foo( 10, 20 )
)
.then( Function.prototype.bind(
  function(x,y){
    console.log( x, y );    // 200 599
  },
  null
) );

```

These tricks may be neat, but ES6 has an even better answer for us: destructuring. The array destructuring assignment form looks like this:

```
Promise.all(
  foo( 10, 20 )
)
.then( function(msgs){
  var [x,y] = msgs;

  console.log( x, y );    // 200 599
} );

```

But best of all, ES6 offers the array parameter destructuring form:

```
Promise.all(
  foo( 10, 20 )
)
.then( function([x,y]){
  console.log( x, y );    // 200 599
} );
```

We've now embraced the one-value-per-Promise mantra, but kept our supporting boilerplate to a minimum!

Note: For more information on ES6 destructuring forms, see the *ES6 & Beyond* title of this series.

Single Resolution

One of the most intrinsic behaviors of Promises is that a Promise can only be resolved once (fulfillment or rejection). For many async use cases, you're only retrieving a value once, so this works fine.

But there's also a lot of async cases that fit into a different model – one that's more akin to events and/or streams of data. It's not clear on the surface how well Promises can fit into such use cases, if at all. Without a significant abstraction on top of Promises, they will completely fall short for handling multiple value resolution.

Imagine a scenario where you might want to fire off a sequence of async steps in response to a stimulus (like an event) that can in fact happen multiple times, like a button click.

This probably won't work the way you want:

```
// `click(..)` binds the `click` event to a DOM element
// `request(..)` is the previously defined Promise-aware Ajax

var p = new Promise( function(resolve,reject){
  click( "#mybtn", resolve );
} );

p.then( function(evt){
  var btnID = evt.currentTarget.id;
  return request( "http://some.url.1/?id=" + btnID );
} )
.then( function(text){
  console.log( text );
} );
```

The behavior here only works if your application calls for the button to be clicked just once. If the button is clicked a second time, the p promise has

already been resolved, so the second `resolve(..)` call would be ignored.

Instead, you'd probably need to invert the paradigm, creating a whole new Promise chain for each event firing:

```
click( "#mybtn", function(evt){
  var btnID = evt.currentTarget.id;

  request( "http://some.url.1/?id=" + btnID )
    .then( function(text){
      console.log( text );
    } );
} );
```

This approach will *work* in that a whole new Promise sequence will be fired off for each "click" event on the button.

But beyond just the ugliness of having to define the entire Promise chain inside the event handler, this design in some respects violates the idea of separation of concerns/capabilities (SoC). You might very well want to define your event handler in a different place in your code from where you define the *response* to the event (the Promise chain). That's pretty awkward to do in this pattern, without helper mechanisms.

Note: Another way of articulating this limitation is that it'd be nice if we could construct some sort of "observable" that we can subscribe a Promise chain to. There are libraries that have created these abstractions (such as RxJS – <http://rxjs.codeplex.com/>), but the abstractions can seem so heavy that you can't even see the nature of Promises anymore. Such heavy abstraction brings important questions to mind such as whether (sans Promises) these mechanisms are as *trustable* as Promises themselves have been designed to be. We'll revisit the "Observable" pattern in Appendix B.

Inertia

One concrete barrier to starting to use Promises in your own code is all the code that currently exists which is not already Promise-aware. If you have lots of callback-based code, it's far easier to just keep coding in that same style.

"A code base in motion (with callbacks) will remain in motion (with callbacks) unless acted upon by a smart, Promises-aware developer."

Promises offer a different paradigm, and as such, the approach to the code can be anywhere from just a little different to, in some cases, radically different. You have to be intentional about it, because Promises will not just naturally shake out from the same ol' ways of doing code that have served you well thus far.

Consider a callback-based scenario like the following:

```
function foo(x,y,cb) {
```

```

    ajax(
      "http://some.url.1/?x=" + x + "&y=" + y,
      cb
    );
}

foo( 11, 31, function(err, text) {
  if (err) {
    console.error( err );
  }
  else {
    console.log( text );
  }
} );

```

Is it immediately obvious what the first steps are to convert this callback-based code to Promise-aware code? Depends on your experience. The more practice you have with it, the more natural it will feel. But certainly, Promises don't just advertise on the label exactly how to do it – there's no one-size-fits-all answer – so the responsibility is up to you.

As we've covered before, we definitely need an Ajax utility that is Promise-aware instead of callback-based, which we could call `request(..)`. You can make your own, as we have already. But the overhead of having to manually define Promise-aware wrappers for every callback-based utility makes it less likely you'll choose to refactor to Promise-aware coding at all.

Promises offer no direct answer to that limitation. Most Promise libraries do offer a helper, however. But even without a library, imagine a helper like this:

```

// polyfill-safe guard check
if (!Promise.wrap) {
  Promise.wrap = function(fn) {
    return function() {
      var args = [].slice.call( arguments );

      return new Promise( function(resolve, reject){
        fn.apply(
          null,
          args.concat( function(err, v){
            if (err) {
              reject( err );
            }
            else {
              resolve( v );
            }
          } )
      );
    };
  };
}

```

```

        } );
    } ;
}
}

```

OK, that's more than just a tiny trivial utility. However, although it may look a bit intimidating, it's not as bad as you'd think. It takes a function that expects an error-first style callback as its last parameter, and returns a new one that automatically creates a Promise to return, and substitutes the callback for you, wired up to the Promise fulfillment/rejection.

Rather than waste too much time talking about *how* this `Promise.wrap(..)` helper works, let's just look at how we use it:

```

var request = Promise.wrap( ajax );

request( "http://some.url.1/" )
.then( ... )
...

```

Wow, that was pretty easy!

`Promise.wrap(..)` does **not** produce a Promise. It produces a function that will produce Promises. In a sense, a Promise-producing function could be seen as a "Promise factory." I propose "promisory" as the name for such a thing ("Promise" + "factory").

The act of wrapping a callback-expecting function to be a Promise-aware function is sometimes referred to as "lifting" or "promisifying". But there doesn't seem to be a standard term for what to call the resultant function other than a "lifted function", so I like "promisory" better as I think it's more descriptive.

Note: Promisory isn't a made-up term. It's a real word, and its definition means to contain or convey a promise. That's exactly what these functions are doing, so it turns out to be a pretty perfect terminology match!

So, `Promise.wrap.ajax` produces an `ajax(..)` promisory we call `request(..)`, and that promisory produces Promises for Ajax responses.

If all functions were already promisories, we wouldn't need to make them ourselves, so the extra step is a tad bit of a shame. But at least the wrapping pattern is (usually) repeatable so we can put it into a `Promise.wrap(..)` helper as shown to aid our promise coding.

So back to our earlier example, we need a promisory for both `ajax(..)` and `foo(..)`:

```

// make a promisory for `ajax(..)`
var request = Promise.wrap( ajax );

// refactor `foo(..)`, but keep it externally

```

```

// callback-based for compatibility with other
// parts of the code for now -- only use
// `request(..)`'s promise internally.
function foo(x,y,cb) {
  request(
    "http://some.url.1/?x=" + x + "&y=" + y
  )
  .then(
    function fulfilled(text){
      cb( null, text );
    },
    cb
  );
}

// now, for this code's purposes, make a
// promisory for `foo(..)`
var betterFoo = Promise.wrap( foo );

// and use the promisory
betterFoo( 11, 31 )
.then(
  function fulfilled(text){
    console.log( text );
  },
  function rejected(err){
    console.error( err );
  }
);

```

Of course, while we're refactoring `foo(..)` to use our new `request(..)` promisory, we could just make `foo(..)` a promisory itself, instead of remaining callback-based and needing to make and use the subsequent `betterFoo(..)` promisory. This decision just depends on whether `foo(..)` needs to stay callback-based compatible with other parts of the code base or not.

Consider:

```

// `foo(..)` is now also a promisory because it
// delegates to the `request(..)` promisory
function foo(x,y) {
  return request(
    "http://some.url.1/?x=" + x + "&y=" + y
  );
}

foo( 11, 31 )

```

```
.then( ... )
```

```
..
```

While ES6 Promises don't natively ship with helpers for such promisory wrapping, most libraries provide them, or you can make your own. Either way, this particular limitation of Promises is addressable without too much pain (certainly compared to the pain of callback hell!).

Promise Uncancelable

Once you create a Promise and register a fulfillment and/or rejection handler for it, there's nothing external you can do to stop that progression if something else happens to make that task moot.

Note: Many Promise abstraction libraries provide facilities to cancel Promises, but this is a terrible idea! Many developers wish Promises had natively been designed with external cancelation capability, but the problem is that it would let one consumer/observer of a Promise affect some other consumer's ability to observe that same Promise. This violates the future-value's trustability (external immutability), but moreover is the embodiment of the "action at a distance" anti-pattern (http://en.wikipedia.org/wiki/Action_at_a_distance_%28computer_programming%29). Regardless of how useful it seems, it will actually lead you straight back into the same nightmares as callbacks.

Consider our Promise timeout scenario from earlier:

```
var p = foo( 42 );  
  
Promise.race( [  
  p,  
  timeoutPromise( 3000 )  
] )  
.then(  
  doSomething,  
  handleError  
);  
  
p.then( function(){  
  // still happens even in the timeout case :(  
} );
```

The "timeout" was external to the promise `p`, so `p` itself keeps going, which we probably don't want.

One option is to invasively define your resolution callbacks:

```
var OK = true;
```

```

var p = foo( 42 );

Promise.race( [
  p,
  timeoutPromise( 3000 )
  .catch( function(err){
    OK = false;
    throw err;
  } )
] )
.then(
  doSomething,
  handleError
);

p.then( function(){
  if (OK) {
    // only happens if no timeout! :)
  }
} );

```

This is ugly. It works, but it's far from ideal. Generally, you should try to avoid such scenarios.

But if you can't, the ugliness of this solution should be a clue that *cancelation* is a functionality that belongs at a higher level of abstraction on top of Promises. I'd recommend you look to Promise abstraction libraries for assistance rather than hacking it yourself.

Note: My *asynquence* Promise abstraction library provides just such an abstraction and an `abort()` capability for the sequence, all of which will be discussed in Appendix A.

A single Promise is not really a flow-control mechanism (at least not in a very meaningful sense), which is exactly what *cancelation* refers to; that's why Promise cancelation would feel awkward.

By contrast, a chain of Promises taken collectively together – what I like to call a “sequence” – *is* a flow control expression, and thus it's appropriate for cancelation to be defined at that level of abstraction.

No individual Promise should be cancelable, but it's sensible for a *sequence* to be cancelable, because you don't pass around a sequence as a single immutable value like you do with a Promise.

Promise Performance

This particular limitation is both simple and complex.

Comparing how many pieces are moving with a basic callback-based async task chain versus a Promise chain, it's clear Promises have a fair bit more going on, which means they are naturally at least a tiny bit slower. Think back to just the simple list of trust guarantees that Promises offer, as compared to the ad hoc solution code you'd have to layer on top of callbacks to achieve the same protections.

More work to do, more guards to protect, means that Promises *are* slower as compared to naked, untrustable callbacks. That much is obvious, and probably simple to wrap your brain around.

But how much slower? Well... that's actually proving to be an incredibly difficult question to answer absolutely, across the board.

Frankly, it's kind of an apples-to-oranges comparison, so it's probably the wrong question to ask. You should actually compare whether an ad-hoc callback system with all the same protections manually layered in is faster than a Promise implementation.

If Promises have a legitimate performance limitation, it's more that they don't really offer a line-item choice as to which trustability protections you want/need or not – you get them all, always.

Nevertheless, if we grant that a Promise is generally a *little bit slower* than its non-Promise, non-trustable callback equivalent – assuming there are places where you feel you can justify the lack of trustability – does that mean that Promises should be avoided across the board, as if your entire application is driven by nothing but must-be-utterly-the-fastest code possible?

Sanity check: if your code is legitimately like that, **is JavaScript even the right language for such tasks?** JavaScript can be optimized to run applications very performantly (see Chapter 5 and Chapter 6). But is obsessing over tiny performance tradeoffs with Promises, in light of all the benefits they offer, *really* appropriate?

Another subtle issue is that Promises make *everything* async, which means that some immediately (synchronously) complete steps still defer advancement of the next step to a Job (see Chapter 1). That means that it's possible that a sequence of Promise tasks could complete ever-so-slightly slower than the same sequence wired up with callbacks.

Of course, the question here is this: are these potential slips in tiny fractions of performance *worth* all the other articulated benefits of Promises we've laid out across this chapter?

My take is that in virtually all cases where you might think Promise performance is slow enough to be concerned, it's actually an anti-pattern to optimize away the benefits of Promise trustability and composability by avoiding them altogether.

Instead, you should default to using them across the code base, and then profile and analyze your application's hot (critical) paths. Are Promises *really* a bot-

tleneck, or are they just a theoretical slowdown? Only *then*, armed with actual valid benchmarks (see Chapter 6) is it responsible and prudent to factor out the Promises in just those identified critical areas.

Promises are a little slower, but in exchange you're getting a lot of trustability, non-Zalgo predictability, and composability built in. Maybe the limitation is not actually their performance, but your lack of perception of their benefits?

Review

Promises are awesome. Use them. They solve the *inversion of control* issues that plague us with callbacks-only code.

They don't get rid of callbacks, they just redirect the orchestration of those callbacks to a trustable intermediary mechanism that sits between us and another utility.

Promise chains also begin to address (though certainly not perfectly) a better way of expressing async flow in sequential fashion, which helps our brains plan and maintain async JS code better. We'll see an even better solution to *that* problem in the next chapter!

You Don't Know JS: Async & Performance

Chapter 4: Generators

In Chapter 2, we identified two key drawbacks to expressing async flow control with callbacks:

- Callback-based async doesn't fit how our brain plans out steps of a task.
- Callbacks aren't trustable or composable because of *inversion of control*.

In Chapter 3, we detailed how Promises uninvert the *inversion of control* of callbacks, restoring trustability/composability.

Now we turn our attention to expressing async flow control in a sequential, synchronous-looking fashion. The “magic” that makes it possible is ES6 **generators**.

Breaking Run-to-Completion

In Chapter 1, we explained an expectation that JS developers almost universally rely on in their code: once a function starts executing, it runs until it completes, and no other code can interrupt and run in between.

As bizarre as it may seem, ES6 introduces a new type of function that does not behave with the run-to-completion behavior. This new type of function is called a “generator.”

To understand the implications, let's consider this example:

```
var x = 1;

function foo() {
  x++;
  bar();           // <-- what about this line?
  console.log( "x:", x );
}

function bar() {
  x++;
}

foo();           // x: 3
```

In this example, we know for sure that `bar()` runs in between `x++` and `console.log(x)`. But what if `bar()` wasn't there? Obviously, the result would be 2 instead of 3.

Now let's twist your brain. What if `bar()` wasn't present, but it could still somehow run between the `x++` and `console.log(x)` statements? How would that be possible?

In **preemptive** multithreaded languages, it would essentially be possible for `bar()` to “interrupt” and run at exactly the right moment between those two statements. But JS is not preemptive, nor is it (currently) multithreaded. And yet, a **cooperative** form of this “interruption” (concurrency) is possible, if `foo()` itself could somehow indicate a “pause” at that part in the code.

Note: I use the word “cooperative” not only because of the connection to classical concurrency terminology (see Chapter 1), but because as you’ll see in the next snippet, the ES6 syntax for indicating a pause point in code is `yield` – suggesting a politely *cooperative* yielding of control.

Here’s the ES6 code to accomplish such cooperative concurrency:

```
var x = 1;

function *foo() {
  x++;
  yield; // pause!
  console.log("x:", x);
}

function bar() {
  x++;
}
```

Note: You will likely see most other JS documentation/code that will format a generator declaration as `function* foo() { .. }` instead of as I’ve done here with `function *foo() { .. }` – the only difference being the stylistic positioning of the `*`. The two forms are functionally/syntactically identical, as is a third `function*foo() { .. }` (no space) form. There are arguments for both styles, but I basically prefer `function *foo...` because it then matches when I reference a generator in writing with `*foo()`. If I said only `foo()`, you wouldn’t know as clearly if I was talking about a generator or a regular function. It’s purely a stylistic preference.

Now, how can we run the code in that previous snippet such that `bar()` executes at the point of the `yield` inside of `*foo()`?

```
// construct an iterator `it` to control the generator
var it = foo();

// start `foo()` here!
it.next(); // 2
x;          // 2
bar();      // 3
x;          // x: 3
it.next(); // x: 3
```

OK, there’s quite a bit of new and potentially confusing stuff in those two

code snippets, so we've got plenty to wade through. But before we explain the different mechanics/syntax with ES6 generators, let's walk through the behavior flow:

1. The `it = foo()` operation does *not* execute the `*foo()` generator yet, but it merely constructs an *iterator* that will control its execution. More on *iterators* in a bit.
2. The first `it.next()` starts the `*foo()` generator, and runs the `x++` on the first line of `*foo()`.
3. `*foo()` pauses at the `yield` statement, at which point that first `it.next()` call finishes. At the moment, `*foo()` is still running and active, but it's in a paused state.
4. We inspect the value of `x`, and it's now 2.
5. We call `bar()`, which increments `x` again with `x++`.
6. We inspect the value of `x` again, and it's now 3.
7. The final `it.next()` call resumes the `*foo()` generator from where it was paused, and runs the `console.log(..)` statement, which uses the current value of `x` of 3.

Clearly, `*foo()` started, but did *not* run-to-completion – it paused at the `yield`. We resumed `*foo()` later, and let it finish, but that wasn't even required.

So, a generator is a special kind of function that can start and stop one or more times, and doesn't necessarily ever have to finish. While it won't be terribly obvious yet why that's so powerful, as we go throughout the rest of this chapter, that will be one of the fundamental building blocks we use to construct generators-as-async-flow-control as a pattern for our code.

Input and Output

A generator function is a special function with the new processing model we just alluded to. But it's still a function, which means it still has some basic tenets that haven't changed – namely, that it still accepts arguments (aka "input"), and that it can still return a value (aka "output"):

```
function *foo(x,y) {  
  return x * y;  
}  
  
var it = foo( 6, 7 );  
  
var res = it.next();  
  
res.value;      // 42
```

We pass in the arguments 6 and 7 to `*foo(..)` as the parameters `x` and `y`, respectively. And `*foo(..)` returns the value 42 back to the calling code.

We now see a difference with how the generator is invoked compared to a nor-

mal function. `foo(6,7)` obviously looks familiar. But subtly, the `*foo(..)` generator hasn't actually run yet as it would have with a function.

Instead, we're just creating an *iterator* object, which we assign to the variable `it`, to control the `*foo(..)` generator. Then we call `it.next()`, which instructs the `*foo(..)` generator to advance from its current location, stopping either at the next `yield` or end of the generator.

The result of that `next(..)` call is an object with a `value` property on it holding whatever value (if anything) was returned from `*foo(..)`. In other words, `yield` caused a value to be sent out from the generator during the middle of its execution, kind of like an intermediate `return`.

Again, it won't be obvious yet why we need this whole indirect *iterator* object to control the generator. We'll get there, I *promise*.

Iteration Messaging In addition to generators accepting arguments and having return values, there's even more powerful and compelling input/output messaging capability built into them, via `yield` and `next(..)`.

Consider:

```
function *foo(x) {
  var y = x * (yield);
  return y;
}

var it = foo( 6 );

// start `foo(..)`
it.next();

var res = it.next( 7 );

res.value;      // 42
```

First, we pass in `6` as the parameter `x`. Then we call `it.next()`, and it starts up `*foo(..)`.

Inside `*foo(..)`, the `var y = x ..` statement starts to be processed, but then it runs across a `yield` expression. At that point, it pauses `*foo(..)` (in the middle of the assignment statement!), and essentially requests the calling code to provide a result value for the `yield` expression. Next, we call `it.next(7)`, which is passing the `7` value back in to be that result of the paused `yield` expression.

So, at this point, the assignment statement is essentially `var y = 6 * 7`. Now, `return y` returns that `42` value back as the result of the `it.next(7)` call.

Notice something very important but also easily confusing, even to seasoned JS developers: depending on your perspective, there's a mismatch between the `yield` and the `next(..)` call. In general, you're going to have one more `next(..)` call than you have `yield` statements – the preceding snippet has one `yield` and two `next(..)` calls.

Why the mismatch?

Because the first `next(..)` always starts a generator, and runs to the first `yield`. But it's the second `next(..)` call that fulfills the first paused `yield` expression, and the third `next(..)` would fulfill the second `yield`, and so on.

Tale of Two Questions Actually, which code you're thinking about primarily will affect whether there's a perceived mismatch or not.

Consider only the generator code:

```
var y = x * (yield);
return y;
```

This `first yield` is basically *asking a question*: “What value should I insert here?”

Who's going to answer that question? Well, the `first next()` has already run to get the generator up to this point, so obviously *it* can't answer the question. So, the `second next(..)` call must answer the question *posed* by the `first yield`.

See the mismatch – second-to-first?

But let's flip our perspective. Let's look at it not from the generator's point of view, but from the iterator's point of view.

To properly illustrate this perspective, we also need to explain that messages can go in both directions – `yield ..` as an expression can send out messages in response to `next(..)` calls, and `next(..)` can send values to a paused `yield` expression. Consider this slightly adjusted code:

```
function *foo(x) {
  var y = x * (yield "Hello");    // <-- yield a value!
  return y;
}

var it = foo( 6 );

var res = it.next();    // first `next()`, don't pass anything
res.value;              // "Hello"

res = it.next( 7 );    // pass `7` to waiting `yield`
res.value;              // 42
```

`yield ..` and `next(..)` pair together as a two-way message passing system **during the execution of the generator**.

So, looking only at the *iterator* code:

```
var res = it.next();      // first `next()`, don't pass anything
res.value;                // "Hello"

res = it.next( 7 );      // pass `7` to waiting `yield`
res.value;                // 42
```

Note: We don't pass a value to the first `next()` call, and that's on purpose. Only a paused `yield` could accept such a value passed by a `next(..)`, and at the beginning of the generator when we call the first `next()`, there **is no paused `yield`** to accept such a value. The specification and all compliant browsers just silently **discard** anything passed to the first `next()`. It's still a bad idea to pass a value, as you're just creating silently "failing" code that's confusing. So, always start a generator with an argument-free `next()`.

The first `next()` call (with nothing passed to it) is basically *asking a question*: "What `next` value does the `*foo(..)` generator have to give me?" And who answers this question? The first `yield "hello"` expression.

See? No mismatch there.

Depending on *who* you think about asking the question, there is either a mismatch between the `yield` and `next(..)` calls, or not.

But wait! There's still an extra `next()` compared to the number of `yield` statements. So, that final `it.next(7)` call is again asking the question about what `next` value the generator will produce. But there's no more `yield` statements left to answer, is there? So who answers?

The `return` statement answers the question!

And if there **is no `return`** in your generator – `return` is certainly not any more required in generators than in regular functions – there's always an assumed/implicit `return`; (aka `return undefined;`), which serves the purpose of default answering the question *posed* by the final `it.next(7)` call.

These questions and answers – the two-way message passing with `yield` and `next(..)` – are quite powerful, but it's not obvious at all how these mechanisms are connected to async flow control. We're getting there!

Multiple Iterators

It may appear from the syntactic usage that when you use an *iterator* to control a generator, you're controlling the declared generator function itself. But there's a subtlety that's easy to miss: each time you construct an *iterator*, you are implicitly constructing an instance of the generator which that *iterator* will control.

You can have multiple instances of the same generator running at the same time, and they can even interact:

```
function *foo() {
  var x = yield 2;
  z++;
  var y = yield (x * z);
  console.log( x, y, z );
}

var z = 1;

var it1 = foo();
var it2 = foo();

var val1 = it1.next().value;           // 2 <-- yield 2
var val2 = it2.next().value;           // 2 <-- yield 2

val1 = it1.next( val2 * 10 ).value;    // 40 <-- x:20, z:2
val2 = it2.next( val1 * 5 ).value;     // 600 <-- x:200, z:3

it1.next( val2 / 2 );                // y:300
                                      // 20 300 3
it2.next( val1 / 4 );                // y:10
                                      // 200 10 3
```

Warning: The most common usage of multiple instances of the same generator running concurrently is not such interactions, but when the generator is producing its own values without input, perhaps from some independently connected resource. We'll talk more about value production in the next section.

Let's briefly walk through the processing:

1. Both instances of `*foo()` are started at the same time, and both `next()` calls reveal a `value` of 2 from the `yield 2` statements, respectively.
2. `val2 * 10` is $2 * 10$, which is sent into the first generator instance `it1`, so that `x` gets value 20. `z` is incremented from 1 to 2, and then $20 * 2$ is `yielded` out, setting `val1` to 40.
3. `val1 * 5` is $40 * 5$, which is sent into the second generator instance `it2`, so that `x` gets value 200. `z` is incremented again, from 2 to 3, and then $200 * 3$ is `yielded` out, setting `val2` to 600.
4. `val2 / 2` is $600 / 2$, which is sent into the first generator instance `it1`, so that `y` gets value 300, then printing out 20 300 3 for its `x y z` values, respectively.
5. `val1 / 4` is $40 / 4$, which is sent into the second generator instance `it2`, so that `y` gets value 10, then printing out 200 10 3 for its `x y z` values, respectively.

That's a “fun” example to run through in your mind. Did you keep it straight?

Interleaving Recall this scenario from the “Run-to-completion” section of Chapter 1:

```
var a = 1;
var b = 2;

function foo() {
    a++;
    b = b * a;
    a = b + 3;
}

function bar() {
    b--;
    a = 8 + b;
    b = a * 2;
}
```

With normal JS functions, of course either `foo()` can run completely first, or `bar()` can run completely first, but `foo()` cannot interleave its individual statements with `bar()`. So, there are only two possible outcomes to the preceding program.

However, with generators, clearly interleaving (even in the middle of statements!) is possible:

```
var a = 1;
var b = 2;

function *foo() {
    a++;
    yield;
    b = b * a;
    a = (yield b) + 3;
}

function *bar() {
    b--;
    yield;
    a = (yield 8) + b;
    b = a * (yield 2);
}
```

Depending on what respective order the *iterators* controlling `*foo()` and `*bar()` are called, the preceding program could produce several different results. In other words, we can actually illustrate (in a sort of fake-ish way) the theoretical

“threaded race conditions” circumstances discussed in Chapter 1, by interleaving the two generator iterations over the same shared variables.

First, let’s make a helper called `step(..)` that controls an *iterator*:

```
function step(gen) {
  var it = gen();
  var last;

  return function() {
    // whatever is `yield`ed out, just
    // send it right back in the next time!
    last = it.next( last ).value;
  };
}
```

`step(..)` initializes a generator to create its `it` *iterator*, then returns a function which, when called, advances the *iterator* by one step. Additionally, the previously `yielded` out value is sent right back in at the *next* step. So, `yield 8` will just become `8` and `yield b` will just be `b` (whatever it was at the time of `yield`).

Now, just for fun, let’s experiment to see the effects of interleaving these different chunks of `*foo()` and `*bar()`. We’ll start with the boring base case, making sure `*foo()` totally finishes before `*bar()` (just like we did in Chapter 1):

```
// make sure to reset `a` and `b`
a = 1;
b = 2;

var s1 = step( foo );
var s2 = step( bar );

// run `*foo()` completely first
s1();
s1();
s1();

// now run `*bar()`
s2();
s2();
s2();
s2();

console.log( a, b );    // 11 22
```

The end result is `11` and `22`, just as it was in the Chapter 1 version. Now let’s mix up the interleaving ordering and see how it changes the final values of `a` and `b`:

```

// make sure to reset `a` and `b`
a = 1;
b = 2;

var s1 = step( foo );
var s2 = step( bar );

s2();      // b--;
s2();      // yield 8
s1();      // a++;
s2();      // a = 8 + b;
           // yield 2
s1();      // b = b * a;
           // yield b
s1();      // a = b + 3;
s2();      // b = a * 2;

```

Before I tell you the results, can you figure out what `a` and `b` are after the preceding program? No cheating!

```
console.log( a, b );    // 12 18
```

Note: As an exercise for the reader, try to see how many other combinations of results you can get back rearranging the order of the `s1()` and `s2()` calls. Don't forget you'll always need three `s1()` calls and four `s2()` calls. Recall the discussion earlier about matching `next()` with `yield` for the reasons why.

You almost certainly won't want to intentionally create *this* level of interleaving confusion, as it creates incredibly difficult to understand code. But the exercise is interesting and instructive to understand more about how multiple generators can run concurrently in the same shared scope, because there will be places where this capability is quite useful.

We'll discuss generator concurrency in more detail at the end of this chapter.

Generator'ing Values

In the previous section, we mentioned an interesting use for generators, as a way to produce values. This is **not** the main focus in this chapter, but we'd be remiss if we didn't cover the basics, especially because this use case is essentially the origin of the name: generators.

We're going to take a slight diversion into the topic of *iterators* for a bit, but we'll circle back to how they relate to generators and using a generator to *generate* values.

Producers and Iterators

Imagine you're producing a series of values where each value has a definable relationship to the previous value. To do this, you're going to need a stateful producer that remembers the last value it gave out.

You can implement something like that straightforwardly using a function closure (see the *Scope & Closures* title of this series):

```
var gimmeSomething = (function(){
    var nextVal;

    return function(){
        if (nextVal === undefined) {
            nextVal = 1;
        }
        else {
            nextVal = (3 * nextVal) + 6;
        }

        return nextVal;
    };
})();

gimmeSomething();      // 1
gimmeSomething();      // 9
gimmeSomething();      // 33
gimmeSomething();      // 105
```

Note: The `nextVal` computation logic here could have been simplified, but conceptually, we don't want to calculate the *next value* (aka `nextVal`) until the *next* `gimmeSomething()` call happens, because in general that could be a resource-leaky design for producers of more persistent or resource-limited values than simple numbers.

Generating an arbitrary number series isn't a terribly realistic example. But what if you were generating records from a data source? You could imagine much the same code.

In fact, this task is a very common design pattern, usually solved by iterators. An *iterator* is a well-defined interface for stepping through a series of values from a producer. The JS interface for iterators, as it is in most languages, is to call `next()` each time you want the next value from the producer.

We could implement the standard *iterator* interface for our number series producer:

```
var something = (function(){
    var nextVal;
```

```

        return {
            // needed for `for..of` loops
            [Symbol.iterator]: function(){ return this; },

            // standard iterator interface method
            next: function(){
                if (nextVal === undefined) {
                    nextVal = 1;
                }
                else {
                    nextVal = (3 * nextVal) + 6;
                }

                return { done:false, value:nextVal };
            }
        };
    })();

something.next().value;      // 1
something.next().value;      // 9
something.next().value;      // 33
something.next().value;      // 105

```

Note: We'll explain why we need the `[Symbol.iterator]: ...` part of this code snippet in the "Iterables" section. Syntactically though, two ES6 features are at play. First, the `[...]` syntax is called a *computed property name* (see the *this & Object Prototypes* title of this series). It's a way in an object literal definition to specify an expression and use the result of that expression as the name for the property. Next, `Symbol.iterator` is one of ES6's predefined special `Symbol` values (see the *ES6 & Beyond* title of this book series).

The `next()` call returns an object with two properties: `done` is a `boolean` value signaling the *iterator's* complete status; `value` holds the iteration value.

ES6 also adds the `for..of` loop, which means that a standard *iterator* can automatically be consumed with native loop syntax:

```

for (var v of something) {
    console.log( v );

    // don't let the loop run forever!
    if (v > 500) {
        break;
    }
}
// 1 9 33 105 321 969

```

Note: Because our `something iterator` always returns `done:false`, this `for..of` loop would run forever, which is why we put the `break` conditional in. It's totally OK for iterators to be never-ending, but there are also cases where the *iterator* will run over a finite set of values and eventually return a `done:true`.

The `for..of` loop automatically calls `next()` for each iteration – it doesn't pass any values in to the `next()` – and it will automatically terminate on receiving a `done:true`. It's quite handy for looping over a set of data.

Of course, you could manually loop over iterators, calling `next()` and checking for the `done:true` condition to know when to stop:

```
for (
  var ret;
  (ret = something.next()) && !ret.done;
) {
  console.log( ret.value );

  // don't let the loop run forever!
  if (ret.value > 500) {
    break;
  }
}
// 1 9 33 105 321 969
```

Note: This manual `for` approach is certainly uglier than the ES6 `for..of` loop syntax, but its advantage is that it affords you the opportunity to pass in values to the `next(..)` calls if necessary.

In addition to making your own *iterators*, many built-in data structures in JS (as of ES6), like `arrays`, also have default *iterators*:

```
var a = [1,3,5,7,9];

for (var v of a) {
  console.log( v );
}
// 1 3 5 7 9
```

The `for..of` loop asks `a` for its *iterator*, and automatically uses it to iterate over `a`'s values.

Note: It may seem a strange omission by ES6, but regular `objects` intentionally do not come with a default *iterator* the way `arrays` do. The reasons go deeper than we will cover here. If all you want is to iterate over the properties of an object (with no particular guarantee of ordering), `Object.keys(..)` returns an `array`, which can then be used like `for (var k of Object.keys(obj)) { ... }`. Such a `for..of` loop over an object's keys would be similar to a `for..in` loop, except that `Object.keys(..)` does not include properties from the

`[[Prototype]]` chain while `for..in` does (see the *this & Object Prototypes* title of this series).

Iterables

The `something` object in our running example is called an *iterator*, as it has the `next()` method on its interface. But a closely related term is *iterable*, which is an `object` that `contains` an *iterator* that can iterate over its values.

As of ES6, the way to retrieve an *iterator* from an *iterable* is that the *iterable* must have a function on it, with the name being the special ES6 symbol value `Symbol.iterator`. When this function is called, it returns an *iterator*. Though not required, generally each call should return a fresh new *iterator*.

`a` in the previous snippet is an *iterable*. The `for..of` loop automatically calls its `Symbol.iterator` function to construct an *iterator*. But we could of course call the function manually, and use the *iterator* it returns:

```
var a = [1,3,5,7,9];

var it = a[Symbol.iterator]();

it.next().value;    // 1
it.next().value;    // 3
it.next().value;    // 5
..
```

In the previous code listing that defined `something`, you may have noticed this line:

```
[Symbol.iterator]: function(){ return this; }
```

That little bit of confusing code is making the `something` value – the interface of the `something` *iterator* – also an *iterable*; it's now both an *iterable* and an *iterator*. Then, we pass `something` to the `for..of` loop:

```
for (var v of something) {
  ..
}
```

The `for..of` loop expects `something` to be an *iterable*, so it looks for and calls its `Symbol.iterator` function. We defined that function to simply `return this`, so it just gives itself back, and the `for..of` loop is none the wiser.

Generator Iterator

Let's turn our attention back to generators, in the context of *iterators*. A generator can be treated as a producer of values that we extract one at a time through an *iterator* interface's `next()` calls.

So, a generator itself is not technically an *iterable*, though it's very similar – when you execute the generator, you get an *iterator* back:

```
function *foo(){ .. }
```

```
var it = foo();
```

We can implement the `something` infinite number series producer from earlier with a generator, like this:

```
function *something() {
  var nextVal;

  while (true) {
    if (nextVal === undefined) {
      nextVal = 1;
    }
    else {
      nextVal = (3 * nextVal) + 6;
    }

    yield nextVal;
  }
}
```

Note: A `while..true` loop would normally be a very bad thing to include in a real JS program, at least if it doesn't have a `break` or `return` in it, as it would likely run forever, synchronously, and block/lock-up the browser UI. However, in a generator, such a loop is generally totally OK if it has a `yield` in it, as the generator will pause at each iteration, yielding back to the main program and/or to the event loop queue. To put it glibly, “generators put the `while..true` back in JS programming!”

That's a fair bit cleaner and simpler, right? Because the generator pauses at each `yield`, the state (scope) of the function `*something()` is kept around, meaning there's no need for the closure boilerplate to preserve variable state across calls.

Not only is it simpler code – we don't have to make our own *iterator* interface – it actually is more reason-able code, because it more clearly expresses the intent. For example, the `while..true` loop tells us the generator is intended to run forever – to keep *generating* values as long as we keep asking for them.

And now we can use our shiny new `*something()` generator with a `for..of` loop, and you'll see it works basically identically:

```
for (var v of something()) {
  console.log( v );

  // don't let the loop run forever!
```

```

if (v > 500) {
  break;
}
// 1 9 33 105 321 969

```

But don't skip over `for (var v of something())` ...! We didn't just reference `something` as a value like in earlier examples, but instead called the `*something()` generator to get its *iterator* for the `for..of` loop to use.

If you're paying close attention, two questions may arise from this interaction between the generator and the loop:

- Why couldn't we say `for (var v of something())` ...? Because `something` here is a generator, which is not an *iterable*. We have to call `something()` to construct a producer for the `for..of` loop to iterate over.
- The `something()` call produces an *iterator*, but the `for..of` loop wants an *iterable*, right? Yep. The generator's *iterator* also has a `Symbol.iterator` function on it, which basically does a `return this`, just like the `something iterable` we defined earlier. In other words, a generator's *iterator* is also an *iterable*!

Stopping the Generator In the previous example, it would appear the *iterator* instance for the `*something()` generator was basically left in a suspended state forever after the `break` in the loop was called.

But there's a hidden behavior that takes care of that for you. "Abnormal completion" (i.e., "early termination") of the `for..of` loop – generally caused by a `break`, `return`, or an uncaught exception – sends a signal to the generator's *iterator* for it to terminate.

Note: Technically, the `for..of` loop also sends this signal to the *iterator* at the normal completion of the loop. For a generator, that's essentially a moot operation, as the generator's *iterator* had to complete first so the `for..of` loop completed. However, custom *iterators* might desire to receive this additional signal from `for..of` loop consumers.

While a `for..of` loop will automatically send this signal, you may wish to send the signal manually to an *iterator*; you do this by calling `return(...)`.

If you specify a `try..finally` clause inside the generator, it will always be run even when the generator is externally completed. This is useful if you need to clean up resources (database connections, etc.):

```

function *something() {
  try {
    var nextVal;

```

```

        while (true) {
            if (nextVal === undefined) {
                nextVal = 1;
            }
            else {
                nextVal = (3 * nextVal) + 6;
            }

            yield nextVal;
        }
    }
    // cleanup clause
    finally {
        console.log( "cleaning up!" );
    }
}

```

The earlier example with `break` in the `for..of` loop will trigger the `finally` clause. But you could instead manually terminate the generator's *iterator* instance from the outside with `return(..)`:

```

var it = something();
for (var v of it) {
    console.log( v );

    // don't let the loop run forever!
    if (v > 500) {
        console.log(
            // complete the generator's iterator
            it.return( "Hello World" ).value
        );
        // no `break` needed here
    }
}
// 1 9 33 105 321 969
// cleaning up!
// Hello World

```

When we call `it.return(..)`, it immediately terminates the generator, which of course runs the `finally` clause. Also, it sets the returned `value` to whatever you passed in to `return(..)`, which is how "Hello World" comes right back out. We also don't need to include a `break` now because the generator's *iterator* is set to `done:true`, so the `for..of` loop will terminate on its next iteration.

Generators owe their namesake mostly to this *consuming produced values* use. But again, that's just one of the uses for generators, and frankly not even the main one we're concerned with in the context of this book.

But now that we more fully understand some of the mechanics of how they work, we can *next* turn our attention to how generators apply to async concurrency.

Iterating Generators Asynchronously

What do generators have to do with async coding patterns, fixing problems with callbacks, and the like? Let's get to answering that important question.

We should revisit one of our scenarios from Chapter 3. Let's recall the callback approach:

```
function foo(x,y,cb) {
  ajax(
    "http://some.url.1/?x=" + x + "&y=" + y,
    cb
  );
}

foo( 11, 31, function(err,text) {
  if (err) {
    console.error( err );
  }
  else {
    console.log( text );
  }
} );
```

If we wanted to express this same task flow control with a generator, we could do:

```
function foo(x,y) {
  ajax(
    "http://some.url.1/?x=" + x + "&y=" + y,
    function(err,data){
      if (err) {
        // throw an error into `*main()`
        it.throw( err );
      }
      else {
        // resume `*main()` with received `data`
        it.next( data );
      }
    }
  );
}

function *main() {
  try {
```

```

        var text = yield foo( 11, 31 );
        console.log( text );
    }
    catch (err) {
        console.error( err );
    }
}

var it = main();

// start it all up!
it.next();

```

At first glance, this snippet is longer, and perhaps a little more complex looking, than the callback snippet before it. But don't let that impression get you off track. The generator snippet is actually **much** better! But there's a lot going on for us to explain.

First, let's look at this part of the code, which is the most important:

```

var text = yield foo( 11, 31 );
console.log( text );

```

Think about how that code works for a moment. We're calling a normal function `foo(..)` and we're apparently able to get back the `text` from the Ajax call, even though it's asynchronous.

How is that possible? If you recall the beginning of Chapter 1, we had almost identical code:

```

var data = ajax( "..url 1.." );
console.log( data );

```

And that code didn't work! Can you spot the difference? It's the `yield` used in a generator.

That's the magic! That's what allows us to have what appears to be blocking, synchronous code, but it doesn't actually block the whole program; it only pauses/blocks the code in the generator itself.

In `yield foo(11,31)`, first the `foo(11,31)` call is made, which returns nothing (aka `undefined`), so we're making a call to request data, but we're actually then doing `yield undefined`. That's OK, because the code is not currently relying on a `yielded` value to do anything interesting. We'll revisit this point later in the chapter.

We're not using `yield` in a message passing sense here, only in a flow control sense to pause/block. Actually, it will have message passing, but only in one direction, after the generator is resumed.

So, the generator pauses at the `yield`, essentially asking the question, "what

value should I return to assign to the variable `text`?” Who’s going to answer that question?

Look at `foo(..)`. If the Ajax request is successful, we call:

```
it.next( data );
```

That’s resuming the generator with the response data, which means that our paused `yield` expression receives that value directly, and then as it restarts the generator code, that value gets assigned to the local variable `text`.

Pretty cool, huh?

Take a step back and consider the implications. We have totally synchronous-looking code inside the generator (other than the `yield` keyword itself), but hidden behind the scenes, inside of `foo(..)`, the operations can complete asynchronously.

That’s huge! That’s a nearly perfect solution to our previously stated problem with callbacks not being able to express asynchrony in a sequential, synchronous fashion that our brains can relate to.

In essence, we are abstracting the asynchrony away as an implementation detail, so that we can reason synchronously/sequentially about our flow control: “Make an Ajax request, and when it finishes print out the response.” And of course, we just expressed two steps in the flow control, but this same capability extends without bounds, to let us express however many steps we need to.

Tip: This is such an important realization, just go back and read the last three paragraphs again to let it sink in!

Synchronous Error Handling

But the preceding generator code has even more goodness to `yield` to us. Let’s turn our attention to the `try..catch` inside the generator:

```
try {
  var text = yield foo( 11, 31 );
  console.log( text );
}
catch (err) {
  console.error( err );
}
```

How does this work? The `foo(..)` call is asynchronously completing, and doesn’t `try..catch` fail to catch asynchronous errors, as we looked at in Chapter 3?

We already saw how the `yield` lets the assignment statement pause to wait for `foo(..)` to finish, so that the completed response can be assigned to `text`. The awesome part is that this `yield` pausing *also* allows the generator to `catch` an

error. We throw that error into the generator with this part of the earlier code listing:

```
if (err) {
  // throw an error into `*main()`
  it.throw( err );
}
```

The `yield`-pause nature of generators means that not only do we get synchronous-looking `return` values from `async` function calls, but we can also synchronously `catch` errors from those `async` function calls!

So we've seen we can throw errors *into* a generator, but what about throwing errors *out of* a generator? Exactly as you'd expect:

```
function *main() {
  var x = yield "Hello World";

  yield x.toLowerCase(); // cause an exception!
}

var it = main();

it.next().value; // Hello World

try {
  it.next( 42 );
}
catch (err) {
  console.error( err ); // TypeError
}
```

Of course, we could have manually thrown an error with `throw ..` instead of causing an exception.

We can even `catch` the same error that we `throw(..)` into the generator, essentially giving the generator a chance to handle it but if it doesn't, the `iterator` code must handle it:

```
function *main() {
  var x = yield "Hello World";

  // never gets here
  console.log( x );
}

var it = main();

it.next();
```

```

try {
  // will `main()` handle this error? we'll see!
  it.throw( "Oops" );
}
catch (err) {
  // nope, didn't handle it!
  console.error( err );           // Oops
}

```

Synchronous-looking error handling (via `try..catch`) with async code is a huge win for readability and reason-ability.

Generators + Promises

In our previous discussion, we showed how generators can be iterated asynchronously, which is a huge step forward in sequential reason-ability over the spaghetti mess of callbacks. But we lost something very important: the trustability and composable of Promises (see Chapter 3)!

Don't worry – we can get that back. The best of all worlds in ES6 is to combine generators (synchronous-looking async code) with Promises (trustable and composable).

But how?

Recall from Chapter 3 the Promise-based approach to our running Ajax example:

```

function foo(x,y) {
  return request(
    "http://some.url.1/?x=" + x + "&y=" + y
  );
}

foo( 11, 31 )
.then(
  function(text){
    console.log( text );
  },
  function(err){
    console.error( err );
  }
);

```

In our earlier generator code for the running Ajax example, `foo(..)` returned nothing (`undefined`), and our *iterator* control code didn't care about that `yielded` value.

But here the Promise-aware `foo(..)` returns a promise after making the Ajax call. That suggests that we could construct a promise with `foo(..)` and then `yield` it from the generator, and then the *iterator* control code would receive that promise.

But what should the *iterator* do with the promise?

It should listen for the promise to resolve (fulfillment or rejection), and then either resume the generator with the fulfillment message or throw an error into the generator with the rejection reason.

Let me repeat that, because it's so important. The natural way to get the most out of Promises and generators is **to yield a Promise**, and wire that Promise to control the generator's *iterator*.

Let's give it a try! First, we'll put the Promise-aware `foo(..)` together with the generator `*main()`:

```
function foo(x,y) {
  return request(
    "http://some.url.1/?x=" + x + "&y=" + y
  );
}

function *main() {
  try {
    var text = yield foo( 11, 31 );
    console.log( text );
  }
  catch (err) {
    console.error( err );
  }
}
```

The most powerful revelation in this refactor is that the code inside `*main()` **did not have to change at all!** Inside the generator, whatever values are yielded out is just an opaque implementation detail, so we're not even aware it's happening, nor do we need to worry about it.

But how are we going to run `*main()` now? We still have some of the implementation plumbing work to do, to receive and wire up the yielded promise so that it resumes the generator upon resolution. We'll start by trying that manually:

```
var it = main();

var p = it.next().value;

// wait for the `p` promise to resolve
p.then(
  function(text){
```

```

        it.next( text );
    },
    function(err){
        it.throw( err );
    }
);

```

Actually, that wasn't so painful at all, was it?

This snippet should look very similar to what we did earlier with the manually wired generator controlled by the error-first callback. Instead of an `if (err) { it.throw...`, the promise already splits fulfillment (success) and rejection (failure) for us, but otherwise the *iterator* control is identical.

Now, we've glossed over some important details.

Most importantly, we took advantage of the fact that we knew that `*main()` only had one Promise-aware step in it. What if we wanted to be able to Promise-drive a generator no matter how many steps it has? We certainly don't want to manually write out the Promise chain differently for each generator! What would be much nicer is if there was a way to repeat (aka "loop" over) the iteration control, and each time a Promise comes out, wait on its resolution before continuing.

Also, what if the generator throws out an error (intentionally or accidentally) during the `it.next(..)` call? Should we quit, or should we `catch` it and send it right back in? Similarly, what if we `it.throw(..)` a Promise rejection into the generator, but it's not handled, and comes right back out?

Promise-Aware Generator Runner

The more you start to explore this path, the more you realize, "wow, it'd be great if there was just some utility to do it for me." And you're absolutely correct. This is such an important pattern, and you don't want to get it wrong (or exhaust yourself repeating it over and over), so your best bet is to use a utility that is specifically designed to *run* Promise-yielding generators in the manner we've illustrated.

Several Promise abstraction libraries provide just such a utility, including my `asynquence` library and its `runner(..)`, which will be discussed in Appendix A of this book.

But for the sake of learning and illustration, let's just define our own standalone utility that we'll call `run(..)`:

```

// thanks to Benjamin Gruenbaum (@benjamingr on GitHub) for
// big improvements here!
function run(gen) {
    var args = [].slice.call( arguments, 1 ), it;

```

```

// initialize the generator in the current context
it = gen.apply( this, args );

// return a promise for the generator completing
return Promise.resolve()
  .then( function handleNext(value){
    // run to the next yielded value
    var next = it.next( value );

    return (function handleResult(next){
      // generator has completed running?
      if (next.done) {
        return next.value;
      }
      // otherwise keep going
      else {
        return Promise.resolve( next.value )
          .then(
            // resume the async loop on
            // success, sending the resolved
            // value back into the generator
            handleNext,
            // if `value` is a rejected
            // promise, propagate error back
            // into the generator for its own
            // error handling
            function handleErr(err) {
              return Promise.resolve(
                it.throw( err )
              )
              .then( handleResult );
            }
          );
      }
    })(next);
  } );
}

```

As you can see, it's a quite a bit more complex than you'd probably want to author yourself, and you especially wouldn't want to repeat this code for each generator you use. So, a utility/library helper is definitely the way to go. Nevertheless, I encourage you to spend a few minutes studying that code listing to get a better sense of how to manage the generator+Promise negotiation.

How would you use `run(...)` with `*main()` in our *running* Ajax example?

```

function *main() {
  // ...
}

run( main );

```

That's it! The way we wired `run(..)`, it will automatically advance the generator you pass to it, asynchronously until completion.

Note: The `run(..)` we defined returns a promise which is wired to resolve once the generator is complete, or receive an uncaught exception if the generator doesn't handle it. We don't show that capability here, but we'll come back to it later in the chapter.

ES7: `async` and `await`? The preceding pattern – generators yielding Promises that then control the generator's *iterator* to advance it to completion – is such a powerful and useful approach, it would be nicer if we could do it without the clutter of the library utility helper (aka `run(..)`).

There's probably good news on that front. At the time of this writing, there's early but strong support for a proposal for more syntactic addition in this realm for the post-ES6, ES7-ish timeframe. Obviously, it's too early to guarantee the details, but there's a pretty decent chance it will shake out similar to the following:

```

function foo(x,y) {
  return request(
    "http://some.url.1/?x=" + x + "&y=" + y
  );
}

async function main() {
  try {
    var text = await foo( 11, 31 );
    console.log( text );
  }
  catch (err) {
    console.error( err );
  }
}

main();

```

As you can see, there's no `run(..)` call (meaning no need for a library utility!) to invoke and drive `main()` – it's just called as a normal function. Also, `main()` isn't declared as a generator function anymore; it's a new kind of function: `async function`. And finally, instead of `yielding` a Promise, we `await` for it to resolve.

The `async` function automatically knows what to do if you `await` a Promise – it will pause the function (just like with generators) until the Promise resolves. We didn't illustrate it in this snippet, but calling an `async` function like `main()` automatically returns a promise that's resolved whenever the function finishes completely.

Tip: The `async` / `await` syntax should look very familiar to readers with experience in C#, because it's basically identical.

The proposal essentially codifies support for the pattern we've already derived, into a syntactic mechanism: combining Promises with sync-looking flow control code. That's the best of both worlds combined, to effectively address practically all of the major concerns we outlined with callbacks.

The mere fact that such a ES7-ish proposal already exists and has early support and enthusiasm is a major vote of confidence in the future importance of this `async` pattern.

Promise Concurrency in Generators

So far, all we've demonstrated is a single-step `async` flow with Promises+generators. But real-world code will often have many `async` steps.

If you're not careful, the sync-looking style of generators may lull you into complacency with how you structure your `async` concurrency, leading to suboptimal performance patterns. So we want to spend a little time exploring the options.

Imagine a scenario where you need to fetch data from two different sources, then combine those responses to make a third request, and finally print out the last response. We explored a similar scenario with Promises in Chapter 3, but let's reconsider it in the context of generators.

Your first instinct might be something like:

```
function *foo() {
  var r1 = yield request( "http://some.url.1" );
  var r2 = yield request( "http://some.url.2" );

  var r3 = yield request(
    "http://some.url.3/?v=" + r1 + "," + r2
  );

  console.log( r3 );
}

// use previously defined `run(..)` utility
run( foo );
```

This code will work, but in the specifics of our scenario, it's not optimal. Can you spot why?

Because the `r1` and `r2` requests can – and for performance reasons, *should* – run concurrently, but in this code they will run sequentially; the "`http://some.url.2`" URL isn't Ajax fetched until after the "`http://some.url.1`" request is finished. These two requests are independent, so the better performance approach would likely be to have them run at the same time.

But how exactly would you do that with a generator and `yield`? We know that `yield` is only a single pause point in the code, so you can't really do two pauses at the same time.

The most natural and effective answer is to base the async flow on Promises, specifically on their capability to manage state in a time-independent fashion (see “Future Value” in Chapter 3).

The simplest approach:

```
function *foo() {
  // make both requests "in parallel"
  var p1 = request( "http://some.url.1" );
  var p2 = request( "http://some.url.2" );

  // wait until both promises resolve
  var r1 = yield p1;
  var r2 = yield p2;

  var r3 = yield request(
    "http://some.url.3/?v=" + r1 + "," + r2
  );

  console.log( r3 );
}

// use previously defined `run(..)` utility
run( foo );
```

Why is this different from the previous snippet? Look at where the `yield` is and is not. `p1` and `p2` are promises for Ajax requests made concurrently (aka “in parallel”). It doesn't matter which one finishes first, because promises will hold onto their resolved state for as long as necessary.

Then we use two subsequent `yield` statements to wait for and retrieve the resolutions from the promises (into `r1` and `r2`, respectively). If `p1` resolves first, the `yield p1` resumes first then waits on the `yield p2` to resume. If `p2` resolves first, it will just patiently hold onto that resolution value until asked, but the `yield p1` will hold on first, until `p1` resolves.

Either way, both `p1` and `p2` will run concurrently, and both have to finish, in either order, before the `r3 = yield request..` Ajax request will be made.

If that flow control processing model sounds familiar, it's basically the same

as what we identified in Chapter 3 as the “gate” pattern, enabled by the `Promise.all([..])` utility. So, we could also express the flow control like this:

```
function *foo() {
  // make both requests "in parallel," and
  // wait until both promises resolve
  var results = yield Promise.all( [
    request( "http://some.url.1" ),
    request( "http://some.url.2" )
  ]);

  var r1 = results[0];
  var r2 = results[1];

  var r3 = yield request(
    "http://some.url.3/?v=" + r1 + "," + r2
  );

  console.log( r3 );
}

// use previously defined `run(..)` utility
run( foo );
```

Note: As we discussed in Chapter 3, we can even use ES6 destructuring assignment to simplify the `var r1 = .. var r2 = ..` assignments, with `var [r1,r2] = results`.

In other words, all of the concurrency capabilities of Promises are available to us in the generator+Promise approach. So in any place where you need more than sequential this-then-that async flow control steps, Promises are likely your best bet.

Promises, Hidden As a word of stylistic caution, be careful about how much Promise logic you include **inside your generators**. The whole point of using generators for asynchrony in the way we’ve described is to create simple, sequential, sync-looking code, and to hide as much of the details of asynchrony away from that code as possible.

For example, this might be a cleaner approach:

```
// note: normal function, not generator
function bar(url1,url2) {
  return Promise.all( [
    request( url1 ),
    request( url2 )
  ]);
}
```

```

}

function *foo() {
  // hide the Promise-based concurrency details
  // inside `bar(..)`
  var results = yield bar(
    "http://some.url.1",
    "http://some.url.2"
  );

  var r1 = results[0];
  var r2 = results[1];

  var r3 = yield request(
    "http://some.url.3/?v=" + r1 + "," + r2
  );

  console.log( r3 );
}

// use previously defined `run(..)` utility
run( foo );

```

Inside `*foo()`, it's cleaner and clearer that all we're doing is just asking `bar(..)` to get us some `results`, and we'll `yield`-wait on that to happen. We don't have to care that under the covers a `Promise.all([..])` Promise composition will be used to make that happen.

We treat asynchrony, and indeed Promises, as an implementation detail.

Hiding your Promise logic inside a function that you merely call from your generator is especially useful if you're going to do a sophisticated series flow-control. For example:

```

function bar() {
  return Promise.all( [
    baz( .. )
    .then( .. ),
    Promise.race( [ .. ] )
  ] )
  .then( .. )
}

```

That kind of logic is sometimes required, and if you dump it directly inside your generator(s), you've defeated most of the reason why you would want to use generators in the first place. We *should* intentionally abstract such details away from our generator code so that they don't clutter up the higher level task

expression.

Beyond creating code that is both functional and performant, you should also strive to make code that is as reasonable and maintainable as possible.

Note: Abstraction is not *always* a healthy thing for programming – many times it can increase complexity in exchange for terseness. But in this case, I believe it's much healthier for your generator+Promise async code than the alternatives. As with all such advice, though, pay attention to your specific situations and make proper decisions for you and your team.

Generator Delegation

In the previous section, we showed calling regular functions from inside a generator, and how that remains a useful technique for abstracting away implementation details (like async Promise flow). But the main drawback of using a normal function for this task is that it has to behave by the normal function rules, which means it cannot pause itself with `yield` like a generator can.

It may then occur to you that you might try to call one generator from another generator, using our `run(..)` helper, such as:

```
function *foo() {
  var r2 = yield request( "http://some.url.2" );
  var r3 = yield request( "http://some.url.3/?v=" + r2 );

  return r3;
}

function *bar() {
  var r1 = yield request( "http://some.url.1" );

  // "delegating" to `*foo()` via `run(..)`
  var r3 = yield run( foo );

  console.log( r3 );
}

run( bar );
```

We run `*foo()` inside of `*bar()` by using our `run(..)` utility again. We take advantage here of the fact that the `run(..)` we defined earlier returns a promise which is resolved when its generator is run to completion (or errors out), so if we `yield` out to a `run(..)` instance the promise from another `run(..)` call, it automatically pauses `*bar()` until `*foo()` finishes.

But there's an even better way to integrate calling `*foo()` into `*bar()`, and it's called `yield`-delegation. The special syntax for `yield`-delegation is: `yield *`

-- (notice the extra *). Before we see it work in our previous example, let's look at a simpler scenario:

```
function *foo() {
  console.log(`~*foo()~ starting`);
  yield 3;
  yield 4;
  console.log(`~*foo()~ finished`);
}

function *bar() {
  yield 1;
  yield 2;
  yield *foo();    // `yield~-delegation!
  yield 5;
}

var it = bar();

it.next().value;    // 1
it.next().value;    // 2
it.next().value;    // `~*foo()~ starting
                    // 3
it.next().value;    // 4
it.next().value;    // `~*foo()~ finished
                    // 5
```

Note: Similar to a note earlier in the chapter where I explained why I prefer `function *foo() ...` instead of `function* foo() ...`, I also prefer – differing from most other documentation on the topic – to say `yield *foo()` instead of `yield* foo()`. The placement of the `*` is purely stylistic and up to your best judgment. But I find the consistency of styling attractive.

How does the `yield *foo()` delegation work?

First, calling `foo()` creates an *iterator* exactly as we've already seen. Then, `yield *` delegates/transfers the *iterator* instance control (of the present `*bar()` generator) over to this other `*foo()` *iterator*.

So, the first two `it.next()` calls are controlling `*bar()`, but when we make the third `it.next()` call, now `*foo()` starts up, and now we're controlling `*foo()` instead of `*bar()`. That's why it's called delegation – `*bar()` delegated its iteration control to `*foo()`.

As soon as the `it iterator` control exhausts the entire `*foo() iterator`, it automatically returns to controlling `*bar()`.

So now back to the previous example with the three sequential Ajax requests:

```
function *foo() {
```

```

var r2 = yield request( "http://some.url.2" );
var r3 = yield request( "http://some.url.3/?v=" + r2 );

return r3;
}

function *bar() {
  var r1 = yield request( "http://some.url.1" );

  // "delegating" to `*foo()` via `yield*`
  var r3 = yield *foo();

  console.log( r3 );
}

run( bar );

```

The only difference between this snippet and the version used earlier is the use of `yield *foo()` instead of the previous `yield run(foo)`.

Note: `yield *` yields iteration control, not generator control; when you invoke the `*foo()` generator, you're now `yield`-delegating to its *iterator*. But you can actually `yield`-delegate to any *iterable*; `yield *[1,2,3]` would consume the default *iterator* for the `[1,2,3]` array value.

Why Delegation?

The purpose of `yield`-delegation is mostly code organization, and in that way is symmetrical with normal function calling.

Imagine two modules that respectively provide methods `foo()` and `bar()`, where `bar()` calls `foo()`. The reason the two are separate is generally because the proper organization of code for the program calls for them to be in separate functions. For example, there may be cases where `foo()` is called standalone, and other places where `bar()` calls `foo()`.

For all these exact same reasons, keeping generators separate aids in program readability, maintenance, and debuggability. In that respect, `yield *` is a syntactic shortcut for manually iterating over the steps of `*foo()` while inside of `*bar()`.

Such manual approach would be especially complex if the steps in `*foo()` were asynchronous, which is why you'd probably need to use that `run(..)` utility to do it. And as we've shown, `yield *foo()` eliminates the need for a sub-instance of the `run(..)` utility (like `run(foo)`).

Delegating Messages

You may wonder how this `yield`-delegation works not just with *iterator* control but with the two-way message passing. Carefully follow the flow of messages in and out, through the `yield`-delegation:

```
function *foo() {
  console.log( "inside `*foo()`:", yield "B" );

  console.log( "inside `*foo()`:", yield "C" );

  return "D";
}

function *bar() {
  console.log( "inside `*bar()`:", yield "A" );

  // `yield`-delegation!
  console.log( "inside `*bar()`:", yield *foo() );

  console.log( "inside `*bar()`:", yield "E" );

  return "F";
}

var it = bar();

console.log( "outside:", it.next().value );
// outside: A

console.log( "outside:", it.next( 1 ).value );
// inside `*bar()`: 1
// outside: B

console.log( "outside:", it.next( 2 ).value );
// inside `*foo()`: 2
// outside: C

console.log( "outside:", it.next( 3 ).value );
// inside `*foo()`: 3
// inside `*bar()`: D
// outside: E

console.log( "outside:", it.next( 4 ).value );
// inside `*bar()`: 4
// outside: F
```

Pay particular attention to the processing steps after the `it.next(3)` call:

1. The `3` value is passed (through the `yield`-delegation in `*bar()`) into the waiting `yield "C"` expression inside of `*foo()`.
2. `*foo()` then calls `return "D"`, but this value doesn't get returned all the way back to the outside `it.next(3)` call.
3. Instead, the `"D"` value is sent as the result of the waiting `yield *foo()` expression inside of `*bar()` – this `yield`-delegation expression has essentially been paused while all of `*foo()` was exhausted. So `"D"` ends up inside of `*bar()` for it to print out.
4. `yield "E"` is called inside of `*bar()`, and the `"E"` value is yielded to the outside as the result of the `it.next(3)` call.

From the perspective of the external *iterator* (`it`), it doesn't appear any differently between controlling the initial generator or a delegated one.

In fact, `yield`-delegation doesn't even have to be directed to another generator; it can just be directed to a non-generator, general *iterable*. For example:

```
function *bar() {
  console.log( "inside `*bar()`:", yield "A" );

  // `yield`-delegation to a non-generator!
  console.log( "inside `*bar()`:", yield *[ "B", "C", "D" ] );

  console.log( "inside `*bar()`:", yield "E" );

  return "F";
}

var it = bar();

console.log( "outside:", it.next().value );
// outside: A

console.log( "outside:", it.next( 1 ).value );
// inside `*bar()`: 1
// outside: B

console.log( "outside:", it.next( 2 ).value );
// outside: C

console.log( "outside:", it.next( 3 ).value );
// outside: D

console.log( "outside:", it.next( 4 ).value );
// inside `*bar()`: undefined
// outside: E
```

```

console.log( "outside:", it.next( 5 ).value );
// inside `*bar()`: 5
// outside: F

```

Notice the differences in where the messages were received/reported between this example and the one previous.

Most strikingly, the default *array iterator* doesn't care about any messages sent in via `next(..)` calls, so the values 2, 3, and 4 are essentially ignored. Also, because that *iterator* has no explicit `return` value (unlike the previously used `*foo()`), the `yield *` expression gets an `undefined` when it finishes.

Exceptions Delegated, Too! In the same way that `yield`-delegation transparently passes messages through in both directions, errors/exceptions also pass in both directions:

```

function *foo() {
  try {
    yield "B";
  }
  catch (err) {
    console.log( "error caught inside `*foo()`:", err );
  }

  yield "C";

  throw "D";
}

function *bar() {
  yield "A";

  try {
    yield *foo();
  }
  catch (err) {
    console.log( "error caught inside `*bar()`:", err );
  }

  yield "E";

  yield *baz();

  // note: can't get here!
  yield "G";
}

```

```

function *baz() {
    throw "F";
}

var it = bar();

console.log( "outside:", it.next().value );
// outside: A

console.log( "outside:", it.next( 1 ).value );
// outside: B

console.log( "outside:", it.throw( 2 ).value );
// error caught inside `*foo()`: 2
// outside: C

console.log( "outside:", it.next( 3 ).value );
// error caught inside `*bar()`: D
// outside: E

try {
    console.log( "outside:", it.next( 4 ).value );
}
catch (err) {
    console.log( "error caught outside:", err );
}
// error caught outside: F

```

Some things to note from this snippet:

1. When we call `it.throw(2)`, it sends the error message 2 into `*bar()`, which delegates that to `*foo()`, which then catches it and handles it gracefully. Then, the `yield "C"` sends "C" back out as the return value from the `it.throw(2)` call.
2. The "D" value that's next thrown from inside `*foo()` propagates out to `*bar()`, which catches it and handles it gracefully. Then the `yield "E"` sends "E" back out as the return value from the `it.next(3)` call.
3. Next, the exception thrown from `*baz()` isn't caught in `*bar()` – though we did catch it outside – so both `*baz()` and `*bar()` are set to a completed state. After this snippet, you would not be able to get the "G" value out with any subsequent `next(..)` call(s) – they will just return `undefined` for `value`.

Delegating Asynchrony

Let's finally get back to our earlier `yield`-delegation example with the multiple sequential Ajax requests:

```
function *foo() {
  var r2 = yield request( "http://some.url.2" );
  var r3 = yield request( "http://some.url.3/?v=" + r2 );

  return r3;
}

function *bar() {
  var r1 = yield request( "http://some.url.1" );

  var r3 = yield *foo();

  console.log( r3 );
}

run( bar );
```

Instead of calling `yield run(foo)` inside of `*bar()`, we just call `yield *foo()`.

In the previous version of this example, the Promise mechanism (controlled by `run(..)`) was used to transport the value from `return r3` in `*foo()` to the local variable `r3` inside `*bar()`. Now, that value is just returned back directly via the `yield *` mechanics.

Otherwise, the behavior is pretty much identical.

Delegating “Recursion”

Of course, `yield`-delegation can keep following as many delegation steps as you wire up. You could even use `yield`-delegation for `async`-capable generator “recursion” – a generator `yield`-delegating to itself:

```
function *foo(val) {
  if (val > 1) {
    // generator recursion
    val = yield *foo( val - 1 );
  }

  return yield request( "http://some.url/?v=" + val );
}

function *bar() {
  var r1 = yield *foo( 3 );
  console.log( r1 );
```

```
}

run( bar );
```

Note: Our `run(..)` utility could have been called with `run(foo, 3)`, because it supports additional parameters being passed along to the initialization of the generator. However, we used a parameter-free `*bar()` here to highlight the flexibility of `yield *`.

What processing steps follow from that code? Hang on, this is going to be quite intricate to describe in detail:

1. `run(bar)` starts up the `*bar()` generator.
2. `foo(3)` creates an *iterator* for `*foo(..)` and passes 3 as its `val` parameter.
3. Because $3 > 1$, `foo(2)` creates another *iterator* and passes in 2 as its `val` parameter.
4. Because $2 > 1$, `foo(1)` creates yet another *iterator* and passes in 1 as its `val` parameter.
5. $1 > 1$ is `false`, so we next call `request(..)` with the 1 value, and get a promise back for that first Ajax call.
6. That promise is `yielded` out, which comes back to the `*foo(2)` generator instance.
7. The `yield *` passes that promise back out to the `*foo(3)` generator instance. Another `yield *` passes the promise out to the `*bar()` generator instance. And yet again another `yield *` passes the promise out to the `run(..)` utility, which will wait on that promise (for the first Ajax request) to proceed.
8. When the promise resolves, its fulfillment message is sent to resume `*bar()`, which passes through the `yield *` into the `*foo(3)` instance, which then passes through the `yield *` to the `*foo(2)` generator instance, which then passes through the `yield *` to the normal `yield` that's waiting in the `*foo(3)` generator instance.
9. That first call's Ajax response is now immediately `returned` from the `*foo(3)` generator instance, which sends that value back as the result of the `yield *` expression in the `*foo(2)` instance, and assigned to its local `val` variable.
10. Inside `*foo(2)`, a second Ajax request is made with `request(..)`, whose promise is `yielded` back to the `*foo(1)` instance, and then `yield *` propagates all the way out to `run(..)` (step 7 again). When the promise resolves, the second Ajax response propagates all the way back into the `*foo(2)` generator instance, and is assigned to its local `val` variable.
11. Finally, the third Ajax request is made with `request(..)`, its promise goes out to `run(..)`, and then its resolution value comes all the way back, which is then `returned` so that it comes back to the waiting `yield *` expression in `*bar()`.

Phew! A lot of crazy mental juggling, huh? You might want to read through that a few more times, and then go grab a snack to clear your head!

Generator Concurrency

As we discussed in both Chapter 1 and earlier in this chapter, two simultaneously running “processes” can cooperatively interleave their operations, and many times this can *yield* (pun intended) very powerful asynchrony expressions.

Frankly, our earlier examples of concurrency interleaving of multiple generators showed how to make it really confusing. But we hinted that there’s places where this capability is quite useful.

Recall a scenario we looked at in Chapter 1, where two different simultaneous Ajax response handlers needed to coordinate with each other to make sure that the data communication was not a race condition. We slotted the responses into the `res` array like this:

```
function response(data) {
  if (data.url == "http://some.url.1") {
    res[0] = data;
  }
  else if (data.url == "http://some.url.2") {
    res[1] = data;
  }
}
```

But how can we use multiple generators concurrently for this scenario?

```
// `request(..)` is a Promise-aware Ajax utility

var res = [];

function *reqData(url) {
  res.push(
    yield request(url)
  );
}
```

Note: We’re going to use two instances of the `*reqData(..)` generator here, but there’s no difference to running a single instance of two different generators; both approaches are reasoned about identically. We’ll see two different generators coordinating in just a bit.

Instead of having to manually sort out `res[0]` and `res[1]` assignments, we’ll use coordinated ordering so that `res.push(..)` properly slots the values in the expected and predictable order. The expressed logic thus should feel a bit cleaner.

But how will we actually orchestrate this interaction? First, let’s just do it manually, with Promises:

```
var it1 = reqData("http://some.url.1");
```

```

var it2 = reqData( "http://some.url.2" );

var p1 = it1.next().value;
var p2 = it2.next().value;

p1
.then( function(data){
  it1.next( data );
  return p2;
} )
.then( function(data){
  it2.next( data );
} );

```

*reqData(..)'s two instances are both started to make their Ajax requests, then paused with `yield`. Then we choose to resume the first instance when `p1` resolves, and then `p2`'s resolution will restart the second instance. In this way, we use Promise orchestration to ensure that `res[0]` will have the first response and `res[1]` will have the second response.

But frankly, this is awfully manual, and it doesn't really let the generators orchestrate themselves, which is where the true power can lie. Let's try it a different way:

```

// `request(..)` is a Promise-aware Ajax utility

var res = [];

function *reqData(url) {
  var data = yield request( url );

  // transfer control
  yield;

  res.push( data );
}

var it1 = reqData( "http://some.url.1" );
var it2 = reqData( "http://some.url.2" );

var p1 = it1.next().value;
var p2 = it2.next().value;

p1.then( function(data){
  it1.next( data );
} );

```

```

p2.then( function(data){
    it2.next( data );
} );
Promise.all( [p1,p2] )
.then( function(){
    it1.next();
    it2.next();
} );

```

OK, this is a bit better (though still manual!), because now the two instances of `*reqData(..)` run truly concurrently, and (at least for the first part) independently.

In the previous snippet, the second instance was not given its data until after the first instance was totally finished. But here, both instances receive their data as soon as their respective responses come back, and then each instance does another `yield` for control transfer purposes. We then choose what order to resume them in the `Promise.all([..])` handler.

What may not be as obvious is that this approach hints at an easier form for a reusable utility, because of the symmetry. We can do even better. Let's imagine using a utility called `runAll(..)`:

```

// `request(..)` is a Promise-aware Ajax utility

var res = [];
runAll(
    function(){
        var p1 = request( "http://some.url.1" );

        // transfer control
        yield;

        res.push( yield p1 );
    },
    function(){
        var p2 = request( "http://some.url.2" );

        // transfer control
        yield;

        res.push( yield p2 );
    }
);

```

Note: We're not including a code listing for `runAll(..)` as it is not only long

enough to bog down the text, but is an extension of the logic we've already implemented in `run(..)` earlier. So, as a good supplementary exercise for the reader, try your hand at evolving the code from `run(..)` to work like the imagined `runAll(..)`. Also, my *asynquence* library provides a previously mentioned `runner(..)` utility with this kind of capability already built in, and will be discussed in Appendix A of this book.

Here's how the processing inside `runAll(..)` would operate:

1. The first generator gets a promise for the first Ajax response from "`http://some.url.1`", then `yields` control back to the `runAll(..)` utility.
2. The second generator runs and does the same for "`http://some.url.2`", `yielding` control back to the `runAll(..)` utility.
3. The first generator resumes, and then `yields` out its promise `p1`. The `runAll(..)` utility does the same in this case as our previous `run(..)`, in that it waits on that promise to resolve, then resumes the same generator (no control transfer!). When `p1` resolves, `runAll(..)` resumes the first generator again with that resolution value, and then `res[0]` is given its value. When the first generator then finishes, that's an implicit transfer of control.
4. The second generator resumes, `yields` out its promise `p2`, and waits for it to resolve. Once it does, `runAll(..)` resumes the second generator with that value, and `res[1]` is set.

In this running example, we use an outer variable called `res` to store the results of the two different Ajax responses – that's our concurrency coordination making that possible.

But it might be quite helpful to further extend `runAll(..)` to provide an inner variable space for the multiple generator instances to *share*, such as an empty object we'll call `data` below. Also, it could take non-Promise values that are `yielded` and hand them off to the next generator.

Consider:

```
// `request(..)` is a Promise-aware Ajax utility

runAll(
  function*(data){
    data.res = [];

    // transfer control (and message pass)
    var url1 = yield "http://some.url.2";

    var p1 = request( url1 ); // "http://some.url.1"

    // transfer control
    yield;
```

```

        data.res.push( yield p1 );
    },
    function*(data){
        // transfer control (and message pass)
        var url2 = yield "http://some.url.1";

        var p2 = request( url2 ); // "http://some.url.2"

        // transfer control
        yield;

        data.res.push( yield p2 );
    }
);

```

In this formulation, the two generators are not just coordinating control transfer, but actually communicating with each other, both through `data.res` and the yielded messages that trade `url1` and `url2` values. That's incredibly powerful!

Such realization also serves as a conceptual base for a more sophisticated asynchrony technique called CSP (Communicating Sequential Processes), which we will cover in Appendix B of this book.

Thunks

So far, we've made the assumption that yielding a Promise from a generator – and having that Promise resume the generator via a helper utility like `run(...)` – was the best possible way to manage asynchrony with generators. To be clear, it is.

But we skipped over another pattern that has some mildly widespread adoption, so in the interest of completeness we'll take a brief look at it.

In general computer science, there's an old pre-JS concept called a "thunk." Without getting bogged down in the historical nature, a narrow expression of a thunk in JS is a function that – without any parameters – is wired to call another function.

In other words, you wrap a function definition around function call – with any parameters it needs – to *defer* the execution of that call, and that wrapping function is a thunk. When you later execute the thunk, you end up calling the original function.

For example:

```

function foo(x,y) {
    return x + y;
}

```

```

function fooThunk() {
    return foo( 3, 4 );
}

// later

```

```
console.log( fooThunk() ); // 7
```

So, a synchronous thunk is pretty straightforward. But what about an async thunk? We can essentially extend the narrow thunk definition to include it receiving a callback.

Consider:

```

function foo(x,y,cb) {
    setTimeout( function(){
        cb( x + y );
    }, 1000 );
}

function fooThunk(cb) {
    foo( 3, 4, cb );
}

// later

```

```
fooThunk( function(sum){
    console.log( sum ); // 7
} );
```

As you can see, `fooThunk(..)` only expects a `cb(..)` parameter, as it already has values 3 and 4 (for `x` and `y`, respectively) pre-specified and ready to pass to `foo(..)`. A thunk is just waiting around patiently for the last piece it needs to do its job: the callback.

You don't want to make thunks manually, though. So, let's invent a utility that does this wrapping for us.

Consider:

```

function thunkify(fn) {
    var args = [].slice.call( arguments, 1 );
    return function(cb) {
        args.push( cb );
        return fn.apply( null, args );
    };
}

```

```

var fooThunk = thunkify( foo, 3, 4 );

// later

fooThunk( function(sum) {
  console.log( sum );      // 7
} );

```

Tip: Here we assume that the original (`foo(..)`) function signature expects its callback in the last position, with any other parameters coming before it. This is a pretty ubiquitous “standard” for async JS function standards. You might call it “callback-last style.” If for some reason you had a need to handle “callback-first style” signatures, you would just make a utility that used `args.unshift(..)` instead of `args.push(..)`.

The preceding formulation of `thunkify(..)` takes both the `foo(..)` function reference, and any parameters it needs, and returns back the thunk itself (`fooThunk(..)`). However, that’s not the typical approach you’ll find to thunks in JS.

Instead of `thunkify(..)` making the thunk itself, typically – if not perplexingly – the `thunkify(..)` utility would produce a function that produces thunks.

Uhhhh... yeah.

Consider:

```

function thunkify(fn) {
  return function() {
    var args = [].slice.call( arguments );
    return function(cb) {
      args.push( cb );
      return fn.apply( null, args );
    };
  };
}

```

The main difference here is the extra `return function() { .. }` layer. Here’s how its usage differs:

```

var whatIsThis = thunkify( foo );

var fooThunk = whatIsThis( 3, 4 );

// later

fooThunk( function(sum) {
  console.log( sum );      // 7
} );

```

Obviously, the big question this snippet implies is what is `whatIsThis` properly called? It's not the thunk, it's the thing that will produce thunks from `foo(..)` calls. It's kind of like a "factory" for "thunks." There doesn't seem to be any kind of standard agreement for naming such a thing.

So, my proposal is "thunkory" ("thunk" + "factory"). So, `thunkify(..)` produces a thunkory, and a thunkory produces thunks. That reasoning is symmetric to my proposal for "promisory" in Chapter 3:

```
var fooThunkory = thunkify( foo );

var fooThunk1 = fooThunkory( 3, 4 );
var fooThunk2 = fooThunkory( 5, 6 );

// later

fooThunk1( function(sum) {
  console.log( sum );      // 7
} );

fooThunk2( function(sum) {
  console.log( sum );      // 11
} );
```

Note: The running `foo(..)` example expects a style of callback that's not "error-first style." Of course, "error-first style" is much more common. If `foo(..)` had some sort of legitimate error-producing expectation, we could change it to expect and use an error-first callback. None of the subsequent `thunkify(..)` machinery cares what style of callback is assumed. The only difference in usage would be `fooThunk1(function(err,sum){...})`.

Exposing the thunkory method – instead of how the earlier `thunkify(..)` hides this intermediary step – may seem like unnecessary complication. But in general, it's quite useful to make thunkories at the beginning of your program to wrap existing API methods, and then be able to pass around and call those thunkories when you need thunks. The two distinct steps preserve a cleaner separation of capability.

To illustrate:

```
// cleaner:
var fooThunkory = thunkify( foo );

var fooThunk1 = fooThunkory( 3, 4 );
var fooThunk2 = fooThunkory( 5, 6 );

// instead of:
var fooThunk1 = thunkify( foo, 3, 4 );
var fooThunk2 = thunkify( foo, 5, 6 );
```

Regardless of whether you like to deal with the thunkories explicitly or not, the usage of thunks `fooThunk1(..)` and `fooThunk2(..)` remains the same.

s/promise/thunk/

So what's all this thunk stuff have to do with generators?

Comparing thunks to promises generally: they're not directly interchangable as they're not equivalent in behavior. Promises are vastly more capable and trustable than bare thunks.

But in another sense, they both can be seen as a request for a value, which may be async in its answering.

Recall from Chapter 3 we defined a utility for promisifying a function, which we called `Promise.wrap(..)` – we could have called it `promisify(..)`, too! This Promise-wrapping utility doesn't produce Promises; it produces promisories that in turn produce Promises. This is completely symmetric to the thunkories and thunks presently being discussed.

To illustrate the symmetry, let's first alter the running `foo(..)` example from earlier to assume an “error-first style” callback:

```
function foo(x,y,cb) {
  setTimeout( function(){
    // assume `cb(..)` as "error-first style"
    cb( null, x + y );
  }, 1000 );
}
```

Now, we'll compare using `thunkify(..)` and `promisify(..)` (aka `Promise.wrap(..)` from Chapter 3):

```
// symmetrical: constructing the question asker
var fooThunkory = thunkify( foo );
var fooPromisory = promisify( foo );

// symmetrical: asking the question
var fooThunk = fooThunkory( 3, 4 );
var fooPromise = fooPromisory( 3, 4 );

// get the thunk answer
fooThunk( function(err,sum){
  if (err) {
    console.error( err );
  }
  else {
    console.log( sum );      // 7
  }
})
```

```

} );

// get the promise answer
fooPromise
.then(
  function(sum){
    console.log( sum );      // 7
  },
  function(err){
    console.error( err );
  }
);

```

Both the thunkory and the promisory are essentially asking a question (for a value), and respectively the thunk `fooThunk` and promise `fooPromise` represent the future answers to that question. Presented in that light, the symmetry is clear.

With that perspective in mind, we can see that generators which `yield` Promises for asynchrony could instead `yield` thunks for asynchrony. All we'd need is a smarter `run(..)` utility (like from before) that can not only look for and wire up to a `yielded` Promise but also to provide a callback to a `yielded` thunk.

Consider:

```

function *foo() {
  var val = yield request( "http://some.url.1" );
  console.log( val );
}

run( foo );

```

In this example, `request(..)` could either be a promisory that returns a promise, or a thunkory that returns a thunk. From the perspective of what's going on inside the generator code logic, we don't care about that implementation detail, which is quite powerful!

So, `request(..)` could be either:

```

// promisory `request(..)` (see Chapter 3)
var request = Promise.wrap( ajax );

// vs.

// thunkory `request(..)`
var request = thunkify( ajax );

```

Finally, as a thunk-aware patch to our earlier `run(..)` utility, we would need logic like this:

```

// ..
// did we receive a thunk back?
else if (typeof next.value == "function") {
  return new Promise( function(resolve,reject){
    // call the thunk with an error-first callback
    next.value( function(err,msg) {
      if (err) {
        reject( err );
      }
      else {
        resolve( msg );
      }
    } );
  } )
  .then( handleNext,
    function handleErr(err) {
      return Promise.resolve(
        it.throw( err )
      )
      .then( handleResult );
    }
  );
}

```

Now, our generators can either call promises to `yield` Promises, or call thunkies to `yield` thunks, and in either case, `run(..)` would handle that value and use it to wait for the completion to resume the generator.

Symmetry wise, these two approaches look identical. However, we should point out that's true only from the perspective of Promises or thunks representing the future value continuation of a generator.

From the larger perspective, thunks do not in and of themselves have hardly any of the trustability or compositability guarantees that Promises are designed with. Using a thunk as a stand-in for a Promise in this particular generator asynchrony pattern is workable but should be seen as less than ideal when compared to all the benefits that Promises offer (see Chapter 3).

If you have the option, prefer `yield pr` rather than `yield th`. But there's nothing wrong with having a `run(..)` utility which can handle both value types.

Note: The `runner(..)` utility in my *asynquence* library, which will be discussed in Appendix A, handles `yields` of Promises, thunks and *asynquence* sequences.

Pre-ES6 Generators

You're hopefully convinced now that generators are a very important addition to the async programming toolbox. But it's a new syntax in ES6, which means you can't just polyfill generators like you can Promises (which are just a new API). So what can we do to bring generators to our browser JS if we don't have the luxury of ignoring pre-ES6 browsers?

For all new syntax extensions in ES6, there are tools – the most common term for them is transpilers, for trans-compilers – which can take your ES6 syntax and transform it into equivalent (but obviously uglier!) pre-ES6 code. So, generators can be transpiled into code that will have the same behavior but work in ES5 and below.

But how? The “magic” of `yield` doesn't obviously sound like code that's easy to transpile. We actually hinted at a solution in our earlier discussion of closure-based *iterators*.

Manual Transformation

Before we discuss the transpilers, let's derive how manual transpilation would work in the case of generators. This isn't just an academic exercise, because doing so will actually help further reinforce how they work.

Consider:

```
// `request(..)` is a Promise-aware Ajax utility

function *foo(url) {
  try {
    console.log( "requesting:", url );
    var val = yield request( url );
    console.log( val );
  }
  catch (err) {
    console.log( "Oops:", err );
    return false;
  }
}

var it = foo( "http://some.url.1" );
```

The first thing to observe is that we'll still need a normal `foo()` function that can be called, and it will still need to return an *iterator*. So, let's sketch out the non-generator transformation:

```
function foo(url) {
  // ..
```

```

// make and return an iterator
return {
  next: function(v) {
    // ..
  },
  throw: function(e) {
    // ..
  }
};

var it = foo( "http://some.url.1" );

```

The next thing to observe is that a generator does its “magic” by suspending its scope/state, but we can emulate that with function closure (see the *Scope & Closures* title of this series). To understand how to write such code, we’ll first annotate different parts of our generator with state values:

```

// `request(..)` is a Promise-aware Ajax utility

function *foo(url) {
  // STATE *1*

  try {
    console.log( "requesting:", url );
    var TMP1 = request( url );

    // STATE *2*
    var val = yield TMP1;
    console.log( val );
  }
  catch (err) {
    // STATE *3*
    console.log( "Oops:", err );
    return false;
  }
}

```

Note: For more accurate illustration, we split up the `val = yield request..` statement into two parts, using the temporary `TMP1` variable. `request(..)` happens in state ***1***, and the assignment of its completion value to `val` happens in state ***2***. We’ll get rid of that intermediate `TMP1` when we convert the code to its non-generator equivalent.

In other words, ***1*** is the beginning state, ***2*** is the state if the `request(..)` succeeds, and ***3*** is the state if the `request(..)` fails. You can probably imagine how any extra `yield` steps would just be encoded as extra states.

Back to our transpiled generator, let's define a variable `state` in the closure we can use to keep track of the state:

```
function foo(url) {
  // manage generator state
  var state;

  //
}
```

Now, let's define an inner function called `process(..)` inside the closure which handles each state, using a `switch` statement:

```
// `request(..)` is a Promise-aware Ajax utility

function foo(url) {
  // manage generator state
  var state;

  // generator-wide variable declarations
  var val;

  function process(v) {
    switch (state) {
      case 1:
        console.log( "requesting:", url );
        return request( url );
      case 2:
        val = v;
        console.log( val );
        return;
      case 3:
        var err = v;
        console.log( "Oops:", err );
        return false;
    }
  }
}

// ..
```

Each state in our generator is represented by its own `case` in the `switch` statement. `process(..)` will be called each time we need to process a new state. We'll come back to how that works in just a moment.

For any generator-wide variable declarations (`val`), we move those to a `var` declaration outside of `process(..)` so they can survive multiple calls to `process(..)`. But the “block scoped” `err` variable is only needed for the *3*

state, so we leave it in place.

In state ***1***, instead of `yield request(..)`, we did `return request(..)`. In terminal state ***2***, there was no explicit `return`, so we just do a `return`; which is the same as `return undefined`. In terminal state ***3***, there was a `return false`, so we preserve that.

Now we need to define the code in the *iterator* functions so they call `process(..)` appropriately:

```
function foo(url) {
  // manage generator state
  var state;

  // generator-wide variable declarations
  var val;

  function process(v) {
    switch (state) {
      case 1:
        console.log( "requesting:", url );
        return request( url );
      case 2:
        val = v;
        console.log( val );
        return;
      case 3:
        var err = v;
        console.log( "Oops:", err );
        return false;
    }
  }

  // make and return an iterator
  return {
    next: function(v) {
      // initial state
      if (!state) {
        state = 1;
        return {
          done: false,
          value: process()
        };
      }
      // yield resumed successfully
      else if (state == 1) {
        state = 2;
      }
    }
  };
}
```

```

        return {
          done: true,
          value: process( v )
        };
      }
      // generator already completed
    else {
      return {
        done: true,
        value: undefined
      };
    }
  },
  "throw": function(e) {
    // the only explicit error handling is in
    // state *1*
    if (state == 1) {
      state = 3;
      return {
        done: true,
        value: process( e )
      };
    }
    // otherwise, an error won't be handled,
    // so just throw it right back out
    else {
      throw e;
    }
  }
};
}

```

How does this code work?

1. The first call to the *iterator*'s `next()` call would move the generator from the uninitialized state to state 1, and then call `process()` to handle that state. The return value from `request(..)`, which is the promise for the Ajax response, is returned back as the `value` property from the `next()` call.
2. If the Ajax request succeeds, the second call to `next(..)` should send in the Ajax response value, which moves our state to 2. `process(..)` is again called (this time with the passed in Ajax response value), and the `value` property returned from `next(..)` will be `undefined`.
3. However, if the Ajax request fails, `throw(..)` should be called with the error, which would move the state from 1 to 3 (instead of 2). Again `process(..)` is called, this time with the error value. That `case` returns `false`, which is set as the `value` property returned from the `throw(..)`

call.

From the outside – that is, interacting only with the *iterator* – this `foo(..)` normal function works pretty much the same as the `*foo(..)` generator would have worked. So we've effectively “transpiled” our ES6 generator to pre-ES6 compatibility!

We could then manually instantiate our generator and control its iterator – calling `var it = foo(..)` and `it.next(..)` and such – or better, we could pass it to our previously defined `run(..)` utility as `run(foo, ..)`.

Automatic Transpilation

The preceding exercise of manually deriving a transformation of our ES6 generator to pre-ES6 equivalent teaches us how generators work conceptually. But that transformation was really intricate and very non-portable to other generators in our code. It would be quite impractical to do this work by hand, and would completely obviate all the benefit of generators.

But luckily, several tools already exist that can automatically convert ES6 generators to things like what we derived in the previous section. Not only do they do the heavy lifting work for us, but they also handle several complications that we glossed over.

One such tool is regenerator (<https://facebook.github.io/regenerator/>), from the smart folks at Facebook.

If we use regenerator to transpile our previous generator, here's the code produced (at the time of this writing):

```
// `request(..)` is a Promise-aware Ajax utility

var foo = regeneratorRuntime.mark(function foo(url) {
  var val;

  return regeneratorRuntime.wrap(function foo$(context$1$0) {
    while (1) switch (context$1$0.prev = context$1$0.next) {
      case 0:
        context$1$0.prev = 0;
        console.log( "requesting:", url );
        context$1$0.next = 4;
        return request( url );
      case 4:
        val = context$1$0.sent;
        console.log( val );
        context$1$0.next = 12;
        break;
      case 8:
        context$1$0.prev = 8;
    }
  });
})
```

```

        context$1$0.t0 = context$1$0.catch(0);
        console.log("Oops:", context$1$0.t0);
        return context$1$0.abrupt("return", false);
    case 12:
    case "end":
        return context$1$0.stop();
    }
}, foo, this, [[0, 8]]);
);

```

There's some obvious similarities here to our manual derivation, such as the `switch` / `case` statements, and we even see `val` pulled out of the closure just as we did.

Of course, one trade-off is that regenerator's transpilation requires a helper library `regeneratorRuntime` that holds all the reusable logic for managing a general generator / *iterator*. A lot of that boilerplate looks different than our version, but even then, the concepts can be seen, like with `context$1$0.next = 4` keeping track of the next state for the generator.

The main takeaway is that generators are not restricted to only being useful in ES6+ environments. Once you understand the concepts, you can employ them throughout your code, and use tools to transform the code to be compatible with older environments.

This is more work than just using a `Promise` API polyfill for pre-ES6 Promises, but the effort is totally worth it, because generators are so much better at expressing async flow control in a reasonable, sensible, synchronous-looking, sequential fashion.

Once you get hooked on generators, you'll never want to go back to the hell of async spaghetti callbacks!

Review

Generators are a new ES6 function type that does not run-to-completion like normal functions. Instead, the generator can be paused in mid-completion (entirely preserving its state), and it can later be resumed from where it left off.

This pause/resume interchange is cooperative rather than preemptive, which means that the generator has the sole capability to pause itself, using the `yield` keyword, and yet the *iterator* that controls the generator has the sole capability (via `next(..)`) to resume the generator.

The `yield` / `next(..)` duality is not just a control mechanism, it's actually a two-way message passing mechanism. A `yield ..` expression essentially pauses waiting for a value, and the next `next(..)` call passes a value (or implicit `undefined`) back to that paused `yield` expression.

The key benefit of generators related to async flow control is that the code inside a generator expresses a sequence of steps for the task in a naturally sync/sequential fashion. The trick is that we essentially hide potential asynchrony behind the `yield` keyword – moving the asynchrony to the code where the generator's *iterator* is controlled.

In other words, generators preserve a sequential, synchronous, blocking code pattern for async code, which lets our brains reason about the code much more naturally, addressing one of the two key drawbacks of callback-based async.

You Don't Know JS: Async & Performance

Chapter 5: Program Performance

This book so far has been all about how to leverage asynchrony patterns more effectively. But we haven't directly addressed why asynchrony really matters to JS. The most obvious explicit reason is **performance**.

For example, if you have two Ajax requests to make, and they're independent, but you need to wait on them both to finish before doing the next task, you have two options for modeling that interaction: serial and concurrent.

You could make the first request and wait to start the second request until the first finishes. Or, as we've seen both with promises and generators, you could make both requests "in parallel," and express the "gate" to wait on both of them before moving on.

Clearly, the latter is usually going to be more performant than the former. And better performance generally leads to better user experience.

It's even possible that asynchrony (interleaved concurrency) can improve just the perception of performance, even if the overall program still takes the same amount of time to complete. User perception of performance is every bit – if not more! – as important as actual measurable performance.

We want to now move beyond localized asynchrony patterns to talk about some bigger picture performance details at the program level.

Note: You may be wondering about micro-performance issues like if `a++` or `++a` is faster. We'll look at those sorts of performance details in the next chapter on "Benchmarking & Tuning."

Web Workers

If you have processing-intensive tasks but you don't want them to run on the main thread (which may slow down the browser/UI), you might have wished that JavaScript could operate in a multithreaded manner.

In Chapter 1, we talked in detail about how JavaScript is single threaded. And that's still true. But a single thread isn't the only way to organize the execution of your program.

Imagine splitting your program into two pieces, and running one of those pieces on the main UI thread, and running the other piece on an entirely separate thread.

What kinds of concerns would such an architecture bring up?

For one, you'd want to know if running on a separate thread meant that it ran in parallel (on systems with multiple CPUs/cores) such that a long-running process

on that second thread would **not** block the main program thread. Otherwise, “virtual threading” wouldn’t be of much benefit over what we already have in JS with `async` concurrency.

And you’d want to know if these two pieces of the program have access to the same shared scope/resources. If they do, then you have all the questions that multithreaded languages (Java, C++, etc.) deal with, such as needing cooperative or preemptive locking (mutexes, etc.). That’s a lot of extra work, and shouldn’t be undertaken lightly.

Alternatively, you’d want to know how these two pieces could “communicate” if they couldn’t share scope/resources.

All these are great questions to consider as we explore a feature added to the web platform circa HTML5 called “Web Workers.” This is a feature of the browser (aka host environment) and actually has almost nothing to do with the JS language itself. That is, JavaScript does not *currently* have any features that support threaded execution.

But an environment like your browser can easily provide multiple instances of the JavaScript engine, each on its own thread, and let you run a different program in each thread. Each of those separate threaded pieces of your program is called a “(Web) Worker.” This type of parallelism is called “task parallelism,” as the emphasis is on splitting up chunks of your program to run in parallel.

From your main JS program (or another Worker), you instantiate a Worker like so:

```
var w1 = new Worker( "http://some.url.1/mycoolworker.js" );
```

The URL should point to the location of a JS file (not an HTML page!) which is intended to be loaded into a Worker. The browser will then spin up a separate thread and let that file run as an independent program in that thread.

Note: The kind of Worker created with such a URL is called a “Dedicated Worker.” But instead of providing a URL to an external file, you can also create an “Inline Worker” by providing a Blob URL (another HTML5 feature); essentially it’s an inline file stored in a single (binary) value. However, Blobs are beyond the scope of what we’ll discuss here.

Workers do not share any scope or resources with each other or the main program – that would bring all the nightmares of threaded programming to the forefront – but instead have a basic event messaging mechanism connecting them.

The `w1` Worker object is an event listener and trigger, which lets you subscribe to events sent by the Worker as well as send events to the Worker.

Here’s how to listen for events (actually, the fixed “`message`” event):

```
w1.addEventListener( "message", function(evt){
  // evt.data
});
```

And you can send the "message" event to the Worker:

```
w1.postMessage( "something cool to say" );
```

Inside the Worker, the messaging is totally symmetrical:

```
// "mycoolworker.js"

addEventListener( "message", function(evt){
    // evt.data
} );

postMessage( "a really cool reply" );
```

Notice that a dedicated Worker is in a one-to-one relationship with the program that created it. That is, the "message" event doesn't need any disambiguation here, because we're sure that it could only have come from this one-to-one relationship – either it came from the Worker or the main page.

Usually the main page application creates the Workers, but a Worker can instantiate its own child Worker(s) – known as subworkers – as necessary. Sometimes this is useful to delegate such details to a sort of "master" Worker that spawns other Workers to process parts of a task. Unfortunately, at the time of this writing, Chrome still does not support subworkers, while Firefox does.

To kill a Worker immediately from the program that created it, call `terminate()` on the Worker object (like `w1` in the previous snippets). Abruptly terminating a Worker thread does not give it any chance to finish up its work or clean up any resources. It's akin to you closing a browser tab to kill a page.

If you have two or more pages (or multiple tabs with the same page!) in the browser that try to create a Worker from the same file URL, those will actually end up as completely separate Workers. Shortly, we'll discuss a way to "share" a Worker.

Note: It may seem like a malicious or ignorant JS program could easily perform a denial-of-service attack on a system by spawning hundreds of Workers, seemingly each with their own thread. While it's true that it's somewhat of a guarantee that a Worker will end up on a separate thread, this guarantee is not unlimited. The system is free to decide how many actual threads/CPUs/cores it really wants to create. There's no way to predict or guarantee how many you'll have access to, though many people assume it's at least as many as the number of CPUs/cores available. I think the safest assumption is that there's at least one other thread besides the main UI thread, but that's about it.

Worker Environment

Inside the Worker, you do not have access to any of the main program's resources. That means you cannot access any of its global variables, nor can you access the page's DOM or other resources. Remember: it's a totally separate thread.

You can, however, perform network operations (Ajax, WebSockets) and set timers. Also, the Worker has access to its own copy of several important global variables/features, including `navigator`, `location`, `JSON`, and `applicationCache`.

You can also load extra JS scripts into your Worker, using `importScripts(..)`:

```
// inside the Worker
importScripts( "foo.js", "bar.js" );
```

These scripts are loaded synchronously, which means the `importScripts(..)` call will block the rest of the Worker's execution until the file(s) are finished loading and executing.

Note: There have also been some discussions about exposing the `<canvas>` API to Workers, which combined with having canvases be Transferables (see the "Data Transfer" section), would allow Workers to perform more sophisticated off-thread graphics processing, which can be useful for high-performance gaming (WebGL) and other similar applications. Although this doesn't exist yet in any browsers, it's likely to happen in the near future.

What are some common uses for Web Workers?

- Processing intensive math calculations
- Sorting large data sets
- Data operations (compression, audio analysis, image pixel manipulations, etc.)
- High-traffic network communications

Data Transfer

You may notice a common characteristic of most of those uses, which is that they require a large amount of information to be transferred across the barrier between threads using the event mechanism, perhaps in both directions.

In the early days of Workers, serializing all data to a string value was the only option. In addition to the speed penalty of the two-way serializations, the other major negative was that the data was being copied, which meant a doubling of memory usage (and the subsequent churn of garbage collection).

Thankfully, we now have a few better options.

If you pass an object, a so-called "Structured Cloning Algorithm" (https://developer.mozilla.org/en-US/docs/Web/Guide/API/DOM/The_structured_clone_algorithm) is used to copy/duplicate the object on the other side. This algorithm is fairly sophisticated and can even handle duplicating objects with circular references. The to-string/from-string performance penalty is not paid, but we still have duplication of memory using this approach. There is support for this in IE10 and above, as well as all the other major browsers.

An even better option, especially for larger data sets, is “Transferable Objects” (<http://updates.html5rocks.com/2011/12/Transferable-Objects-Lightning-Fast>). What happens is that the object’s “ownership” is transferred, but the data itself is not moved. Once you transfer away an object to a Worker, it’s empty or inaccessible in the originating location – that eliminates the hazards of threaded programming over a shared scope. Of course, transfer of ownership can go in both directions.

There really isn’t much you need to do to opt into a Transferable Object; any data structure that implements the Transferable interface (<https://developer.mozilla.org/en-US/docs/Web/API/Transferable>) will automatically be transferred this way (support Firefox & Chrome).

For example, typed arrays like `Uint8Array` (see the *ES6 & Beyond* title of this series) are “Transferables.” This is how you’d send a Transferable Object using `postMessage(..)`:

```
// `foo` is a `Uint8Array` for instance
postMessage( foo.buffer, [ foo.buffer ] );
```

The first parameter is the raw buffer and the second parameter is a list of what to transfer.

Browsers that don’t support Transferable Objects simply degrade to structured cloning, which means performance reduction rather than outright feature breakage.

Shared Workers

If your site or app allows for loading multiple tabs of the same page (a common feature), you may very well want to reduce the resource usage of their system by preventing duplicate dedicated Workers; the most common limited resource in this respect is a socket network connection, as browsers limit the number of simultaneous connections to a single host. Of course, limiting multiple connections from a client also eases your server resource requirements.

In this case, creating a single centralized Worker that all the page instances of your site or app can *share* is quite useful.

That’s called a `SharedWorker`, which you create like so (support for this is limited to Firefox and Chrome):

```
var w1 = new SharedWorker( "http://some.url.1/mycoolworker.js" );
```

Because a shared Worker can be connected to or from more than one program instance or page on your site, the Worker needs a way to know which program a message comes from. This unique identification is called a “port” – think network socket ports. So the calling program must use the `port` object of the Worker for communication:

```
w1.port.addEventListener( "message", handleMessages );

// ...

w1.port.postMessage( "something cool" );
```

Also, the port connection must be initialized, as:

```
w1.port.start();
```

Inside the shared Worker, an extra event must be handled: "connect". This event provides the port object for that particular connection. The most convenient way to keep multiple connections separate is to use closure (see *Scope & Closures* title of this series) over the port, as shown next, with the event listening and transmitting for that connection defined inside the handler for the "connect" event:

```
// inside the shared Worker
addEventListener( "connect", function(evt){
    // the assigned port for this connection
    var port = evt.ports[0];

    port.addEventListener( "message", function(evt){
        // ...

        port.postMessage( ... );

        // ...
    } );

    // initialize the port connection
    port.start();
} );
```

Other than that difference, shared and dedicated Workers have the same capabilities and semantics.

Note: Shared Workers survive the termination of a port connection if other port connections are still alive, whereas dedicated Workers are terminated whenever the connection to their initiating program is terminated.

Polyfilling Web Workers

Web Workers are very attractive performance-wise for running JS programs in parallel. However, you may be in a position where your code needs to run in older browsers that lack support. Because Workers are an API and not a syntax, they can be polyfilled, to an extent.

If a browser doesn't support Workers, there's simply no way to fake multi-

threading from the performance perspective. Iframes are commonly thought of to provide a parallel environment, but in all modern browsers they actually run on the same thread as the main page, so they're not sufficient for faking parallelism.

As we detailed in Chapter 1, JS's asynchronicity (not parallelism) comes from the event loop queue, so you can force faked Workers to be asynchronous using timers (`setTimeout(..)`, etc.). Then you just need to provide a polyfill for the Worker API. There are some listed here (<https://github.com/Modernizr/Modernizr/wiki/HTML5-Cross-Browser-Polyfills#web-workers>), but frankly none of them look great.

I've written a sketch of a polyfill for `Worker` here (<https://gist.github.com/getify/1b26accb1a09aa53ad25>). It's basic, but it should get the job done for simple `Worker` support, given that the two-way messaging works correctly as well as "`onerror`" handling. You could probably also extend it with more features, such as `terminate()` or faked Shared Workers, as you see fit.

Note: You can't fake synchronous blocking, so this polyfill just disallows use of `importScripts(..)`. Another option might have been to parse and transform the Worker's code (once Ajax loaded) to handle rewriting to some asynchronous form of an `importScripts(..)` polyfill, perhaps with a promise-aware interface.

SIMD

Single instruction, multiple data (SIMD) is a form of "data parallelism," as contrasted to "task parallelism" with Web Workers, because the emphasis is not really on program logic chunks being parallelized, but rather multiple bits of data being processed in parallel.

With SIMD, threads don't provide the parallelism. Instead, modern CPUs provide SIMD capability with "vectors" of numbers – think: type specialized arrays – as well as instructions that can operate in parallel across all the numbers; these are low-level operations leveraging instruction-level parallelism.

The effort to expose SIMD capability to JavaScript is primarily spearheaded by Intel (<https://01.org/node/1495>), namely by Mohammad Haghagh (at the time of this writing), in cooperation with Firefox and Chrome teams. SIMD is on an early standards track with a good chance of making it into a future revision of JavaScript, likely in the ES7 timeframe.

SIMD JavaScript proposes to expose short vector types and APIs to JS code, which on those SIMD-enabled systems would map the operations directly through to the CPU equivalents, with fallback to non-parallelized operation "shims" on non-SIMD systems.

The performance benefits for data-intensive applications (signal analysis, matrix operations on graphics, etc.) with such parallel math processing are quite obvious!

Early proposal forms of the SIMD API at the time of this writing look like this:

```
var v1 = SIMD.float32x4( 3.14159, 21.0, 32.3, 55.55 );
var v2 = SIMD.float32x4( 2.1, 3.2, 4.3, 5.4 );

var v3 = SIMD.int32x4( 10, 101, 1001, 10001 );
var v4 = SIMD.int32x4( 10, 20, 30, 40 );

SIMD.float32x4.mul( v1, v2 );    // [ 6.597339, 67.2, 138.89, 299.97 ]
SIMD.int32x4.add( v3, v4 );    // [ 20, 121, 1031, 10041 ]
```

Shown here are two different vector data types, 32-bit floating-point numbers and 32-bit integer numbers. You can see that these vectors are sized exactly to four 32-bit elements, as this matches the SIMD vector sizes (128-bit) available in most modern CPUs. It's also possible we may see an x8 (or larger!) version of these APIs in the future.

Besides `mul()` and `add()`, many other operations are likely to be included, such as `sub()`, `div()`, `abs()`, `neg()`, `sqrt()`, `reciprocal()`, `reciprocalSqrt()` (arithmetic), `shuffle()` (rearrange vector elements), `and()`, `or()`, `xor()`, `not()` (logical), `equal()`, `greaterThan()`, `lessThan()` (comparison), `shiftLeft()`, `shiftRightLogical()`, `shiftRightArithmetic()` (shifts), `fromFloat32x4()`, and `fromInt32x4()` (conversions).

Note: There's an official "prolyfill" (hopeful, expectant, future-leaning polyfill) for the SIMD functionality available (https://github.com/johnmccutchan/ecmascript_simd), which illustrates a lot more of the planned SIMD capability than we've illustrated in this section.

asm.js

“asm.js” (<http://asmjs.org/>) is a label for a highly optimizable subset of the JavaScript language. By carefully avoiding certain mechanisms and patterns that are *hard* to optimize (garbage collection, coercion, etc.), asm.js-styled code can be recognized by the JS engine and given special attention with aggressive low-level optimizations.

Distinct from other program performance mechanisms discussed in this chapter, asm.js isn't necessarily something that needs to be adopted into the JS language specification. There *is* an asm.js specification (<http://asmjs.org/spec/latest/>), but it's mostly for tracking an agreed upon set of candidate inferences for optimization rather than a set of requirements of JS engines.

There's not currently any new syntax being proposed. Instead, asm.js suggests ways to recognize existing standard JS syntax that conforms to the rules of asm.js and let engines implement their own optimizations accordingly.

There's been some disagreement between browser vendors over exactly how asm.js should be activated in a program. Early versions of the asm.js exper-

iment required a "use asm"; pragma (similar to strict mode's "use strict";) to help clue the JS engine to be looking for asm.js optimization opportunities and hints. Others have asserted that asm.js should just be a set of heuristics that engines automatically recognize without the author having to do anything extra, meaning that existing programs could theoretically benefit from asm.js-style optimizations without doing anything special.

How to Optimize with asm.js

The first thing to understand about asm.js optimizations is around types and coercion (see the *Types & Grammar* title of this series). If the JS engine has to track multiple different types of values in a variable through various operations, so that it can handle coercions between types as necessary, that's a lot of extra work that keeps the program optimization suboptimal.

Note: We're going to use asm.js-style code here for illustration purposes, but be aware that it's not commonly expected that you'll author such code by hand. asm.js is more intended to a compilation target from other tools, such as Emscripten (<https://github.com/kripken/emscripten/wiki>). It's of course possible to write your own asm.js code, but that's usually a bad idea because the code is very low level and managing it can be very time consuming and error prone. Nevertheless, there may be cases where you'd want to hand tweak your code for asm.js optimization purposes.

There are some “tricks” you can use to hint to an asm.js-aware JS engine what the intended type is for variables/operations, so that it can skip these coercion tracking steps.

For example:

```
var a = 42;  
// ..
```

```
var b = a;
```

In that program, the `b = a` assignment leaves the door open for type divergence in variables. However, it could instead be written as:

```
var a = 42;  
// ..
```

```
var b = a | 0;
```

Here, we've used the `|` (“binary OR”) with value `0`, which has no effect on the value other than to make sure it's a 32-bit integer. That code run in a normal JS engine works just fine, but when run in an asm.js-aware JS engine it *can* signal

that `b` should always be treated as a 32-bit integer, so the coercion tracking can be skipped.

Similarly, the addition operation between two variables can be restricted to a more performant integer addition (instead of floating point):

```
(a + b) | 0
```

Again, the asm.js-aware JS engine can see that hint and infer that the `+` operation should be 32-bit integer addition because the end result of the whole expression would automatically be 32-bit integer conformed anyway.

asm.js Modules

One of the biggest detractors to performance in JS is around memory allocation, garbage collection, and scope access. asm.js suggests one of the ways around these issues is to declare a more formalized asm.js “module” – do not confuse these with ES6 modules; see the *ES6 & Beyond* title of this series.

For an asm.js module, you need to explicitly pass in a tightly conformed namespace – this is referred to in the spec as `stdlib`, as it should represent standard libraries needed – to import necessary symbols, rather than just using globals via lexical scope. In the base case, the `window` object is an acceptable `stdlib` object for asm.js module purposes, but you could and perhaps should construct an even more restricted one.

You also must declare a “heap” – which is just a fancy term for a reserved spot in memory where variables can already be used without asking for more memory or releasing previously used memory – and pass that in, so that the asm.js module won’t need to do anything that would cause memory churn; it can just use the pre-reserved space.

A “heap” is likely a typed `ArrayBuffer`, such as:

```
var heap = new ArrayBuffer( 0x10000 ); // 64k heap
```

Using that pre-reserved 64k of binary space, an asm.js module can store and retrieve values in that buffer without any memory allocation or garbage collection penalties. For example, the `heap` buffer could be used inside the module to back an array of 64-bit float values like this:

```
var arr = new Float64Array( heap );
```

OK, so let’s make a quick, silly example of an asm.js-styled module to illustrate how these pieces fit together. We’ll define a `foo(..)` that takes a start (`x`) and end (`y`) integer for a range, and calculates all the inner adjacent multiplications of the values in the range, and then finally averages those values together:

```
function fooASM(stdlib,foreign,heap) {
  "use asm";
```

```

var arr = new stdlib.Int32Array( heap );

function foo(x,y) {
    x = x | 0;
    y = y | 0;

    var i = 0;
    var p = 0;
    var sum = 0;
    var count = ((y|0) - (x|0)) | 0;

    // calculate all the inner adjacent multiplications
    for (i = x | 0;
        (i | 0) < (y | 0);
        p = (p + 8) | 0, i = (i + 1) | 0
    ) {
        // store result
        arr[ p >> 3 ] = (i * (i + 1)) | 0;
    }

    // calculate average of all intermediate values
    for (i = 0, p = 0;
        (i | 0) < (count | 0);
        p = (p + 8) | 0, i = (i + 1) | 0
    ) {
        sum = (sum + arr[ p >> 3 ]) | 0;
    }

    return +(sum / count);
}

return {
    foo: foo
};
}

var heap = new ArrayBuffer( 0x1000 );
var foo = fooASM( window, null, heap ).foo;

foo( 10, 20 );           // 233

```

Note: This asm.js example is hand authored for illustration purposes, so it doesn't represent the same code that would be produced from a compilation tool targeting asm.js. But it does show the typical nature of asm.js code, especially the type hinting and use of the `heap` buffer for temporary variable storage.

The first call to `fooASM(...)` is what sets up our asm.js module with its `heap`

allocation. The result is a `foo(..)` function we can call as many times as necessary. Those `foo(..)` calls should be specially optimized by an `asm.js`-aware JS engine. Importantly, the preceding code is completely standard JS and would run just fine (without special optimization) in a non-`asm.js` engine.

Obviously, the nature of restrictions that make `asm.js` code so optimizable reduces the possible uses for such code significantly. `asm.js` won't necessarily be a general optimization set for any given JS program. Instead, it's intended to provide an optimized way of handling specialized tasks such as intensive math operations (e.g., those used in graphics processing for games).

Review

The first four chapters of this book are based on the premise that `async` coding patterns give you the ability to write more performant code, which is generally a very important improvement. But `async` behavior only gets you so far, because it's still fundamentally bound to a single event loop thread.

So in this chapter we've covered several program-level mechanisms for improving performance even further.

`Web Workers` let you run a JS file (aka program) in a separate thread using `async` events to message between the threads. They're wonderful for offloading long-running or resource-intensive tasks to a different thread, leaving the main UI thread more responsive.

`SIMD` proposes to map CPU-level parallel math operations to JavaScript APIs for high-performance data-parallel operations, like number processing on large data sets.

Finally, `asm.js` describes a small subset of JavaScript that avoids the hard-to-optimize parts of JS (like garbage collection and coercion) and lets the JS engine recognize and run such code through aggressive optimizations. `asm.js` could be hand authored, but that's extremely tedious and error prone, akin to hand authoring assembly language (hence the name). Instead, the main intent is that `asm.js` would be a good target for cross-compilation from other highly optimized program languages – for example, `Emscripten` (<https://github.com/kripken/emscripten/wiki>) transpiling C/C++ to JavaScript.

While not covered explicitly in this chapter, there are even more radical ideas under very early discussion for JavaScript, including approximations of direct threaded functionality (not just hidden behind data structure APIs). Whether that happens explicitly, or we just see more parallelism creep into JS behind the scenes, the future of more optimized program-level performance in JS looks really *promising*.

You Don't Know JS: Async & Performance

Chapter 6: Benchmarking & Tuning

As the first four chapters of this book were all about performance as a coding pattern (asynchrony and concurrency), and Chapter 5 was about performance at the macro program architecture level, this chapter goes after the topic of performance at the micro level, focusing on single expressions/statements.

One of the most common areas of curiosity – indeed, some developers can get quite obsessed about it – is in analyzing and testing various options for how to write a line or chunk of code, and which one is faster.

We're going to look at some of these issues, but it's important to understand from the outset that this chapter is **not** about feeding the obsession of micro-performance tuning, like whether some given JS engine can run `++a` faster than `a++`. The more important goal of this chapter is to figure out what kinds of JS performance matter and which ones don't, *and how to tell the difference*.

But even before we get there, we need to explore how to most accurately and reliably test JS performance, because there's tons of misconceptions and myths that have flooded our collective cult knowledge base. We've got to sift through all that junk to find some clarity.

Benchmarking

OK, time to start dispelling some misconceptions. I'd wager the vast majority of JS developers, if asked to benchmark the speed (execution time) of a certain operation, would initially go about it something like this:

```
var start = (new Date()).getTime(); // or `Date.now()`  
  
// do some operation  
  
var end = (new Date()).getTime();  
  
console.log( "Duration:", (end - start) );
```

Raise your hand if that's roughly what came to your mind. Yep, I thought so. There's a lot wrong with this approach, but don't feel bad; **we've all been there.**

What did that measurement tell you, exactly? Understanding what it does and doesn't say about the execution time of the operation in question is key to learning how to appropriately benchmark performance in JavaScript.

If the duration reported is 0, you may be tempted to believe that it took less than a millisecond. But that's not very accurate. Some platforms don't have single millisecond precision, but instead only update the timer in larger increments.

For example, older versions of windows (and thus IE) had only 15ms precision, which means the operation has to take at least that long for anything other than 0 to be reported!

Moreover, whatever duration is reported, the only thing you really know is that the operation took approximately that long on that exact single run. You have near-zero confidence that it will always run at that speed. You have no idea if the engine or system had some sort of interference at that exact moment, and that at other times the operation could run faster.

What if the duration reported is 4? Are you more sure it took about four milliseconds? Nope. It might have taken less time, and there may have been some other delay in getting either `start` or `end` timestamps.

More troublingly, you also don't know that the circumstances of this operation test aren't overly optimistic. It's possible that the JS engine figured out a way to optimize your isolated test case, but in a more real program such optimization would be diluted or impossible, such that the operation would run slower than your test.

So... what do we know? Unfortunately, with those realizations stated, **we know very little**. Something of such low confidence isn't even remotely good enough to build your determinations on. Your "benchmark" is basically useless. And worse, it's dangerous in that it implies false confidence, not just to you but also to others who don't think critically about the conditions that led to those results.

Repetition

"OK," you now say, "Just put a loop around it so the whole test takes longer." If you repeat an operation 100 times, and that whole loop reportedly takes a total of 137ms, then you can just divide by 100 and get an average duration of 1.37ms for each operation, right?

Well, not exactly.

A straight mathematical average by itself is definitely not sufficient for making judgments about performance which you plan to extrapolate to the breadth of your entire application. With a hundred iterations, even a couple of outliers (high or low) can skew the average, and then when you apply that conclusion repeatedly, you even further inflate the skew beyond credulity.

Instead of just running for a fixed number of iterations, you can instead choose to run the loop of tests until a certain amount of time has passed. That might be more reliable, but how do you decide how long to run? You might guess that it should be some multiple of how long your operation should take to run once. Wrong.

Actually, the length of time to repeat across should be based on the accuracy of the timer you're using, specifically to minimize the chances of inaccuracy. The

less precise your timer, the longer you need to run to make sure you've minimized the error percentage. A 15ms timer is pretty bad for accurate benchmarking; to minimize its uncertainty (aka “error rate”) to less than 1%, you need to run your each cycle of test iterations for 750ms. A 1ms timer only needs a cycle to run for 50ms to get the same confidence.

But then, that's just a single sample. To be sure you're factoring out the skew, you'll want lots of samples to average across. You'll also want to understand something about just how slow the worst sample is, how fast the best sample is, how far apart those best and worse cases were, and so on. You'll want to know not just a number that tells you how fast something ran, but also to have some quantifiable measure of how trustable that number is.

Also, you probably want to combine these different techniques (as well as others), so that you get the best balance of all the possible approaches.

That's all bare minimum just to get started. If you've been approaching performance benchmarking with anything less serious than what I just glossed over, well... “you don't know: proper benchmarking.”

Benchmark.js

Any relevant and reliable benchmark should be based on statistically sound practices. I am not going to write a chapter on statistics here, so I'll hand wave around some terms: standard deviation, variance, margin of error. If you don't know what those terms really mean – I took a stats class back in college and I'm still a little fuzzy on them – you are not actually qualified to write your own benchmarking logic.

Luckily, smart folks like John-David Dalton and Mathias Bynens do understand these concepts, and wrote a statistically sound benchmarking tool called Benchmark.js (<http://benchmarkjs.com/>). So I can end the suspense by simply saying: “just use that tool.”

I won't repeat their whole documentation for how Benchmark.js works; they have fantastic API Docs (<http://benchmarkjs.com/docs>) you should read. Also there are some great (<http://calendar.perfplanet.com/2010/bulletproof-javascript-benchmarks/>) writeups (<http://monsur.hossa.in/2012/12/11/benchmarkjs.html>) on more of the details and methodology.

But just for quick illustration purposes, here's how you could use Benchmark.js to run a quick performance test:

```
function foo() {
  // operation(s) to test
}

var bench = new Benchmark(
  "foo test",           // test name
  function() {
    // test code
  }
);
```

```

    foo,                      // function to test (just contents)
    {
        // ...
    }
);

bench.hz;                  // number of operations per second
bench.stats.moe;           // margin of error
bench.stats.variance;       // variance across samples
// ...

```

There's *lots* more to learn about using Benchmark.js besides this glance I'm including here. But the point is that it's handling all of the complexities of setting up a fair, reliable, and valid performance benchmark for a given piece of JavaScript code. If you're going to try to test and benchmark your code, this library is the first place you should turn.

We're showing here the usage to test a single operation like X, but it's fairly common that you want to compare X to Y. This is easy to do by simply setting up two different tests in a "Suite" (a Benchmark.js organizational feature). Then, you run them head-to-head, and compare the statistics to conclude whether X or Y was faster.

Benchmark.js can of course be used to test JavaScript in a browser (see the "jsPerf.com" section later in this chapter), but it can also run in non-browser environments (Node.js, etc.).

One largely untapped potential use-case for Benchmark.js is to use it in your Dev or QA environments to run automated performance regression tests against critical path parts of your application's JavaScript. Similar to how you might run unit test suites before deployment, you can also compare the performance against previous benchmarks to monitor if you are improving or degrading application performance.

Setup/Teardown In the previous code snippet, we glossed over the "extra options" `{ .. }` object. But there are two options we should discuss: `setup` and `teardown`.

These two options let you define functions to be called before and after your test case runs.

It's incredibly important to understand that your `setup` and `teardown` code **does not run for each test iteration**. The best way to think about it is that there's an outer loop (repeating cycles), and an inner loop (repeating test iterations). `setup` and `teardown` are run at the beginning and end of each *outer* loop (aka cycle) iteration, but not inside the inner loop.

Why does this matter? Let's imagine you have a test case that looks like this:

```
a = a + "w";
b = a.charAt( 1 );
```

Then, you set up your test setup as follows:

```
var a = "x";
```

Your temptation is probably to believe that `a` is starting out as "x" for each test iteration.

But it's not! It's starting `a` at "x" for each test cycle, and then your repeated `+ "w"` concatenations will be making a larger and larger `a` value, even though you're only ever accessing the character "w" at the 1 position.

Where this most commonly bites you is when you make side effect changes to something like the DOM, like appending a child element. You may think your parent element is set as empty each time, but it's actually getting lots of elements added, and that can significantly sway the results of your tests.

Context Is King

Don't forget to check the context of a particular performance benchmark, especially a comparison between X and Y tasks. Just because your test reveals that X is faster than Y doesn't mean that the conclusion "X is faster than Y" is actually relevant.

For example, let's say a performance test reveals that X runs 10,000,000 operations per second, and Y runs at 8,000,000 operations per second. You could claim that Y is 20% slower than X, and you'd be mathematically correct, but your assertion doesn't hold as much water as you'd think.

Let's think about the results more critically: 10,000,000 operations per second is 10,000 operations per millisecond, and 10 operations per microsecond. In other words, a single operation takes 0.1 microseconds, or 100 nanoseconds. It's hard to fathom just how small 100ns is, but for comparison, it's often cited that the human eye isn't generally capable of distinguishing anything less than 100ms, which is one million times slower than the 100ns speed of the X operation.

Even recent scientific studies showing that maybe the brain can process as quick as 13ms (about 8x faster than previously asserted) would mean that X is still running 125,000 times faster than the human brain can perceive a distinct thing happening. **X is going really, really fast.**

But more importantly, let's talk about the difference between X and Y, the 2,000,000 operations per second difference. If X takes 100ns, and Y takes 80ns, the difference is 20ns, which in the best case is still one 650-thousandth of the interval the human brain can perceive.

What's my point? **None of this performance difference matters, at all!**

But wait, what if this operation is going to happen a whole bunch of times in a row? Then the difference could add up, right?

OK, so what we're asking then is, how likely is it that operation X is going to be run over and over again, one right after the other, and that this has to happen 650,000 times just to get a sliver of a hope the human brain could perceive it. More likely, it'd have to happen 5,000,000 to 10,000,000 times together in a tight loop to even approach relevance.

While the computer scientist in you might protest that this is possible, the louder voice of realism in you should sanity check just how likely or unlikely that really is. Even if it is relevant in rare occasions, it's irrelevant in most situations.

The vast majority of your benchmark results on tiny operations – like the `++x` vs `x++` myth – **are just totally bogus** for supporting the conclusion that X should be favored over Y on a performance basis.

Engine Optimizations

You simply cannot reliably extrapolate that if X was 10 microseconds faster than Y in your isolated test, that means X is always faster than Y and should always be used. That's not how performance works. It's vastly more complicated.

For example, let's imagine (purely hypothetical) that you test some microperformance behavior such as comparing:

```
var twelve = "12";
var foo = "foo";

// test 1
var X1 = parseInt( twelve );
var X2 = parseInt( foo );

// test 2
var Y1 = Number( twelve );
var Y2 = Number( foo );
```

If you understand what `parseInt(..)` does compared to `Number(..)`, you might intuit that `parseInt(..)` potentially has “more work” to do, especially in the `foo` case. Or you might intuit that they should have the same amount of work to do in the `foo` case, as both should be able to stop at the first character “`f`”.

Which intuition is correct? I honestly don't know. But I'll make the case it doesn't matter what your intuition is. What might the results be when you test it? Again, I'm making up a pure hypothetical here, I haven't actually tried, nor do I care.

Let's pretend the test comes back that `X` and `Y` are statistically identical. Have you then confirmed your intuition about the "`f`" character thing? Nope.

It's possible in our hypothetical that the engine might recognize that the variables `twelve` and `foo` are only being used in one place in each test, and so it might decide to inline those values. Then it may realize that `Number("12")` can just be replaced by `12`. And maybe it comes to the same conclusion with `parseInt(..)`, or maybe not.

Or an engine's dead-code removal heuristic could kick in, and it could realize that variables `X` and `Y` aren't being used, so declaring them is irrelevant, so it doesn't end up doing anything at all in either test.

And all that's just made with the mindset of assumptions about a single test run. Modern engines are fantastically more complicated than what we're intuiting here. They do all sorts of tricks, like tracing and tracking how a piece of code behaves over a short period of time, or with a particularly constrained set of inputs.

What if the engine optimizes a certain way because of the fixed input, but in your real program you give more varied input and the optimization decisions shake out differently (or not at all!)? Or what if the engine kicks in optimizations because it sees the code being run tens of thousands of times by the benchmarking utility, but in your real program it will only run a hundred times in near proximity, and under those conditions the engine determines the optimizations are not worth it?

And all those optimizations we just hypothesized about might happen in our constrained test but maybe the engine wouldn't do them in a more complex program (for various reasons). Or it could be reversed – the engine might not optimize such trivial code but may be more inclined to optimize it more aggressively when the system is already more taxed by a more sophisticated program.

The point I'm trying to make is that you really don't know for sure exactly what's going on under the covers. All the guesses and hypothesis you can muster don't amount to hardly anything concrete for really making such decisions.

Does that mean you can't really do any useful testing? **Definitely not!**

What this boils down to is that testing *not real* code gives you *not real* results. In so much as is possible and practical, you should test actual real, non-trivial snippets of your code, and under as best of real conditions as you can actually hope to. Only then will the results you get have a chance to approximate reality.

Microbenchmarks like `++x` vs `x++` are so incredibly likely to be bogus, we might as well just flatly assume them as such.

jsPerf.com

While Benchmark.js is useful for testing the performance of your code in whatever JS environment you're running, it cannot be stressed enough that you need to compile test results from lots of different environments (desktop browsers, mobile devices, etc.) if you want to have any hope of reliable test conclusions.

For example, Chrome on a high-end desktop machine is not likely to perform anywhere near the same as Chrome mobile on a smartphone. And a smartphone with a full battery charge is not likely to perform anywhere near the same as a smartphone with 2% battery life left, when the device is starting to power down the radio and processor.

If you want to make assertions like "X is faster than Y" in any reasonable sense across more than just a single environment, you're going to need to actually test as many of those real world environments as possible. Just because Chrome executes some X operation faster than Y doesn't mean that all browsers do. And of course you also probably will want to cross-reference the results of multiple browser test runs with the demographics of your users.

There's an awesome website for this purpose called jsPerf (<http://jsperf.com>). It uses the Benchmark.js library we talked about earlier to run statistically accurate and reliable tests, and makes the test on an openly available URL that you can pass around to others.

Each time a test is run, the results are collected and persisted with the test, and the cumulative test results are graphed on the page for anyone to see.

When creating a test on the site, you start out with two test cases to fill in, but you can add as many as you need. You also have the ability to set up `setup` code that is run at the beginning of each test cycle and `teardown` code run at the end of each cycle.

Note: A trick for doing just one test case (if you're benchmarking a single approach instead of a head-to-head) is to fill in the second test input boxes with placeholder text on first creation, then edit the test and leave the second test blank, which will delete it. You can always add more test cases later.

You can define the initial page setup (importing libraries, defining utility helper functions, declaring variables, etc.). There are also options for defining setup and teardown behavior if needed – consult the "Setup/Teardown" section in the Benchmark.js discussion earlier.

Sanity Check

jsPerf is a fantastic resource, but there's an awful lot of tests published that when you analyze them are quite flawed or bogus, for any of a variety of reasons as outlined so far in this chapter.

Consider:

```

// Case 1
var x = [];
for (var i=0; i<10; i++) {
  x[i] = "x";
}

// Case 2
var x = [];
for (var i=0; i<10; i++) {
  x[x.length] = "x";
}

// Case 3
var x = [];
for (var i=0; i<10; i++) {
  x.push("x");
}

```

Some observations to ponder about this test scenario:

- It's extremely common for devs to put their own loops into test cases, and they forget that Benchmark.js already does all the repetition you need. There's a really strong chance that the `for` loops in these cases are totally unnecessary noise.
- The declaring and initializing of `x` is included in each test case, possibly unnecessarily. Recall from earlier that if `x = []` were in the `setup` code, it wouldn't actually be run before each test iteration, but instead once at the beginning of each cycle. That means `x` would continue growing quite large, not just the size 10 implied by the `for` loops.

So is the intent to make sure the tests are constrained only to how the JS engine behaves with very small arrays (size 10)? That *could* be the intent, but if it is, you have to consider if that's not focusing far too much on nuanced internal implementation details.

On the other hand, does the intent of the test embrace the context that the arrays will actually be growing quite large? Is the JS engines' behavior with larger arrays relevant and accurate when compared with the intended real world usage?

- Is the intent to find out how much `x.length` or `x.push(..)` add to the performance of the operation to append to the `x` array? OK, that might be a valid thing to test. But then again, `push(..)` is a function call, so of course it's going to be slower than `[..]` access. Arguably, cases 1 and 2 are fairer than case 3.

Here's another example that illustrates a common apples-to-oranges flaw:

```
// Case 1
```

```

var x = ["John", "Albert", "Sue", "Frank", "Bob"];
x.sort();

// Case 2
var x = ["John", "Albert", "Sue", "Frank", "Bob"];
x.sort( function mySort(a,b){
    if (a < b) return -1;
    if (a > b) return 1;
    return 0;
} );

```

Here, the obvious intent is to find out how much slower the custom `mySort(..)` comparator is than the built-in default comparator. But by specifying the function `mySort(..)` as inline function expression, you've created an unfair/bogus test. Here, the second case is not only testing a custom user JS function, **but it's also testing creating a new function expression for each iteration.**

Would it surprise you to find out that if you run a similar test but update it to isolate only for creating an inline function expression versus using a pre-declared function, the inline function expression creation can be from 2% to 20% slower!?

Unless your intent with this test *is* to consider the inline function expression creation “cost,” a better/fairer test would put `mySort(..)`’s declaration in the page setup – don’t put it in the test `setup` as that’s unnecessary redeclaration for each cycle – and simply reference it by name in the test case: `x.sort(mySort)`.

Building on the previous example, another pitfall is in opaquely avoiding or adding “extra work” to one test case that creates an apples-to-oranges scenario:

```

// Case 1
var x = [12, -14, 0, 3, 18, 0, 2.9];
x.sort();

// Case 2
var x = [12, -14, 0, 3, 18, 0, 2.9];
x.sort( function mySort(a,b){
    return a - b;
} );

```

Setting aside the previously mentioned inline function expression pitfall, the second case’s `mySort(..)` works in this case because you have provided it numbers, but would have of course failed with strings. The first case doesn’t throw an error, but it actually behaves differently and has a different outcome! It should be obvious, but: **a different outcome between two test cases almost certainly invalidates the entire test!**

But beyond the different outcomes, in this case, the built in `sort(..)`’s comparator is actually doing “extra work” that `mySort()` does not, in that the built-in one coerces the compared values to strings and does lexicographic com-

parison. The first snippet results in `[-14, 0, 0, 12, 18, 2.9, 3]` while the second snippet results (likely more accurately based on intent) in `[-14, 0, 0, 2.9, 3, 12, 18]`.

So that test is unfair because it's not actually doing the same task between the cases. Any results you get are bogus.

These same pitfalls can even be much more subtle:

```
// Case 1
var x = false;
var y = x ? 1 : 2;

// Case 2
var x;
var y = x ? 1 : 2;
```

Here, the intent might be to test the performance impact of the coercion to a Boolean that the `? :` operator will do if the `x` expression is not already a Boolean (see the *Types & Grammar* title of this book series). So, you're apparently OK with the fact that there is extra work to do the coercion in the second case.

The subtle problem? You're setting `x`'s value in the first case and not setting it in the other, so you're actually doing work in the first case that you're not doing in the second. To eliminate any potential (albeit minor) skew, try:

```
// Case 1
var x = false;
var y = x ? 1 : 2;

// Case 2
var x = undefined;
var y = x ? 1 : 2;
```

Now there's an assignment in both cases, so the thing you want to test – the coercion of `x` or not – has likely been more accurately isolated and tested.

Writing Good Tests

Let me see if I can articulate the bigger point I'm trying to make here.

Good test authoring requires careful analytical thinking about what differences exist between two test cases and whether the differences between them are *intentional* or *unintentional*.

Intentional differences are of course normal and OK, but it's too easy to create unintentional differences that skew your results. You have to be really, really careful to avoid that skew. Moreover, you may intend a difference but it may not be obvious to other readers of your test what your intent was, so they may doubt (or trust!) your test incorrectly. How do you fix that?

Write better, clearer tests. But also, take the time to document (using the jsPerf.com “Description” field and/or code comments) exactly what the intent of your test is, even to the nuanced detail. Call out the intentional differences, which will help others and your future self to better identify unintentional differences that could be skewing the test results.

Isolate things which aren’t relevant to your test by pre-declaring them in the page or test setup settings so they’re outside the timed parts of the test.

Instead of trying to narrow in on a tiny snippet of your real code and benchmarking just that piece out of context, tests and benchmarks are better when they include a larger (while still relevant) context. Those tests also tend to run slower, which means any differences you spot are more relevant in context.

Micropreformance

OK, until now we’ve been dancing around various micropreformance issues and generally looking disfavorably upon obsessing about them. I want to take just a moment to address them directly.

The first thing you need to get more comfortable with when thinking about performance benchmarking your code is that the code you write is not always the code the engine actually runs. We briefly looked at that topic back in Chapter 1 when we discussed statement reordering by the compiler, but here we’re going to suggest the compiler can sometimes decide to run different code than you wrote, not just in different orders but different in substance.

Let’s consider this piece of code:

```
var foo = 41;

(function(){
  (function(){
    (function(baz){
      var bar = foo + baz;
      // ..
    })(1);
  })();
})();
```

You may think about the `foo` reference in the innermost function as needing to do a three-level scope lookup. We covered in the *Scope & Closures* title of this book series how lexical scope works, and the fact that the compiler generally caches such lookups so that referencing `foo` from different scopes doesn’t really practically “cost” anything extra.

But there’s something deeper to consider. What if the compiler realizes that `foo` isn’t referenced anywhere else but that one location, and it further notices that the value never is anything except the 41 as shown?

Isn't it quite possible and acceptable that the JS compiler could decide to just remove the `foo` variable entirely, and *inline* the value, such as this:

```
(function(){  
  (function(){  
    (function(baz){  
      var bar = 41 + baz;  
      // ...  
    })(1);  
  })();  
})();
```

Note: Of course, the compiler could probably also do a similar analysis and rewrite with the `baz` variable here, too.

When you begin to think about your JS code as being a hint or suggestion to the engine of what to do, rather than a literal requirement, you realize that a lot of the obsession over discrete syntactic minutia is most likely unfounded.

Another example:

```
function factorial(n) {  
  if (n < 2) return 1;  
  return n * factorial(n - 1);  
}  
  
factorial(5); // 120
```

Ah, the good ol' fashioned “factorial” algorithm! You might assume that the JS engine will run that code mostly as is. And to be honest, it might – I’m not really sure.

But as an anecdote, the same code expressed in C and compiled with advanced optimizations would result in the compiler realizing that the call `factorial(5)` can just be replaced with the constant value 120, eliminating the function and call entirely!

Moreover, some engines have a practice called “unrolling recursion,” where it can realize that the recursion you’ve expressed can actually be done “easier” (i.e., more optimally) with a loop. It’s possible the preceding code could be *rewritten* by a JS engine to run as:

```
function factorial(n) {  
  if (n < 2) return 1;  
  
  var res = 1;  
  for (var i=n; i>1; i--) {  
    res *= i;  
  }  
  return res;
```

```
}

factorial( 5 );      // 120
```

Now, let's imagine that in the earlier snippet you had been worried about whether `n * factorial(n-1)` or `n *= factorial(--n)` runs faster. Maybe you even did a performance benchmark to try to figure out which was better. But you miss the fact that in the bigger context, the engine may not run either line of code because it may unroll the recursion!

Speaking of `--`, `--n` versus `n--` is often cited as one of those places where you can optimize by choosing the `--n` version, because theoretically it requires less effort down at the assembly level of processing.

That sort of obsession is basically nonsense in modern JavaScript. That's the kind of thing you should be letting the engine take care of. You should write the code that makes the most sense. Compare these three `for` loops:

```
// Option 1
for (var i=0; i<10; i++) {
    console.log( i );
}

// Option 2
for (var i=0; i<10; ++i) {
    console.log( i );
}

// Option 3
for (var i=-1; ++i<10; ) {
    console.log( i );
}
```

Even if you have some theory where the second or third option is more performant than the first option by a tiny bit, which is dubious at best, the third loop is more confusing because you have to start with `-1` for `i` to account for the fact that `++i` pre-increment is used. And the difference between the first and second options is really quite irrelevant.

It's entirely possible that a JS engine may see a place where `i++` is used and realize that it can safely replace it with the `++i` equivalent, which means your time spent deciding which one to pick was completely wasted and the outcome moot.

Here's another common example of silly microperformance obsession:

```
var x = [ .. ];

// Option 1
for (var i=0; i < x.length; i++) {
```

```

        // ...
}

// Option 2
for (var i=0, len = x.length; i < len; i++) {
    // ...
}

```

The theory here goes that you should cache the length of the `x` array in the variable `len`, because ostensibly it doesn't change, to avoid paying the price of `x.length` being consulted for each iteration of the loop.

If you run performance benchmarks around `x.length` usage compared to caching it in a `len` variable, you'll find that while the theory sounds nice, in practice any measured differences are statistically completely irrelevant.

In fact, in some engines like v8, it can be shown (<http://mrale.ph/blog/2014/12/24/array-length-caching.html>) that you could make things slightly worse by pre-caching the length instead of letting the engine figure it out for you. Don't try to outsmart your JavaScript engine, you'll probably lose when it comes to performance optimizations.

Not All Engines Are Alike

The different JS engines in various browsers can all be “spec compliant” while having radically different ways of handling code. The JS specification doesn't require anything performance related – well, except ES6's “Tail Call Optimization” covered later in this chapter.

The engines are free to decide that one operation will receive its attention to optimize, perhaps trading off for lesser performance on another operation. It can be very tenuous to find an approach for an operation that always runs faster in all browsers.

There's a movement among some in the JS dev community, especially those who work with Node.js, to analyze the specific internal implementation details of the v8 JavaScript engine and make decisions about writing JS code that is tailored to take best advantage of how v8 works. You can actually achieve a surprisingly high degree of performance optimization with such endeavors, so the payoff for the effort can be quite high.

Some commonly cited examples (<https://github.com/petkaantonov/bluebird/wiki/Optimization-killers>) for v8:

- Don't pass the `arguments` variable from one function to any other function, as such “leakage” slows down the function implementation.
- Isolate a `try..catch` in its own function. Browsers struggle with optimizing any function with a `try..catch` in it, so moving that construct to its

own function means you contain the de-optimization harm while letting the surrounding code be optimizable.

But rather than focus on those tips specifically, let's sanity check the v8-only optimization approach in a general sense.

Are you genuinely writing code that only needs to run in one JS engine? Even if your code is entirely intended for Node.js *right now*, is the assumption that v8 will *always* be the used JS engine reliable? Is it possible that someday a few years from now, there's another server-side JS platform besides Node.js that you choose to run your code on? What if what you optimized for before is now a much slower way of doing that operation on the new engine?

Or what if your code always stays running on v8 from here on out, but v8 decides at some point to change the way some set of operations works such that what used to be fast is now slow, and vice versa?

These scenarios aren't just theoretical, either. It used to be that it was faster to put multiple string values into an array and then call `join("")` on the array to concatenate the values than to just use `+` concatenation directly with the values. The historical reason for this is nuanced, but it has to do with internal implementation details about how string values were stored and managed in memory.

As a result, "best practice" advice at the time disseminated across the industry suggesting developers always use the array `join(..)` approach. And many followed.

Except, somewhere along the way, the JS engines changed approaches for internally managing strings, and specifically put in optimizations for `+` concatenation. They didn't slow down `join(..)` per se, but they put more effort into helping `+` usage, as it was still quite a bit more widespread.

Note: The practice of standardizing or optimizing some particular approach based mostly on its existing widespread usage is often called (metaphorically) "paving the cowpath."

Once that new approach to handling strings and concatenation took hold, unfortunately all the code out in the wild that was using array `join(..)` to concatenate strings was then sub-optimal.

Another example: at one time, the Opera browser differed from other browsers in how it handled the boxing/unboxing of primitive wrapper objects (see the *Types & Grammar* title of this book series). As such, their advice to developers was to use a `String` object instead of the primitive `string` value if properties like `length` or methods like `charAt(..)` needed to be accessed. This advice may have been correct for Opera at the time, but it was literally completely opposite for other major contemporary browsers, as they had optimizations specifically for the `string` primitives and not their object wrapper counterparts.

I think these various gotchas are at least possible, if not likely, for code even today. So I'm very cautious about making wide ranging performance optimizations in my JS code based purely on engine implementation details, **especially if those details are only true of a single engine.**

The reverse is also something to be wary of: you shouldn't necessarily change a piece of code to work around one engine's difficulty with running a piece of code in an acceptably performant way.

Historically, IE has been the brunt of many such frustrations, given that there have been plenty of scenarios in older IE versions where it struggled with some performance aspect that other major browsers of the time seemed not to have much trouble with. The string concatenation discussion we just had was actually a real concern back in the IE6 and IE7 days, where it was possible to get better performance out of `join(..)` than `+`.

But it's troublesome to suggest that just one browser's trouble with performance is justification for using a code approach that quite possibly could be sub-optimal in all other browsers. Even if the browser in question has a large market share for your site's audience, it may be more practical to write the proper code and rely on the browser to update itself with better optimizations eventually.

“There is nothing more permanent than a temporary hack.” Chances are, the code you write now to work around some performance bug will probably outlive the performance bug in the browser itself.

In the days when a browser only updated once every five years, that was a tougher call to make. But as it stands now, browsers across the board are updating at a much more rapid interval (though obviously the mobile world still lags), and they're all competing to optimize web features better and better.

If you run across a case where a browser *does* have a performance wart that others don't suffer from, make sure to report it to them through whatever means you have available. Most browsers have open public bug trackers suitable for this purpose.

Tip: I'd only suggest working around a performance issue in a browser if it was a really drastic show-stopper, not just an annoyance or frustration. And I'd be very careful to check that the performance hack didn't have noticeable negative side effects in another browser.

Big Picture

Instead of worrying about all these microperformance nuances, we should instead be looking at big-picture types of optimizations.

How do you know what's big picture or not? You have to first understand if your code is running on a critical path or not. If it's not on the critical path, chances are your optimizations are not worth much.

Ever heard the admonition, “that’s premature optimization!”? It comes from a famous quote from Donald Knuth: “premature optimization is the root of all evil.”. Many developers cite this quote to suggest that most optimizations are “premature” and are thus a waste of effort. The truth is, as usual, more nuanced.

Here is Knuth’s quote, in context:

Programmers waste enormous amounts of time thinking about, or worrying about, the speed of **noncritical** parts of their programs, and these attempts at efficiency actually have a strong negative impact when debugging and maintenance are considered. We should forget about small efficiencies, say about 97% of the time: premature optimization is the root of all evil. Yet we should not pass up our opportunities in that **critical** 3%. [emphasis added]

(http://web.archive.org/web/20130731202547/http://pplab.snu.ac.kr/courses/adv_pl05/papers/p261-knuth.pdf, Computing Surveys, Vol 6, No 4, December 1974)

I believe it’s a fair paraphrasing to say that Knuth *meant*: “non-critical path optimization is the root of all evil.” So the key is to figure out if your code is on the critical path – you should optimize it! – or not.

I’d even go so far as to say this: no amount of time spent optimizing critical paths is wasted, no matter how little is saved; but no amount of optimization on noncritical paths is justified, no matter how much is saved.

If your code is on the critical path, such as a “hot” piece of code that’s going to be run over and over again, or in UX critical places where users will notice, like an animation loop or CSS style updates, then you should spare no effort in trying to employ relevant, measurably significant optimizations.

For example, consider a critical path animation loop that needs to coerce a string value to a number. There are of course multiple ways to do that (see the *Types & Grammar* title of this book series), but which one if any is the fastest?

```
var x = "42";    // need number `42`  
  
// Option 1: let implicit coercion automatically happen  
var y = x / 2;  
  
// Option 2: use `parseInt(..)`  
var y = parseInt( x, 0 ) / 2;  
  
// Option 3: use `Number(..)`  
var y = Number( x ) / 2;  
  
// Option 4: use `+` unary operator  
var y = +x / 2;
```

```
// Option 5: use `/` unary operator
var y = (x | 0) / 2;
```

Note: I will leave it as an exercise to the reader to set up a test if you’re interested in examining the minute differences in performance among these options.

When considering these different options, as they say, “One of these things is not like the others.” `parseInt(..)` does the job, but it also does a lot more – it parses the string rather than just coercing. You can probably guess, correctly, that `parseInt(..)` is a slower option, and you should probably avoid it.

Of course, if `x` can ever be a value that **needs parsing**, such as “42px” (like from a CSS style lookup), then `parseInt(..)` really is the only suitable option!

`Number(..)` is also a function call. From a behavioral perspective, it’s identical to the `+ unary operator` option, but it may in fact be a little slower, requiring more machinery to execute the function. Of course, it’s also possible that the JS engine recognizes this behavioral symmetry and just handles the inlining of `Number(..)`’s behavior (aka `+x`) for you!

But remember, obsessing about `+x` versus `x | 0` is in most cases likely a waste of effort. This is a microperformance issue, and one that you shouldn’t let dictate/upgrade the readability of your program.

While performance is very important in critical paths of your program, it’s not the only factor. Among several options that are roughly similar in performance, readability should be another important concern.

Tail Call Optimization (TCO)

As we briefly mentioned earlier, ES6 includes a specific requirement that ventures into the world of performance. It’s related to a specific form of optimization that can occur with function calls: *tail call optimization*.

Briefly, a “tail call” is a function call that appears at the “tail” of another function, such that after the call finishes, there’s nothing left to do (except perhaps return its result value).

For example, here’s a non-recursive setup with tail calls:

```
function foo(x) {
    return x;
}

function bar(y) {
    return foo(y + 1);    // tail call
}

function baz() {
    return 1 + bar(40);  // not tail call
}
```

```

}

baz();           // 42

```

`foo(y+1)` is a tail call in `bar(..)` because after `foo(..)` finishes, `bar(..)` is also finished except in this case returning the result of the `foo(..)` call. However, `bar(40)` is *not* a tail call because after it completes, its result value must be added to 1 before `baz()` can return it.

Without getting into too much nitty-gritty detail, calling a new function requires an extra amount of reserved memory to manage the call stack, called a “stack frame.” So the preceding snippet would generally require a stack frame for each of `baz()`, `bar(..)`, and `foo(..)` all at the same time.

However, if a TCO-capable engine can realize that the `foo(y+1)` call is in *tail position* meaning `bar(..)` is basically complete, then when calling `foo(..)`, it doesn’t need to create a new stack frame, but can instead reuse the existing stack frame from `bar(..)`. That’s not only faster, but it also uses less memory.

That sort of optimization isn’t a big deal in a simple snippet, but it becomes a *much bigger deal* when dealing with recursion, especially if the recursion could have resulted in hundreds or thousands of stack frames. With TCO the engine can perform all those calls with a single stack frame!

Recursion is a hairy topic in JS because without TCO, engines have had to implement arbitrary (and different!) limits to how deep they will let the recursion stack get before they stop it, to prevent running out of memory. With TCO, recursive functions with *tail position* calls can essentially run unbounded, because there’s never any extra usage of memory!

Consider that recursive `factorial(..)` from before, but rewritten to make it TCO friendly:

```

function factorial(n) {
  function fact(n,res) {
    if (n < 2) return res;

    return fact( n - 1, n * res );
  }

  return fact( n, 1 );
}

factorial( 5 );      // 120

```

This version of `factorial(..)` is still recursive, but it’s also optimizable with TCO, because both inner `fact(..)` calls are in *tail position*.

Note: It’s important to note that TCO only applies if there’s actually a tail call. If you write recursive functions without tail calls, the performance will

still fall back to normal stack frame allocation, and the engines' limits on such recursive call stacks will still apply. Many recursive functions can be rewritten as we just showed with `factorial(..)`, but it takes careful attention to detail.

One reason that ES6 requires engines to implement TCO rather than leaving it up to their discretion is because the *lack of TCO* actually tends to reduce the chances that certain algorithms will be implemented in JS using recursion, for fear of the call stack limits.

If the lack of TCO in the engine would just gracefully degrade to slower performance in all cases, it wouldn't probably have been something that ES6 needed to *require*. But because the lack of TCO can actually make certain programs impractical, it's more an important feature of the language than just a hidden implementation detail.

ES6 guarantees that from now on, JS developers will be able to rely on this optimization across all ES6+ compliant browsers. That's a win for JS performance!

Review

Effectively benchmarking performance of a piece of code, especially to compare it to another option for that same code to see which approach is faster, requires careful attention to detail.

Rather than rolling your own statistically valid benchmarking logic, just use the `Benchmark.js` library, which does that for you. But be careful about how you author tests, because it's far too easy to construct a test that seems valid but that's actually flawed – even tiny differences can skew the results to be completely unreliable.

It's important to get as many test results from as many different environments as possible to eliminate hardware/device bias. jsPerf.com is a fantastic website for crowdsourcing performance benchmark test runs.

Many common performance tests unfortunately obsess about irrelevant micropreformance details like `x++` versus `++x`. Writing good tests means understanding how to focus on big picture concerns, like optimizing on the critical path, and avoiding falling into traps like different JS engines' implementation details.

Tail call optimization (TCO) is a required optimization as of ES6 that will make some recursive patterns practical in JS where they would have been impossible otherwise. TCO allows a function call in the *tail position* of another function to execute without needing any extra resources, which means the engine no longer needs to place arbitrary restrictions on call stack depth for recursive algorithms.

You Don't Know JS: Async & Performance

Appendix A: *asynquence* Library

Chapters 1 and 2 went into quite a bit of detail about typical asynchronous programming patterns and how they're commonly solved with callbacks. But we also saw why callbacks are fatally limited in capability, which led us to Chapters 3 and 4, with Promises and generators offering a much more solid, trustable, and reasonable base to build your asynchrony on.

I referenced my own asynchronous library *asynquence* (<http://github.com/getify/asynquence>) – “async” + “sequence” = “asynquence” – several times in this book, and I want to now briefly explain how it works and why its unique design is important and helpful.

In the next appendix, we'll explore some advanced async patterns, but you'll probably want a library to make those palatable enough to be useful. We'll use *asynquence* to express those patterns, so you'll want to spend a little time here getting to know the library first.

asynquence is obviously not the only option for good async coding; certainly there are many great libraries in this space. But *asynquence* provides a unique perspective by combining the best of all these patterns into a single library, and moreover is built on a single basic abstraction: the (async) sequence.

My premise is that sophisticated JS programs often need bits and pieces of various different asynchronous patterns woven together, and this is usually left entirely up to each developer to figure out. Instead of having to bring in two or more different async libraries that focus on different aspects of asynchrony, *asynquence* unifies them into variated sequence steps, with just one core library to learn and deploy.

I believe the value is strong enough with *asynquence* to make async flow control programming with Promise-style semantics super easy to accomplish, so that's why we'll exclusively focus on that library here.

To begin, I'll explain the design principles behind *asynquence*, and then we'll illustrate how its API works with code examples.

Sequences, Abstraction Design

Understanding *asynquence* begins with understanding a fundamental abstraction: any series of steps for a task, whether they separately are synchronous or asynchronous, can be collectively thought of as a “sequence”. In other words, a sequence is a container that represents a task, and is comprised of individual (potentially async) steps to complete that task.

Each step in the sequence is controlled under the covers by a Promise (see Chapter 3). That is, every step you add to a sequence implicitly creates a Promise

that is wired to the previous end of the sequence. Because of the semantics of Promises, every single step advancement in a sequence is asynchronous, even if you synchronously complete the step.

Moreover, a sequence will always proceed linearly from step to step, meaning that step 2 always comes after step 1 finishes, and so on.

Of course, a new sequence can be forked off an existing sequence, meaning the fork only occurs once the main sequence reaches that point in the flow. Sequences can also be combined in various ways, including having one sequence subsumed by another sequence at a particular point in the flow.

A sequence is kind of like a Promise chain. However, with Promise chains, there is no “handle” to grab that references the entire chain. Whichever Promise you have a reference to only represents the current step in the chain plus any other steps hanging off it. Essentially, you cannot hold a reference to a Promise chain unless you hold a reference to the first Promise in the chain.

There are many cases where it turns out to be quite useful to have a handle that references the entire sequence collectively. The most important of those cases is with sequence `abort`/`cancel`. As we covered extensively in Chapter 3, Promises themselves should never be able to be canceled, as this violates a fundamental design imperative: external immutability.

But sequences have no such immutability design principle, mostly because sequences are not passed around as future-value containers that need immutable value semantics. So sequences are the proper level of abstraction to handle `abort`/`cancel` behavior. *asynquence* sequences can be `abort()`ed at any time, and the sequence will stop at that point and not go for any reason.

There’s plenty more reasons to prefer a sequence abstraction on top of Promise chains, for flow control purposes.

First, Promise chaining is a rather manual process – one that can get pretty tedious once you start creating and chaining Promises across a wide swath of your programs – and this tedium can act counterproductively to dissuade the developer from using Promises in places where they are quite appropriate.

Abstractions are meant to reduce boilerplate and tedium, so the sequence abstraction is a good solution to this problem. With Promises, your focus is on the individual step, and there’s little assumption that you will keep the chain going. With sequences, the opposite approach is taken, assuming the sequence will keep having more steps added indefinitely.

This abstraction complexity reduction is especially powerful when you start thinking about higher-order Promise patterns (beyond `race(..)` and `all(..)`).

For example, in the middle of a sequence, you may want to express a step that is conceptually like a `try..catch` in that the step will always result in success, either the intended main success resolution or a positive nonerror signal for the

caught error. Or, you might want to express a step that is like a retry/until loop, where it keeps trying the same step over and over until success occurs.

These sorts of abstractions are quite nontrivial to express using only Promise primitives, and doing so in the middle of an existing Promise chain is not pretty. But if you abstract your thinking to a sequence, and consider a step as a wrapper around a Promise, that step wrapper can hide such details, freeing you to think about the flow control in the most sensible way without being bothered by the details.

Second, and perhaps more importantly, thinking of async flow control in terms of steps in a sequence allows you to abstract out the details of what types of asynchronicity are involved with each individual step. Under the covers, a Promise will always control the step, but above the covers, that step can look either like a continuation callback (the simple default), or like a real Promise, or as a run-to-completion generator, or ... Hopefully, you get the picture.

Third, sequences can more easily be twisted to adapt to different modes of thinking, such as event-, stream-, or reactive-based coding. *asynquence* provides a pattern I call “reactive sequences” (which we’ll cover later) as a variation on the “reactive observable” ideas in RxJS (“Reactive Extensions”), that lets a repeatable event fire off a new sequence instance each time. Promises are one-shot-only, so it’s quite awkward to express repetitious asynchrony with Promises alone.

Another alternate mode of thinking inverts the resolution/control capability in a pattern I call “iterable sequences”. Instead of each individual step internally controlling its own completion (and thus advancement of the sequence), the sequence is inverted so the advancement control is through an external iterator, and each step in the *iterable sequence* just responds to the `next(..)` iterator control.

We’ll explore all of these different variations as we go throughout the rest of this appendix, so don’t worry if we ran over those bits far too quickly just now.

The takeaway is that sequences are a more powerful and sensible abstraction for complex asynchrony than just Promises (Promise chains) or just generators, and *asynquence* is designed to express that abstraction with just the right level of sugar to make async programming more understandable and more enjoyable.

***asynquence* API**

To start off, the way you create a sequence (an *asynquence* instance) is with the `ASQ(..)` function. An `ASQ()` call with no parameters creates an empty initial sequence, whereas passing one or more values or functions to `ASQ(..)` sets up the sequence with each argument representing the initial steps of the sequence.

Note: For the purposes of all code examples here, I will use the *asynquence* top-level identifier in global browser usage: `ASQ`. If you include and use *asynquence*

through a module system (browser or server), you of course can define whichever symbol you prefer, and *asynquence* won't care!

Many of the API methods discussed here are built into the core of *asynquence*, but others are provided through including the optional "contrib" plug-ins package. See the documentation for *asynquence* for whether a method is built in or defined via plug-in: <http://github.com/getify/asynquence>

Steps

If a function represents a normal step in the sequence, that function is invoked with the first parameter being the continuation callback, and any subsequent parameters being any messages passed on from the previous step. The step will not complete until the continuation callback is called. Once it's called, any arguments you pass to it will be sent along as messages to the next step in the sequence.

To add an additional normal step to the sequence, call `then(..)` (which has essentially the exact same semantics as the `ASQ(..)` call):

```
ASQ(
  // step 1
  function(done){
    setTimeout( function(){
      done( "Hello" );
    }, 100 );
  },
  // step 2
  function(done,greeting) {
    setTimeout( function(){
      done( greeting + " World" );
    }, 100 );
  }
)
// step 3
.then( function(done,msg){
  setTimeout( function(){
    done( msg.toUpperCase() );
  }, 100 );
} )
// step 4
.then( function(done,msg){
  console.log( msg );           // HELLO WORLD
} );
```

Note: Though the name `then(..)` is identical to the native Promises API, this `then(..)` is different. You can pass as few or as many functions or values to `then(..)` as you'd like, and each is taken as a separate step. There's no

two-callback fulfilled/rejected semantics involved.

Unlike with Promises, where to chain one Promise to the next you have to create and `return` that Promise from a `then(..)` fulfillment handler, with *asynquence*, all you need to do is call the continuation callback – I always call it `done()` but you can name it whatever suits you – and optionally pass it completion messages as arguments.

Each step defined by `then(..)` is assumed to be asynchronous. If you have a step that's synchronous, you can either just call `done(..)` right away, or you can use the simpler `val(..)` step helper:

```
// step 1 (sync)
ASQ( function(done){
    done( "Hello" );      // manually synchronous
} )
// step 2 (sync)
.val( function(greeting){
    return greeting + " World";
} )
// step 3 (async)
.then( function(done,msg){
    setTimeout( function(){
        done( msg.toUpperCase() );
    }, 100 );
} )
// step 4 (sync)
.val( function(msg){
    console.log( msg );
} );
```

As you can see, `val(..)`-invoked steps don't receive a continuation callback, as that part is assumed for you – and the parameter list is less cluttered as a result! To send a message along to the next step, you simply use `return`.

Think of `val(..)` as representing a synchronous “value-only” step, which is useful for synchronous value operations, logging, and the like.

Errors

One important difference with *asynquence* compared to Promises is with error handling.

With Promises, each individual Promise (step) in a chain can have its own independent error, and each subsequent step has the ability to handle the error or not. The main reason for this semantic comes (again) from the focus on individual Promises rather than on the chain (sequence) as a whole.

I believe that most of the time, an error in one part of a sequence is generally

not recoverable, so the subsequent steps in the sequence are moot and should be skipped. So, by default, an error at any step of a sequence throws the entire sequence into error mode, and the rest of the normal steps are ignored.

If you *do* need to have a step where its error is recoverable, there are several different API methods that can accommodate, such as `try(..)` – previously mentioned as a kind of `try..catch` step – or `until(..)` – a retry loop that keeps attempting the step until it succeeds or you manually `break()` the loop. `asynquence` even has `pThen(..)` and `pCatch(..)` methods, which work identically to how normal Promise `then(..)` and `catch(..)` work (see Chapter 3), so you can do localized mid-sequence error handling if you so choose.

The point is, you have both options, but the more common one in my experience is the default. With Promises, to get a chain of steps to ignore all steps once an error occurs, you have to take care not to register a rejection handler at any step; otherwise, that error gets swallowed as handled, and the sequence may continue (perhaps unexpectedly). This kind of desired behavior is a bit awkward to properly and reliably handle.

To register a sequence error notification handler, `asynquence` provides an `or(..)` sequence method, which also has an alias of `onerror(..)`. You can call this method anywhere in the sequence, and you can register as many handlers as you'd like. That makes it easy for multiple different consumers to listen in on a sequence to know if it failed or not; it's kind of like an error event handler in that respect.

Just like with Promises, all JS exceptions become sequence errors, or you can programmatically signal a sequence error:

```
var sq = ASQ( function(done){
  setTimeout( function(){
    // signal an error for the sequence
    done.fail( "Oops" );
  }, 100 );
} )
.then( function(done){
  // will never get here
} )
.or( function(err){
  console.log( err );           // Oops
} )
.then( function(done){
  // won't get here either
} );
// later

sq.or( function(err){
```

```

        console.log( err );           // Oops
    } );

```

Another really important difference with error handling in *asynquence* compared to native Promises is the default behavior of “unhandled exceptions”. As we discussed at length in Chapter 3, a rejected Promise without a registered rejection handler will just silently hold (aka swallow) the error; you have to remember to always end a chain with a final `catch(..)`.

In *asynquence*, the assumption is reversed.

If an error occurs on a sequence, and it **at that moment** has no error handlers registered, the error is reported to the `console`. In other words, unhandled rejections are by default always reported so as not to be swallowed and missed.

As soon as you register an error handler against a sequence, it opts that sequence out of such reporting, to prevent duplicate noise.

There may, in fact, be cases where you want to create a sequence that may go into the error state before you have a chance to register the handler. This isn’t common, but it can happen from time to time.

In those cases, you can also **opt a sequence instance out** of error reporting by calling `defer()` on the sequence. You should only opt out of error reporting if you are sure that you’re going to eventually handle such errors:

```

var sq1 = ASQ( function(done){
    doesnt.Exist();           // will throw exception to console
} );

var sq2 = ASQ( function(done){
    doesnt.Exist();           // will throw only a sequence error
} )
// opt-out of error reporting
.defer();

setTimeout( function(){
    sq1.or( function(err){
        console.log( err ); // ReferenceError
    } );
    sq2.or( function(err){
        console.log( err ); // ReferenceError
    } );
}, 100 );

// ReferenceError (from sq1)

```

This is better error handling behavior than Promises themselves have, because it’s the Pit of Success, not the Pit of Failure (see Chapter 3).

Note: If a sequence is piped into (aka subsumed by) another sequence – see “Combining Sequences” for a complete description – then the source sequence is opted out of error reporting, but now the target sequence’s error reporting or lack thereof must be considered.

Parallel Steps

Not all steps in your sequences will have just a single (async) task to perform; some will need to perform multiple steps “in parallel” (concurrently). A step in a sequence in which multiple substeps are processing concurrently is called a `gate(..)` – there’s an `all(..)` alias if you prefer – and is directly symmetric to native `Promise.all(..)`.

If all the steps in the `gate(..)` complete successfully, all success messages will be passed to the next sequence step. If any of them generate errors, the whole sequence immediately goes into an error state.

Consider:

```
ASQ( function(done){
  setTimeout( done, 100 );
} )
.gate(
  function(done){
    setTimeout( function(){
      done( "Hello" );
    }, 100 );
  },
  function(done){
    setTimeout( function(){
      done( "World", "!" );
    }, 100 );
  }
)
.val( function(msg1,msg2){
  console.log( msg1 );    // Hello
  console.log( msg2 );    // [ "World", "!" ]
} );
```

For illustration, let’s compare that example to native Promises:

```
new Promise( function(resolve,reject){
  setTimeout( resolve, 100 );
} )
.then( function(){
  return Promise.all( [
    new Promise( function(resolve,reject){
      setTimeout( function(){
```

```

        resolve( "Hello" );
    }, 100 );
} ),
new Promise( function(resolve,reject){
    setTimeout( function(){
        // note: we need a [ ] array here
        resolve( [ "World", "!" ] );
    }, 100 );
} )
] );
} )
.then( function(msgs){
    console.log( msgs[0] ); // Hello
    console.log( msgs[1] ); // [ "World", "!" ]
} );

```

Yuck. Promises require a lot more boilerplate overhead to express the same asynchronous flow control. That's a great illustration of why the *asynquence* API and abstraction make dealing with Promise steps a lot nicer. The improvement only goes higher the more complex your asynchrony is.

Step Variations There are several variations in the contrib plug-ins on *asynquence*'s `gate(..)` step type that can be quite helpful:

- `any(..)` is like `gate(..)`, except just one segment has to eventually succeed to proceed on the main sequence.
- `first(..)` is like `any(..)`, except as soon as any segment succeeds, the main sequence proceeds (ignoring subsequent results from other segments).
- `race(..)` (symmetric with `Promise.race(..)`) is like `first(..)`, except the main sequence proceeds as soon as any segment completes (either success or failure).
- `last(..)` is like `any(..)`, except only the latest segment to complete successfully sends its message(s) along to the main sequence.
- `none(..)` is the inverse of `gate(..)`: the main sequence proceeds only if all the segments fail (with all segment error message(s) transposed as success message(s) and vice versa).

Let's first define some helpers to make illustration cleaner:

```

function success1(done) {
    setTimeout( function(){
        done( 1 );
    }, 100 );
}

function success2(done) {
    setTimeout( function(){

```

```

        done( 2 );
    }, 100 );
}

function failure3(done) {
    setTimeout( function(){
        done.fail( 3 );
    }, 100 );
}

function output(msg) {
    console.log( msg );
}

Now, let's demonstrate these gate(..) step variations:

ASQ().race(
    failure3,
    success1
)
.or( output );           // 3

ASQ().any(
    success1,
    failure3,
    success2
)
.val( function(){
    var args = [].slice.call( arguments );
    console.log(
        args           // [ 1, undefined, 2 ]
    );
} );

ASQ().first(
    failure3,
    success1,
    success2
)
.val( output );           // 1

ASQ().last(
    failure3,
    success1,

```

```

        success2
)
.val( output );      // 2

ASQ().none(
    failure3
)
.val( output )      // 3
.none(
    failure3
    success1
)
.or( output );      // 1

```

Another step variation is `map(..)`, which lets you asynchronously map elements of an array to different values, and the step doesn't proceed until all the mappings are complete. `map(..)` is very similar to `gate(..)`, except it gets the initial values from an array instead of from separately specified functions, and also because you define a single function callback to operate on each value:

```

function double(x,done) {
    setTimeout( function(){
        done( x * 2 );
    }, 100 );
}

ASQ().map( [1,2,3], double )
.val( output );      // [2,4,6]

```

Also, `map(..)` can receive either of its parameters (the array or the callback) from messages passed from the previous step:

```

function plusOne(x,done) {
    setTimeout( function(){
        done( x + 1 );
    }, 100 );
}

ASQ( [1,2,3] )
.map( double )      // message ` [1,2,3] ` comes in
.map( plusOne )      // message ` [2,4,6] ` comes in
.val( output );      // [3,5,7]

```

Another variation is `waterfall(..)`, which is kind of like a mixture between `gate(..)`'s message collection behavior but `then(..)`'s sequential processing.

Step 1 is first executed, then the success message from step 1 is given to step 2, and then both success messages go to step 3, and then all three success messages

go to step 4, and so on, such that the messages sort of collect and cascade down the “waterfall”.

Consider:

```
function double(done) {
  var args = [].slice.call( arguments, 1 );
  console.log( args );

  setTimeout( function(){
    done( args[args.length - 1] * 2 );
  }, 100 );
}

ASQ( 3 )
.waterrfall(
  double,          // [ 3 ]
  double,          // [ 6 ]
  double,          // [ 6, 12 ]
  double          // [ 6, 12, 24 ]
)
.val( function(){
  var args = [].slice.call( arguments );
  console.log( args );  // [ 6, 12, 24, 48 ]
} );
```

If at any point in the “waterfall” an error occurs, the whole sequence immediately goes into an error state.

Error Tolerance Sometimes you want to manage errors at the step level and not let them necessarily send the whole sequence into the error state. *asynquence* offers two step variations for that purpose.

`try(..)` attempts a step, and if it succeeds, the sequence proceeds as normal, but if the step fails, the failure is turned into a success message formated as `{ catch: .. }` with the error message(s) filled in:

```
ASQ()
.try( success1 )
.val( output )          // 1
.try( failure3 )
.val( output )          // { catch: 3 }
.or( function(err){
  // never gets here
} );
```

You could instead set up a retry loop using `until(..)`, which tries the step and if it fails, retries the step again on the next event loop tick, and so on.

This retry loop can continue indefinitely, but if you want to break out of the loop, you can call the `break()` flag on the completion trigger, which sends the main sequence into an error state:

```
var count = 0;

ASQ( 3 )
.until( double )
.val( output )                                // 6
.until( function(done){
    count++;

    setTimeout( function(){
        if (count < 5) {
            done.fail();
        }
        else {
            // break out of the `until(..)` retry loop
            done.break( "Oops" );
        }
    }, 100 );
} )
.or( output );                                // Oops
```

Promise-Style Steps If you would prefer to have, inline in your sequence, Promise-style semantics like Promises' `then(..)` and `catch(..)` (see Chapter 3), you can use the `pThen` and `pCatch` plug-ins:

```
ASQ( 21 )
.pThen( function(msg){
    return msg * 2;
} )
.pThen( output )                                // 42
.pThen( function(){
    // throw an exception
    doesnt.Exist();
} )
.pCatch( function(err){
    // caught the exception (rejection)
    console.log( err );           // ReferenceError
} )
.val( function(){
    // main sequence is back in a
    // success state because previous
    // exception was caught by
    // `pCatch(..)`
```

```
});
```

`pThen(..)` and `pCatch(..)` are designed to run in the sequence, but behave as if it was a normal Promise chain. As such, you can either resolve genuine Promises or *asynquence* sequences from the “fulfillment” handler passed to `pThen(..)` (see Chapter 3).

Forking Sequences

One feature that can be quite useful about Promises is that you can attach multiple `then(..)` handler registrations to the same promise, effectively “forking” the flow-control at that promise:

```
var p = Promise.resolve( 21 );

// fork 1 (from `p`)
p.then( function(msg){
    return msg * 2;
} )
.then( function(msg){
    console.log( msg );      // 42
} )

// fork 2 (from `p`)
p.then( function(msg){
    console.log( msg );      // 21
} );
```

The same “forking” is easy in *asynquence* with `fork()`:

```
var sq = ASQ(..).then(..).then(..);

var sq2 = sq.fork();

// fork 1
sq.then(..)...

// fork 2
sq2.then(..)...;
```

Combining Sequences

The reverse of `fork()`ing, you can combine two sequences by subsuming one into another, using the `seq(..)` instance method:

```
var sq = ASQ( function(done){
    setTimeout( function(){
        done( "Hello World" );
    }, 200 );
```

```

} );

ASQ( function(done){
  setTimeout( done, 100 );
} )
// subsume `sq` sequence into this sequence
.seq( sq )
.val( function(msg){
  console.log( msg );      // Hello World
} )

```

`seq(..)` can either accept a sequence itself, as shown here, or a function. If a function, it's expected that the function when called will return a sequence, so the preceding code could have been done with:

```

// ...
.seq( function(){
  return sq;
} )
// ...

```

Also, that step could instead have been accomplished with a `pipe(..)`:

```

// ...
.then( function(done){
  // pipe `sq` into the `done` continuation callback
  sq.pipe( done );
} )
// ...

```

When a sequence is subsumed, both its success message stream and its error stream are piped in.

Note: As mentioned in an earlier note, piping (manually with `pipe(..)` or automatically with `seq(..)`) opts the source sequence out of error-reporting, but doesn't affect the error reporting status of the target sequence.

Value and Error Sequences

If any step of a sequence is just a normal value, that value is just mapped to that step's completion message:

```

var sq = ASQ( 42 );

sq.val( function(msg){
  console.log( msg );      // 42
} );

```

If you want to make a sequence that's automatically errored:

```

var sq = ASQ.failed( "Oops" );

ASQ()
  .seq( sq )
  .val( function(msg){
    // won't get here
  } )
  .or( function(err){
    console.log( err );      // Oops
  } );

```

You also may want to automatically create a delayed-value or a delayed-error sequence. Using the `after` and `failAfter` contrib plug-ins, this is easy:

```

var sq1 = ASQ.after( 100, "Hello", "World" );
var sq2 = ASQ.failAfter( 100, "Oops" );

sq1.val( function(msg1,msg2){
  console.log( msg1, msg2 );      // Hello World
} );

sq2.or( function(err){
  console.log( err );           // Oops
} );

```

You can also insert a delay in the middle of a sequence using `after(..)`:

```

ASQ( 42 )
// insert a delay into the sequence
  .after( 100 )
  .val( function(msg){
    console.log( msg );      // 42
  } );

```

Promises and Callbacks

I think *asynquence* sequences provide a lot of value on top of native Promises, and for the most part you'll find it more pleasant and more powerful to work at that level of abstraction. However, integrating *asynquence* with other non-*asynquence* code will be a reality.

You can easily subsume a promise (e.g., `thenable` – see Chapter 3) into a sequence using the `promise(..)` instance method:

```

var p = Promise.resolve( 42 );

ASQ()
  .promise( p )           // could also: `function(){ return p; }`
  .val( function(msg){

```

```
    console.log( msg ); // 42
} );
```

And to go the opposite direction and fork/vend a promise from a sequence at a certain step, use the `toPromise` contrib plug-in:

```
var sq = ASQ.after( 100, "Hello World" );

sq.toPromise()
// this is a standard promise chain now
.then( function(msg){
    return msg.toUpperCase();
} )
.then( function(msg){
    console.log( msg );      // HELLO WORLD
} );
```

To adapt `asynquence` to systems using callbacks, there are several helper facilities. To automatically generate an “error-first style” callback from your sequence to wire into a callback-oriented utility, use `errfcb`:

```
var sq = ASQ( function(done){
    // note: expecting "error-first style" callback
    someAsyncFuncWithCB( 1, 2, done.errfcb )
} )
.val( function(msg){
    // ...
} )
.or( function(err){
    // ...
} );

// note: expecting "error-first style" callback
anotherAsyncFuncWithCB( 1, 2, sq.errfcb() );
```

You also may want to create a sequence-wrapped version of a utility – compare to “promisory” in Chapter 3 and “thunkory” in Chapter 4 – and `asynquence` provides `ASQ.wrap(..)` for that purpose:

```
var coolUtility = ASQ.wrap( someAsyncFuncWithCB );

coolUtility( 1, 2 )
.val( function(msg){
    // ...
} )
.or( function(err){
    // ...
} );
```

Note: For the sake of clarity (and for fun!), let's coin yet another term, for a sequence-producing function that comes from `ASQ.wrap(..)`, like `coolUtility` here. I propose “sequory” (“sequence” + “factory”).

Iterable Sequences

The normal paradigm for a sequence is that each step is responsible for completing itself, which is what advances the sequence. Promises work the same way.

The unfortunate part is that sometimes you need external control over a Promise/step, which leads to awkward “capability extraction”.

Consider this Promises example:

```
var domready = new Promise( function(resolve,reject){  
  // don't want to put this here, because  
  // it belongs logically in another part  
  // of the code  
  document.addEventListener( "DOMContentLoaded", resolve );  
} );  
  
// ..  
  
domready.then( function(){  
  // DOM is ready!  
} );
```

The “capability extraction” anti-pattern with Promises looks like this:

```
var ready;  
  
var domready = new Promise( function(resolve,reject){  
  // extract the `resolve()` capability  
  ready = resolve;  
} );  
  
// ..  
  
domready.then( function(){  
  // DOM is ready!  
} );  
  
// ..  
  
document.addEventListener( "DOMContentLoaded", ready );
```

Note: This anti-pattern is an awkward code smell, in my opinion, but some developers like it, for reasons I can't grasp.

asynquence offers an inverted sequence type I call “iterable sequences”, which externalizes the control capability (it’s quite useful in use cases like the `domready`):

```
// note: `domready` here is an *iterator* that
// controls the sequence
var domready = ASQ.iterable();

// ...

domready.val( function(){
  // DOM is ready
} );

// ...

document.addEventListener( "DOMContentLoaded", domready.next );
```

There’s more to iterable sequences than what we see in this scenario. We’ll come back to them in Appendix B.

Running Generators

In Chapter 4, we derived a utility called `run(..)` which can run generators to completion, listening for `yielded` Promises and using them to `async resume` the generator. *asynquence* has just such a utility built in, called `runner(..)`.

Let’s first set up some helpers for illustration:

```
function doublePr(x) {
  return new Promise( function(resolve,reject){
    setTimeout( function(){
      resolve( x * 2 );
    }, 100 );
  } );
}

function doubleSeq(x) {
  return ASQ( function(done){
    setTimeout( function(){
      done( x * 2 )
    }, 100 );
  } );
}
```

Now, we can use `runner(..)` as a step in the middle of a sequence:

```
ASQ( 10, 11 )
.runner( function*(token){
  var x = token.messages[0] + token.messages[1];
```

```

    // yield a real promise
    x = yield doublePr( x );

    // yield a sequence
    x = yield doubleSeq( x );

    return x;
} )
.val( function(msg){
    console.log( msg );           // 84
} );

```

Wrapped Generators

You can also create a self-packaged generator – that is, a normal function that runs your specified generator and returns a sequence for its completion – by `ASQ.wrap(..)`ing it:

```

var foo = ASQ.wrap( function*(token){
    var x = token.messages[0] + token.messages[1];

    // yield a real promise
    x = yield doublePr( x );

    // yield a sequence
    x = yield doubleSeq( x );

    return x;
}, { gen: true } );

// ...

foo( 8, 9 )
.val( function(msg){
    console.log( msg );           // 68
} );

```

There's a lot more awesome that `runner(..)` is capable of, but we'll come back to that in Appendix B.

Review

asynquence is a simple abstraction – a sequence is a series of (async) steps – on top of Promises, aimed at making working with various asynchronous patterns much easier, without any compromise in capability.

There are other goodies in the *asynquence* core API and its contrib plug-ins

beyond what we saw in this appendix, but we'll leave that as an exercise for the reader to go check the rest of the capabilities out.

You've now seen the essence and spirit of *asynquence*. The key take away is that a sequence is comprised of steps, and those steps can be any of dozens of different variations on Promises, or they can be a generator-run, or... The choice is up to you, you have all the freedom to weave together whatever async flow control logic is appropriate for your tasks. No more library switching to catch different async patterns.

If these *asynquence* snippets have made sense to you, you're now pretty well up to speed on the library; it doesn't take that much to learn, actually!

If you're still a little fuzzy on how it works (or why!), you'll want to spend a little more time examining the previous examples and playing around with *asynquence* yourself, before going on to the next appendix. Appendix B will push *asynquence* into several more advanced and powerful async patterns.

You Don't Know JS: Async & Performance

Appendix B: Advanced Async Patterns

Appendix A introduced the *asynquence* library for sequence-oriented async flow control, primarily based on Promises and generators.

Now we'll explore other advanced asynchronous patterns built on top of that existing understanding and functionality, and see how *asynquence* makes those sophisticated async techniques easy to mix and match in our programs without needing lots of separate libraries.

Iterable Sequences

We introduced *asynquence*'s iterable sequences in the previous appendix, but we want to revisit them in more detail.

To refresh, recall:

```
var domready = ASQ iterable();

// ...

domready.val( function(){
  // DOM is ready
} );

// ...

document.addEventListener( "DOMContentLoaded", domready.next );
```

Now, let's define a sequence of multiple steps as an iterable sequence:

```
var steps = ASQ iterable();

steps
  .then( function STEP1(x){
    return x * 2;
  })
  .then( function STEP2(x){
    return x + 3;
  })
  .then( function STEP3(x){
    return x * 4;
  });

steps.next( 8 ).value; // 16
steps.next( 16 ).value; // 19
```

```
steps.next( 19 ).value; // 76
steps.next().done; // true
```

As you can see, an iterable sequence is a standard-compliant *iterator* (see Chapter 4). So, it can be iterated with an ES6 `for..of` loop, just like a generator (or any other *iterable*) can:

```
var steps = ASQ iterable();

steps
  .then( function STEP1(){ return 2; } )
  .then( function STEP2(){ return 4; } )
  .then( function STEP3(){ return 6; } )
  .then( function STEP4(){ return 8; } )
  .then( function STEP5(){ return 10; } );

for (var v of steps) {
  console.log( v );
}
// 2 4 6 8 10
```

Beyond the event triggering example shown in the previous appendix, iterable sequences are interesting because in essence they can be seen as a stand-in for generators or Promise chains, but with even more flexibility.

Consider a multiple Ajax request example – we've seen the same scenario in Chapters 3 and 4, both as a Promise chain and as a generator, respectively – expressed as an iterable sequence:

```
// sequence-aware ajax
var request = ASQ wrap( ajax );

ASQ( "http://some.url.1" )
  .runner(
    ASQ iterable()

    .then( function STEP1(token){
      var url = token.messages[0];
      return request( url );
    } )

    .then( function STEP2(resp){
      return ASQ().gate(
        request( "http://some.url.2/?v=" + resp ),
        request( "http://some.url.3/?v=" + resp )
      );
    } )
```

```

        .then( function STEP3(r1,r2){ return r1 + r2; } )
    )
    .val( function(msg){
        console.log( msg );
    } );
}

```

The iterable sequence expresses a sequential series of (sync or async) steps that looks awfully similar to a Promise chain – in other words, it's much cleaner looking than just plain nested callbacks, but not quite as nice as the `yield`-based sequential syntax of generators.

But we pass the iterable sequence into `ASQ#runner(...)`, which runs it to completion the same as if it was a generator. The fact that an iterable sequence behaves essentially the same as a generator is notable for a couple of reasons.

First, iterable sequences are kind of a pre-ES6 equivalent to a certain subset of ES6 generators, which means you can either author them directly (to run anywhere), or you can author ES6 generators and transpile/convert them to iterable sequences (or Promise chains for that matter!).

Thinking of an async-run-to-completion generator as just syntactic sugar for a Promise chain is an important recognition of their isomorphic relationship.

Before we move on, we should note that the previous snippet could have been expressed in *asynquence* as:

```

ASQ( "http://some.url.1" )
    .seq( /*STEP 1*/ request )
    .seq( function STEP2(resp){
        return ASQ().gate(
            request( "http://some.url.2/?v=" + resp ),
            request( "http://some.url.3/?v=" + resp )
        );
    } )
    .val( function STEP3(r1,r2){ return r1 + r2; } )
    .val( function(msg){
        console.log( msg );
    } );
}

```

Moreover, step 2 could have even been expressed as:

```

.gate(
    function STEP2a(done,resp) {
        request( "http://some.url.2/?v=" + resp )
        .pipe( done );
    },
    function STEP2b(done,resp) {
        request( "http://some.url.3/?v=" + resp )
        .pipe( done );
    }
);

```

```
    }  
})
```

So, why would we go to the trouble of expressing our flow control as an iterable sequence in a `ASQ#runner(..)` step, when it seems like a simpler/flatter *asyquence* chain does the job well?

Because the iterable sequence form has an important trick up its sleeve that gives us more capability. Read on.

Extending Iterable Sequences

Generators, normal *asyquence* sequences, and Promise chains, are all **eagerly evaluated** – whatever flow control is expressed initially *is* the fixed flow that will be followed.

However, iterable sequences are **lazily evaluated**, which means that during execution of the iterable sequence, you can extend the sequence with more steps if desired.

Note: You can only append to the end of an iterable sequence, not inject into the middle of the sequence.

Let's first look at a simpler (synchronous) example of that capability to get familiar with it:

```
function double(x) {  
  x *= 2;  
  
  // should we keep extending?  
  if (x < 500) {  
    isq.then(double);  
  }  
  
  return x;  
}  
  
// setup single-step iterable sequence  
var isq = ASQ iterable().then(double);  
  
for (var v = 10, ret;  
  (ret = isq.next(v)) && !ret.done;  
) {  
  v = ret.value;  
  console.log(v);  
}
```

The iterable sequence starts out with only one defined step (`isq.then(double)`), but the sequence keeps extending itself under certain conditions (`x < 500`).

Both *asynquence* sequences and Promise chains technically *can* do something similar, but we'll see in a little bit why their capability is insufficient.

Though this example is rather trivial and could otherwise be expressed with a `while` loop in a generator, we'll consider more sophisticated cases.

For instance, you could examine the response from an Ajax request and if it indicates that more data is needed, you conditionally insert more steps into the iterable sequence to make the additional request(s). Or you could conditionally add a value-formatting step to the end of your Ajax handling.

Consider:

```
var steps = ASQ iterable()

.then( function STEP1(token){
  var url = token.messages[0].url;

  // was an additional formatting step provided?
  if (token.messages[0].format) {
    steps.then( token.messages[0].format );
  }

  return request( url );
} )

.then( function STEP2(resp){
  // add another Ajax request to the sequence?
  if (/x1/.test( resp )) {
    steps.then( function STEP5(text){
      return request(
        "http://some.url.4/?v=" + text
      );
    } );
  }
}

return ASQ().gate(
  request( "http://some.url.2/?v=" + resp ),
  request( "http://some.url.3/?v=" + resp )
);
} )

.then( function STEP3(r1,r2){ return r1 + r2; } );
```

You can see in two different places where we conditionally extend `steps` with `steps.then(..)`. And to run this `steps` iterable sequence, we just wire it into our main program flow with an *asynquence* sequence (called `main` here) using `ASQ#runner(..)`:

```

var main = ASQ( {
  url: "http://some.url.1",
  format: function STEP4(text){
    return text.toUpperCase();
  }
} )
.runner( steps )
.val( function(msg){
  console.log( msg );
} );

```

Can the flexibility (conditional behavior) of the `steps` iterable sequence be expressed with a generator? Kind of, but we have to rearrange the logic in a slightly awkward way:

```

function *steps(token) {
  // **STEP 1**
  var resp = yield request( token.messages[0].url );

  // **STEP 2**
  var rvals = yield ASQ().gate(
    request( "http://some.url.2/?v=" + resp ),
    request( "http://some.url.3/?v=" + resp )
  );

  // **STEP 3**
  var text = rvals[0] + rvals[1];

  // **STEP 4**
  // was an additional formatting step provided?
  if (token.messages[0].format) {
    text = yield token.messages[0].format( text );
  }

  // **STEP 5**
  // need another Ajax request added to the sequence?
  if (/foobar/.test( resp )) {
    text = yield request(
      "http://some.url.4/?v=" + text
    );
  }

  return text;
}

// note: `*steps()` can be run by the same `ASQ` sequence
// as `steps` was previously

```

Setting aside the already identified benefits of the sequential, synchronous-looking syntax of generators (see Chapter 4), the `steps` logic had to be reordered in the `*steps()` generator form, to fake the dynamicism of the extendable iterable sequence `steps`.

What about expressing the functionality with Promises or sequences, though? You *can* do something like this:

```
var steps = something( ... )
.then( ... )
.then( function(..){
  // ...

  // extending the chain, right?
  steps = steps.then( ... );

  ...
})

.then( ... );
```

The problem is subtle but important to grasp. So, consider trying to wire up our `steps` Promise chain into our main program flow – this time expressed with Promises instead of *asynquence*:

```
var main = Promise.resolve( {
  url: "http://some.url.1",
  format: function STEP4(text){
    return text.toUpperCase();
  }
})
.then( function(..){
  return steps;           // hint!
})
.val( function(msg){
  console.log( msg );
});
```

Can you spot the problem now? Look closely!

There's a race condition for sequence steps ordering. When you `return steps`, at that moment `steps` *might* be the originally defined promise chain, or it might now point to the extended promise chain via the `steps = steps.then(..)` call, depending on what order things happen.

Here are the two possible outcomes:

- If `steps` is still the original promise chain, once it's later “extended” by `steps = steps.then(..)`, that extended promise on the end of the chain is **not** considered by the `main` flow, as it's already tapped the `steps` chain. This is the unfortunately limiting **eager evaluation**.

- If `steps` is already the extended promise chain, it works as we expect in that the extended promise is what `main` taps.

Other than the obvious fact that a race condition is intolerable, the first case is the concern; it illustrates **eager evaluation** of the promise chain. By contrast, we easily extended the iterable sequence without such issues, because iterable sequences are **lazily evaluated**.

The more dynamic you need your flow control, the more iterable sequences will shine.

Tip: Check out more information and examples of iterable sequences on the *asynquence* site (<https://github.com/getify/asynquence/blob/master/README.md#iterable-sequences>).

Event Reactive

It should be obvious from (at least!) Chapter 3 that Promises are a very powerful tool in your `async` toolbox. But one thing that's clearly lacking is in their capability to handle streams of events, as a Promise can only be resolved once. And frankly, this exact same weakness is true of plain *asynquence* sequences, as well.

Consider a scenario where you want to fire off a series of steps every time a certain event is fired. A single Promise or sequence cannot represent all occurrences of that event. So, you have to create a whole new Promise chain (or sequence) for *each* event occurrence, such as:

```
listener.on( "foobar", function(data){

  // create a new event handling promise chain
  new Promise( function(resolve,reject){
    // ...
  } )
  .then( ... )
  .then( ... );

} );
```

The base functionality we need is present in this approach, but it's far from a desirable way to express our intended logic. There are two separate capabilities conflated in this paradigm: the event listening, and responding to the event; separation of concerns would implore us to separate out these capabilities.

The carefully observant reader will see this problem as somewhat symmetrical to the problems we detailed with callbacks in Chapter 2; it's kind of an inversion of control problem.

Imagine uninverting this paradigm, like so:

```

var observable = listener.on( "foobar" );

// later
observable
.then( ... )
.then( ... );

// elsewhere
observable
.then( ... )
.then( ... );

```

The `observable` value is not exactly a Promise, but you can *observe* it much like you can observe a Promise, so it's closely related. In fact, it can be observed many times, and it will send out notifications every time its event ("foobar") occurs.

Tip: This pattern I've just illustrated is a **massive simplification** of the concepts and motivations behind reactive programming (aka RP), which has been implemented/expounded upon by several great projects and languages. A variation on RP is functional reactive programming (FRP), which refers to applying functional programming techniques (immutability, referential integrity, etc.) to streams of data. "Reactive" refers to spreading this functionality out over time in response to events. The interested reader should consider studying "Reactive Observables" in the fantastic "Reactive Extensions" library ("RxJS" for JavaScript) by Microsoft (<http://rxjs.codeplex.com/>); it's much more sophisticated and powerful than I've just shown. Also, Andre Staltz has an excellent write-up (<https://gist.github.com/staltz/868e7e9bc2a7b8c1f754>) that pragmatically lays out RP in concrete examples.

ES7 Observables

At the time of this writing, there's an early ES7 proposal for a new data type called "Observable" (<https://github.com/jhusain/asyncgenerator#introducing-observable>), which in spirit is similar to what we've laid out here, but is definitely more sophisticated.

The notion of this kind of Observable is that the way you "subscribe" to the events from a stream is to pass in a generator – actually the *iterator* is the interested party – whose `next(..)` method will be called for each event.

You could imagine it sort of like this:

```

// `someEventStream` is a stream of events, like from
// mouse clicks, and the like.

var observer = new Observer( someEventStream, function*(){
  while (var evt = yield) {

```

```

        console.log( evt );
    }
} );

```

The generator you pass in will `yield` pause the `while` loop waiting for the next event. The *iterator* attached to the generator instance will have its `next(..)` called each time `someEventStream` has a new event published, and so that event data will resume your generator/*iterator* with the `evt` data.

In the subscription to events functionality here, it's the *iterator* part that matters, not the generator. So conceptually you could pass in practically any iterable, including `ASQ iterable()` iterable sequences.

Interestingly, there are also proposed adapters to make it easy to construct Observables from certain types of streams, such as `fromEvent(..)` for DOM events. If you look at a suggested implementation of `fromEvent(..)` in the earlier linked ES7 proposal, it looks an awful lot like the `ASQ.react(..)` we'll see in the next section.

Of course, these are all early proposals, so what shakes out may very well look/behave differently than shown here. But it's exciting to see the early alignments of concepts across different libraries and language proposals!

Reactive Sequences

With that crazy brief summary of Observables (and F/RP) as our inspiration and motivation, I will now illustrate an adaptation of a small subset of "Reactive Observables," which I call "Reactive Sequences."

First, let's start with how to create an Observable, using an *asynquence* plug-in utility called `react(..)`:

```

var observable = ASQ.react( function setup(next){
  listener.on( "foobar", next );
} );

```

Now, let's see how to define a sequence that "reacts" – in F/RP, this is typically called "subscribing" – to that `observable`:

```

observable
  .seq( .. )
  .then( .. )
  .val( .. );

```

So, you just define the sequence by chaining off the Observable. That's easy, huh?

In F/RP, the stream of events typically channels through a set of functional transforms, like `scan(..)`, `map(..)`, `reduce(..)`, and so on. With reactive sequences, each event channels through a new instance of the sequence. Let's look at a more concrete example:

```

ASQ.react( function setup(next){
  document.getElementById( "mybtn" )
    .addEventListener( "click", next, false );
} )
  .seq( function(evt){
    var btnID = evt.target.id;
    return request(
      "http://some.url.1/?id=" + btnID
    );
} )
  .val( function(text){
    console.log( text );
} );

```

The “reactive” portion of the reactive sequence comes from assigning one or more event handlers to invoke the event trigger (calling `next(..)`).

The “sequence” portion of the reactive sequence is exactly like the sequences we’ve already explored: each step can be whatever asynchronous technique makes sense, from continuation callback to Promise to generator.

Once you set up a reactive sequence, it will continue to initiate instances of the sequence as long as the events keep firing. If you want to stop a reactive sequence, you can call `stop()`.

If a reactive sequence is `stop()`’d, you likely want the event handler(s) to be unregistered as well; you can register a teardown handler for this purpose:

```

var sq = ASQ.react( function setup(next,registerTeardown){
  var btn = document.getElementById( "mybtn" );

  btn.addEventListener( "click", next, false );

  // will be called once `sq.stop()` is called
  registerTeardown( function(){
    btn.removeEventListener( "click", next, false );
  } );
} )
  .seq( .. )
  .then( .. )
  .val( .. );

// later
sq.stop();

```

Note: The `this` binding reference inside the `setup(..)` handler is the same `sq` reactive sequence, so you can use the `this` reference to add to the reactive sequence definition, call methods like `stop()`, and so on.

Here's an example from the Node.js world, using reactive sequences to handle incoming HTTP requests:

```
var server = http.createServer();
server.listen(8000);

// reactive observer
var request = ASQ.react( function setup(next,registerTeardown){
    server.addListener( "request", next );
    server.addListener( "close", this.stop );

    registerTeardown( function(){
        server.removeListener( "request", next );
        server.removeListener( "close", request.stop );
    } );
});

// respond to requests
request
.seq( pullFromDatabase )
.val( function(data,res){
    res.end( data );
} );

// node teardown
process.on( "SIGINT", request.stop );
```

The `next(..)` trigger can also adapt to node streams easily, using `onStream(..)` and `unStream(..)`:

```
ASQ.react( function setup(next){
    var fstream = fs.createReadStream( "/some/file" );

    // pipe the stream's "data" event to `next(..)`
    next.onStream( fstream );

    // listen for the end of the stream
    fstream.on( "end", function(){
        next.unStream( fstream );
    } );
};

.seq( .. )
.then( .. )
.val( .. );
```

You can also use sequence combinations to compose multiple reactive sequence streams:

```

var sq1 = ASQ.react(..).seq(..).then(..);
var sq2 = ASQ.react(..).seq(..).then(..);

var sq3 = ASQ.react(..)
  .gate(
    sq1,
    sq2
  )
  .then(..);

```

The main takeaway is that `ASQ.react(..)` is a lightweight adaptation of F/RP concepts, enabling the wiring of an event stream to a sequence, hence the term “reactive sequence.” Reactive sequences are generally capable enough for basic reactive uses.

Note: Here’s an example of using `ASQ.react(..)` in managing UI state (<http://jsbin.com/rozipaki/6/edit?js,output>), and another example of handling HTTP request/response streams with `ASQ.react(..)` (<https://gist.github.com/getify/bba5ec0de9d6047b720e>).

Generator Coroutine

Hopefully Chapter 4 helped you get pretty familiar with ES6 generators. In particular, we want to revisit the “Generator Concurrency” discussion, and push it even further.

We imagined a `runAll(..)` utility that could take two or more generators and run them concurrently, letting them cooperatively `yield` control from one to the next, with optional message passing.

In addition to being able to run a single generator to completion, the `ASQ#runner(..)` we discussed in Appendix A is a similar implementation of the concepts of `runAll(..)`, which can run multiple generators concurrently to completion.

So let’s see how we can implement the concurrent Ajax scenario from Chapter 4:

```

ASQ(
  "http://some.url.2"
)
.runner(
  function*(token){
    // transfer control
    yield token;

    var url1 = token.messages[0]; // "http://some.url.1"
    // clear out messages to start fresh
  }
)

```

```

    token.messages = [];

    var p1 = request( url1 );

    // transfer control
    yield token;

    token.messages.push( yield p1 );
},
function*(token){
    var url2 = token.messages[0]; // "http://some.url.2"

    // message pass and transfer control
    token.messages[0] = "http://some.url.1";
    yield token;

    var p2 = request( url2 );

    // transfer control
    yield token;

    token.messages.push( yield p2 );

    // pass along results to next sequence step
    return token.messages;
}
)
.val( function(res){
    // `res[0]` comes from "http://some.url.1"
    // `res[1]` comes from "http://some.url.2"
} );

```

The main differences between `ASQ#runner(..)` and `runAll(..)` are as follows:

- Each generator (coroutine) is provided an argument we call `token`, which is the special value to `yield` when you want to explicitly transfer control to the next coroutine.
- `token.messages` is an array that holds any messages passed in from the previous sequence step. It's also a data structure that you can use to share messages between coroutines.
- yielding a Promise (or sequence) value does not transfer control, but instead pauses the coroutine processing until that value is ready.
- The last `returned` or `yielded` value from the coroutine processing run will be forward passed to the next step in the sequence.

It's also easy to layer helpers on top of the base `ASQ#runner(..)` functionality to suit different uses.

State Machines

One example that may be familiar to many programmers is state machines. You can, with the help of a simple cosmetic utility, create an easy-to-express state machine processor.

Let's imagine such a utility. We'll call it `state(..)`, and will pass it two arguments: a state value and a generator that handles that state. `state(..)` will do the dirty work of creating and returning an adapter generator to pass to `ASQ#runner(..)`.

Consider:

```
function state(val,handler) {
  // make a coroutine handler for this state
  return function*(token) {
    // state transition handler
    function transition(to) {
      token.messages[0] = to;
    }

    // set initial state (if none set yet)
    if (token.messages.length < 1) {
      token.messages[0] = val;
    }

    // keep going until final state (false) is reached
    while (token.messages[0] !== false) {
      // current state matches this handler?
      if (token.messages[0] === val) {
        // delegate to state handler
        yield *handler( transition );
      }

      // transfer control to another state handler?
      if (token.messages[0] !== false) {
        yield token;
      }
    }
  };
}
```

If you look closely, you'll see that `state(..)` returns back a generator that accepts a `token`, and then it sets up a `while` loop that will run until the state machine reaches its final state (which we arbitrarily pick as the `false` value); that's exactly the kind of generator we want to pass to `ASQ#runner(..)`!

We also arbitrarily reserve the `token.messages[0]` slot as the place where the

current state of our state machine will be tracked, which means we can even seed the initial state as the value passed in from the previous step in the sequence.

How do we use the `state(...)` helper along with `ASQ#runner(...)`?

```
var prevState;

ASQ(
  /* optional: initial state value */
  2
)
// run our state machine
// transitions: 2 -> 3 -> 1 -> 3 -> false
.runner(
  // state `1` handler
  state( 1, function *stateOne(transition){
    console.log( "in state 1" );

    prevState = 1;
    yield transition( 3 ); // goto state `3`
  } ),

  // state `2` handler
  state( 2, function *stateTwo(transition){
    console.log( "in state 2" );

    prevState = 2;
    yield transition( 3 ); // goto state `3`
  } ),

  // state `3` handler
  state( 3, function *stateThree(transition){
    console.log( "in state 3" );

    if (prevState === 2) {
      prevState = 3;
      yield transition( 1 ); // goto state `1`
    }
    // all done!
    else {
      yield "That's all folks!";

      prevState = 3;
      yield transition( false ); // terminal state
    }
  } )
)
```

```
// state machine complete, so move on
.val( function(msg){
  console.log( msg ); // That's all folks!
} );
```

It's important to note that the `*stateOne(..)`, `*stateTwo(..)`, and `*stateThree(..)` generators themselves are reinvoked each time that state is entered, and they finish when you `transition(..)` to another value. While not shown here, of course these state generator handlers can be asynchronously paused by `yielding` Promises/sequences/thunks.

The underneath hidden generators produced by the `state(..)` helper and actually passed to `ASQ#runner(..)` are the ones that continue to run concurrently for the length of the state machine, and each of them handles cooperatively `yielding` control to the next, and so on.

Note: See this “ping pong” example (<http://jsbin.com/qutabu/1/edit?js,output>) for more illustration of using cooperative concurrency with generators driven by `ASQ#runner(..)`.

Communicating Sequential Processes (CSP)

“Communicating Sequential Processes” (CSP) was first described by C. A. R. Hoare in a 1978 academic paper (<http://dl.acm.org/citation.cfm?doid=359576.359585>), and later in a 1985 book (<http://www.usingcsp.com/>) of the same name. CSP describes a formal method for concurrent “processes” to interact (aka “communicate”) during processing.

You may recall that we examined concurrent “processes” back in Chapter 1, so our exploration of CSP here will build upon that understanding.

Like most great concepts in computer science, CSP is heavily steeped in academic formalism, expressed as a process algebra. However, I suspect symbolic algebra theorems won’t make much practical difference to the reader, so we will want to find some other way of wrapping our brains around CSP.

I will leave much of the formal description and proof of CSP to Hoare’s writing, and to many other fantastic writings since. Instead, we will try to just briefly explain the idea of CSP in as un-academic and hopefully intuitively understandable a way as possible.

Message Passing

The core principle in CSP is that all communication/interaction between otherwise independent processes must be through formal message passing. Perhaps counter to your expectations, CSP message passing is described as a synchronous action, where the sender process and the receiver process have to mutually be ready for the message to be passed.

How could such synchronous messaging possibly be related to asynchronous programming in JavaScript?

The concreteness of relationship comes from the nature of how ES6 generators are used to produce synchronous-looking actions that under the covers can indeed either be synchronous or (more likely) asynchronous.

In other words, two or more concurrently running generators can appear to synchronously message each other while preserving the fundamental asynchrony of the system because each generator's code is paused (aka "blocked") waiting on resumption of an asynchronous action.

How does this work?

Imagine a generator (aka "process") called "A" that wants to send a message to generator "B." First, "A" `yields` the message (thus pausing "A") to be sent to "B." When "B" is ready and takes the message, "A" is then resumed (unblocked).

Symmetrically, imagine a generator "A" that wants a message `from` "B." "A" `yields` its request (thus pausing "A") for the message from "B," and once "B" sends a message, "A" takes the message and is resumed.

One of the more popular expressions of this CSP message passing theory comes from ClojureScript's `core.async` library, and also from the *go* language. These takes on CSP embody the described communication semantics in a conduit that is opened between processes called a "channel."

Note: The term *channel* is used in part because there are modes in which more than one value can be sent at once into the "buffer" of the channel; this is similar to what you may think of as a stream. We won't go into depth about it here, but it can be a very powerful technique for managing streams of data.

In the simplest notion of CSP, a channel that we create between "A" and "B" would have a method called `take(..)` for blocking to receive a value, and a method called `put(..)` for blocking to send a value.

This might look like:

```
var ch = channel();

function *foo() {
  var msg = yield take( ch );

  console.log( msg );
}

function *bar() {
  yield put( ch, "Hello World" );

  console.log( "message sent" );
}
```

```

run( foo );
run( bar );
// Hello World
// "message sent"

```

Compare this structured, synchronous(-looking) message passing interaction to the informal and unstructured message sharing that `ASQ#runner(..)` provides through the `token.messages` array and cooperative yielding. In essence, `yield put(..)` is a single operation that both sends the value and pauses execution to transfer control, whereas in earlier examples we did those as separate steps.

Moreover, CSP stresses that you don't really explicitly "transfer control," but rather you design your concurrent routines to block expecting either a value received from the channel, or to block expecting to try to send a message on the channel. The blocking around receiving or sending messages is how you coordinate sequencing of behavior between the coroutines.

Note: Fair warning: this pattern is very powerful but it's also a little mind twisting to get used to at first. You will want to practice this a bit to get used to this new way of thinking about coordinating your concurrency.

There are several great libraries that have implemented this flavor of CSP in JavaScript, most notably "js-csp" (<https://github.com/ubolonton/js-csp>), which James Long (<http://twitter.com/jlongster>) forked (<https://github.com/jlongster/js-csp>) and has written extensively about (<http://jlongster.com/Taming-the-Asynchronous-Beast-with-CSP-in-JavaScript>). Also, it cannot be stressed enough how amazing the many writings of David Nolen (<http://twitter.com/swannodette>) are on the topic of adapting ClojureScript's go-style core.async CSP into JS generators (<http://swannodette.github.io/2013/08/24/es6-generators-and-csp>).

asynquence CSP emulation

Because we've been discussing async patterns here in the context of my *asynquence* library, you might be interested to see that we can fairly easily add an emulation layer on top of `ASQ#runner(..)` generator handling as a nearly perfect porting of the CSP API and behavior. This emulation layer ships as an optional part of the "asynquence-contrib" package alongside *asynquence*.

Very similar to the `state(..)` helper from earlier, `ASQ.csp.go(..)` takes a generator – in go/core.async terms, it's known as a goroutine – and adapts it to use with `ASQ#runner(..)` by returning a new generator.

Instead of being passed a `token`, your goroutine receives an initially created channel (`ch` below) that all goroutines in this run will share. You can create more channels (which is often quite helpful!) with `ASQ.csp.chan(..)`.

In CSP, we model all asynchrony in terms of blocking on channel messages, rather than blocking waiting for a Promise/sequence/thunk to complete.

So, instead of `yielding` the Promise returned from `request(..)`, `request(..)` should return a channel that you `take(..)` a value from. In other words, a single-value channel is roughly equivalent in this context/usage to a Promise/sequence.

Let's first make a channel-aware version of `request(..)`:

```
function request(url) {
  var ch = ASQ.csp.channel();
  ajax(url).then(function(content){
    // `putAsync(..)` is a version of `put(..)` that
    // can be used outside of a generator. It returns
    // a promise for the operation's completion. We
    // don't use that promise here, but we could if
    // we needed to be notified when the value had
    // been `take(..)`n.
    ASQ.csp.putAsync(ch, content);
  });
  return ch;
}
```

From Chapter 3, “promisory” is a Promise-producing utility, “thunkory” from Chapter 4 is a thunk-producing utility, and finally, in Appendix A we invented “sequory” for a sequence-producing utility.

Naturally, we need to coin a symmetric term here for a channel-producing utility. So let's unsurprisingly call it a “chanory” (“channel” + “factory”). As an exercise for the reader, try your hand at defining a `channelify(..)` utility similar to `Promise.wrap(..)/promisify(..)` (Chapter 3), `thunkify(..)` (Chapter 4), and `ASQ.wrap(..)` (Appendix A).

Now consider the concurrent Ajax example using *asyquence*-flavored CSP:

```
ASQ()
.runner(
  ASQ.csp.go(function*(ch){
    yield ASQ.csp.put(ch, "http://some.url.2");

    var url1 = yield ASQ.csp.take(ch);
    // "http://some.url.1"

    var res1 = yield ASQ.csp.take(request(url1));
    yield ASQ.csp.put(ch, res1);
  }),
  ASQ.csp.go(function*(ch){
    var url2 = yield ASQ.csp.take(ch);
    // "http://some.url.2"
  })
)
```

```

yield ASQ.csp.put( ch, "http://some.url.1" );

var res2 = yield ASQ.csp.take( request( url2 ) );
var res1 = yield ASQ.csp.take( ch );

// pass along results to next sequence step
ch.buffer_size = 2;
ASQ.csp.put( ch, res1 );
ASQ.csp.put( ch, res2 );
} )
)
.val( function(res1,res2){
  // `res1` comes from "http://some.url.1"
  // `res2` comes from "http://some.url.2"
} );

```

The message passing that trades the URL strings between the two goroutines is pretty straightforward. The first goroutine makes an Ajax request to the first URL, and that response is put onto the `ch` channel. The second goroutine makes an Ajax request to the second URL, then gets the first response `res1` off the `ch` channel. At that point, both responses `res1` and `res2` are completed and ready.

If there are any remaining values in the `ch` channel at the end of the goroutine run, they will be passed along to the next step in the sequence. So, to pass out message(s) from the final goroutine, `put(..)` them into `ch`. As shown, to avoid the blocking of those final `put(..)`s, we switch `ch` into buffering mode by setting its `buffer_size` to 2 (default: 0).

Note: See many more examples of using *asynquence*-flavored CSP here (<https://gist.github.com/getify/e0d04f1f5aa24b1947ae>).

Review

Promises and generators provide the foundational building blocks upon which we can build much more sophisticated and capable asynchrony.

asynquence has utilities for implementing *iterable sequences*, *reactive sequences* (aka “Observables”), *concurrent coroutines*, and even *CSP goroutines*.

Those patterns, combined with the continuation-callback and Promise capabilities, gives *asynquence* a powerful mix of different asynchronous functionalities, all integrated in one clean async flow control abstraction: the sequence.

You Don't Know JS: ES6 & Beyond

Foreword

Kyle Simpson is a thorough pragmatist.

I can't think of higher praise than this. To me, these are two of the most important qualities that a software developer must have. That's right: *must*, not *should*. Kyle's keen ability to tease apart layers of the JavaScript programming language and present them in understandable and meaningful portions is second to none.

ES6 & Beyond will be familiar to readers of the *You Don't Know JS* series: they can expect to be deeply immersed in everything from the obvious, to the very subtle – revealing semantics that were either taken for granted or never even considered. Until now, the *You Don't Know JS* book series has covered material that has at least some degree of familiarity to its readers. They have either seen or heard about the subject matter; they may even have experience with it. This volume covers material that only a very small portion of the JavaScript developer community has been exposed to: the evolutionary changes to the language introduced in the ECMAScript 2015 Language Specification.

Over the last couple years, I've witnessed Kyle's tireless efforts to familiarize himself with this material to a level of expertise that is rivaled by only a handful of his professional peers. That's quite a feat, considering that at the time of this writing, the language specification document hasn't been formally published! But what I've said is true, and I've read every word that Kyle's written for this book. I've followed every change, and each time, the content only gets better and provides yet a deeper level of understanding.

This book is about shaking up your sense of understanding by exposing you to the new and unknown. The intention is to evolve your knowledge in step with your tools by bestowing you with new capabilities. It exists to give you the confidence to fully embrace the next major era of JavaScript programming.

Rick Waldron @rwaldron Open Web Engineer at Bocoup Ecma/TC39 Representative for jQuery

You Don't Know JS

Preface

I'm sure you noticed, but "JS" in the book series title is not an abbreviation for words used to curse about JavaScript, though cursing at the language's quirks is something we can probably all identify with!

From the earliest days of the web, JavaScript has been a foundational technology that drives interactive experience around the content we consume. While flickering mouse trails and annoying pop-up prompts may be where JavaScript started, nearly 2 decades later, the technology and capability of JavaScript has grown many orders of magnitude, and few doubt its importance at the heart of the world's most widely available software platform: the web.

But as a language, it has perpetually been a target for a great deal of criticism, owing partly to its heritage but even more to its design philosophy. Even the name evokes, as Brendan Eich once put it, "dumb kid brother" status next to its more mature older brother "Java". But the name is merely an accident of politics and marketing. The two languages are vastly different in many important ways. "JavaScript" is as related to "Java" as "Carnival" is to "Car".

Because JavaScript borrows concepts and syntax idioms from several languages, including proud C-style procedural roots as well as subtle, less obvious Scheme/Lisp-style functional roots, it is exceedingly approachable to a broad audience of developers, even those with just little to no programming experience. The "Hello World" of JavaScript is so simple that the language is inviting and easy to get comfortable with in early exposure.

While JavaScript is perhaps one of the easiest languages to get up and running with, its eccentricities make solid mastery of the language a vastly less common occurrence than in many other languages. Where it takes a pretty in-depth knowledge of a language like C or C++ to write a full-scale program, full-scale production JavaScript can, and often does, barely scratch the surface of what the language can do.

Sophisticated concepts which are deeply rooted into the language tend instead to surface themselves in *seemingly* simplistic ways, such as passing around functions as callbacks, which encourages the JavaScript developer to just use the language as-is and not worry too much about what's going on under the hood.

It is simultaneously a simple, easy-to-use language that has broad appeal, and a complex and nuanced collection of language mechanics which without careful study will elude *true understanding* even for the most seasoned of JavaScript developers.

Therein lies the paradox of JavaScript, the Achilles' Heel of the language, the challenge we are presently addressing. Because JavaScript *can* be used without understanding, the understanding of the language is often never attained.

Mission

If at every point that you encounter a surprise or frustration in JavaScript, your response is to add it to the blacklist, as some are accustomed to doing, you soon will be relegated to a hollow shell of the richness of JavaScript.

While this subset has been famously dubbed “The Good Parts”, I would implore you, dear reader, to instead consider it the “The Easy Parts”, “The Safe Parts”, or even “The Incomplete Parts”.

This *You Don’t Know JavaScript* book series offers a contrary challenge: learn and deeply understand *all* of JavaScript, even and especially “The Tough Parts”.

Here, we address head on the tendency of JS developers to learn “just enough” to get by, without ever forcing themselves to learn exactly how and why the language behaves the way it does. Furthermore, we eschew the common advice to *retreat* when the road gets rough.

I am not content, nor should you be, at stopping once something *just works*, and not really knowing *why*. I gently challenge you to journey down that bumpy “road less traveled” and embrace all that JavaScript is and can do. With that knowledge, no technique, no framework, no popular buzzword acronym of the week, will be beyond your understanding.

These books each take on specific core parts of the language which are most commonly misunderstood or under-understood, and dive very deep and exhaustively into them. You should come away from reading with a firm confidence in your understanding, not just of the theoretical, but the practical “what you need to know” bits.

The JavaScript you know *right now* is probably *parts* handed down to you by others who’ve been burned by incomplete understanding. *That* JavaScript is but a shadow of the true language. You don’t *really* know JavaScript, *yet*, but if you dig into this series, you *will*. Read on, my friends. JavaScript awaits you.

Summary

JavaScript is awesome. It’s easy to learn partially, and much harder to learn completely (or even *sufficiently*). When developers encounter confusion, they usually blame the language instead of their lack of understanding. These books aim to fix that, inspiring a strong appreciation for the language you can now, and *should*, deeply *know*.

Note: Many of the examples in this book assume modern (and future-reaching) JavaScript engine environments, such as ES6. Some code may not work as described if run in older (pre-ES6) engines.

You Don't Know JS: ES6 & Beyond

Chapter 1: ES? Now & Future

Before you dive into this book, you should have a solid working proficiency over JavaScript up to the most recent standard (at the time of this writing), which is commonly called *ES5* (technically ES 5.1). Here, we plan to talk squarely about the upcoming *ES6*, as well as cast our vision beyond to understand how JS will evolve moving forward.

If you are still looking for confidence with JavaScript, I highly recommend you read the other titles in this series first:

- *Up & Going*: Are you new to programming and JS? This is the roadmap you need to consult as you start your learning journey.
- *Scope & Closures*: Did you know that JS lexical scope is based on compiler (not interpreter!) semantics? Can you explain how closures are a direct result of lexical scope and functions as values?
- *this & Object Prototypes*: Can you recite the four simple rules for how `this` is bound? Have you been muddling through fake “classes” in JS instead of adopting the simpler “behavior delegation” design pattern? Ever heard of *objects linked to other objects* (OLOO)?
- *Types & Grammar*: Do you know the built-in types in JS, and more importantly, do you know how to properly and safely use coercion between types? How comfortable are you with the nuances of JS grammar/syntax?
- *Async & Performance*: Are you still using callbacks to manage your asynchrony? Can you explain what a promise is and why/how it solves “callback hell”? Do you know how to use generators to improve the legibility of async code? What exactly constitutes mature optimization of JS programs and individual operations?

If you've already read all those titles and you feel pretty comfortable with the topics they cover, it's time we dive into the evolution of JS to explore all the changes coming not only soon but farther over the horizon.

Unlike ES5, ES6 is not just a modest set of new APIs added to the language. It incorporates a whole slew of new syntactic forms, some of which may take quite a bit of getting used to. There's also a variety of new organization forms and new API helpers for various data types.

ES6 is a radical jump forward for the language. Even if you think you know JS in ES5, ES6 is full of new stuff you *don't know yet*, so get ready! This book explores all the major themes of ES6 that you need to get up to speed on, and even gives you a glimpse of future features coming down the track that you should be aware of.

Warning: All code in this book assumes an ES6+ environment. At the time of this writing, ES6 support varies quite a bit in browsers and JS environments

(like Node.js), so your mileage may vary.

Versioning

The JavaScript standard is referred to officially as “ECMAScript” (abbreviated “ES”), and up until just recently has been versioned entirely by ordinal number (i.e., “5” for “5th edition”).

The earliest versions, ES1 and ES2, were not widely known or implemented. ES3 was the first widespread baseline for JavaScript, and constitutes the JavaScript standard for browsers like IE6-8 and older Android 2.x mobile browsers. For political reasons beyond what we’ll cover here, the ill-fated ES4 never came about.

In 2009, ES5 was officially finalized (later ES5.1 in 2011), and settled as the widespread standard for JS for the modern revolution and explosion of browsers, such as Firefox, Chrome, Opera, Safari, and many others.

Leading up to the expected *next* version of JS (slipped from 2013 to 2014 and then 2015), the obvious and common label in discourse has been ES6.

However, late into the ES6 specification timeline, suggestions have surfaced that versioning may in the future switch to a year-based schema, such as ES2016 (aka ES7) to refer to whatever version of the specification is finalized before the end of 2016. Some disagree, but ES6 will likely maintain its dominant mindshare over the late-change substitute ES2015. However, ES2016 may in fact signal the new year-based schema.

It has also been observed that the pace of JS evolution is much faster even than single-year versioning. As soon as an idea begins to progress through standards discussions, browsers start prototyping the feature, and early adopters start experimenting with the code.

Usually well before there’s an official stamp of approval, a feature is de facto standardized by virtue of this early engine/tooling prototyping. So it’s also valid to consider the future of JS versioning to be per-feature rather than per-arbitrary-collection-of-major-features (as it is now) or even per-year (as it may become).

The takeaway is that the version labels stop being as important, and JavaScript starts to be seen more as an evergreen, living standard. The best way to cope with this is to stop thinking about your code base as being “ES6-based,” for instance, and instead consider it feature by feature for support.

Transpiling

Made even worse by the rapid evolution of features, a problem arises for JS developers who at once may both strongly desire to use new features while at

the same time being slapped with the reality that their sites/apps may need to support older browsers without such support.

The way ES5 appears to have played out in the broader industry, the typical mindset was that code bases waited to adopt ES5 until most if not all pre-ES5 environments had fallen out of their support spectrum. As a result, many are just recently (at the time of this writing) starting to adopt things like `strict` mode, which landed in ES5 over five years ago.

It's widely considered to be a harmful approach for the future of the JS ecosystem to wait around and trail the specification by so many years. All those responsible for evolving the language desire for developers to begin basing their code on the new features and patterns as soon as they stabilize in specification form and browsers have a chance to implement them.

So how do we resolve this seeming contradiction? The answer is tooling, specifically a technique called *transpiling* (transformation + compiling). Roughly, the idea is to use a special tool to transform your ES6 code into equivalent (or close!) matches that work in ES5 environments.

For example, consider shorthand property definitions (see “Object Literal Extensions” in Chapter 2). Here’s the ES6 form:

```
var foo = [1,2,3];

var obj = {
  foo      // means `foo: foo`
};

obj.foo;    // [1,2,3]
```

But (roughly) here’s how that transpiles:

```
var foo = [1,2,3];

var obj = {
  foo: foo
};

obj.foo;    // [1,2,3]
```

This is a minor but pleasant transformation that lets us shorten the `foo: foo` in an object literal declaration to just `foo`, if the names are the same.

Transpilers perform these transformations for you, usually in a build workflow step similar to how you perform linting, minification, and other similar operations.

Shims/Polyfills

Not all new ES6 features need a transpiler. Polyfills (aka shims) are a pattern for defining equivalent behavior from a newer environment into an older environment, when possible. Syntax cannot be polyfilled, but APIs often can be.

For example, `Object.is(..)` is a new utility for checking strict equality of two values but without the nuanced exceptions that `==` has for `NaN` and `-0` values. The polyfill for `Object.is(..)` is pretty easy:

```
if (!Object.is) {
  Object.is = function(v1, v2) {
    // test for ` -0 `
    if (v1 === 0 && v2 === 0) {
      return 1 / v1 === 1 / v2;
    }
    // test for `NaN`
    if (v1 !== v1) {
      return v2 !== v2;
    }
    // everything else
    return v1 === v2;
  };
}
```

Tip: Pay attention to the outer `if` statement guard wrapped around the polyfill. This is an important detail, which means the snippet only defines its fallback behavior for older environments where the API in question isn't already defined; it would be very rare that you'd want to overwrite an existing API.

There's a great collection of ES6 shims called "ES6 Shim" (<https://github.com/paulmillr/es6-shim/>) that you should definitely adopt as a standard part of any new JS project!

It is assumed that JS will continue to evolve constantly, with browsers rolling out support for features continually rather than in large chunks. So the best strategy for keeping updated as it evolves is to just introduce polyfill shims into your code base, and a transpiler step into your build workflow, right now and get used to that new reality.

If you decide to keep the status quo and just wait around for all browsers without a feature supported to go away before you start using the feature, you're always going to be way behind. You'll sadly be missing out on all the innovations designed to make writing JavaScript more effective, efficient, and robust.

Review

ES6 (some may try to call it ES2015) is just landing as of the time of this writing, and it has lots of new stuff you need to learn!

But it's even more important to shift your mindset to align with the new way that JavaScript is going to evolve. It's not just waiting around for years for some official document to get a vote of approval, as many have done in the past.

Now, JavaScript features land in browsers as they become ready, and it's up to you whether you'll get on the train early or whether you'll be playing costly catch-up games years from now.

Whatever labels that future JavaScript adopts, it's going to move a lot quicker than it ever has before. Transpilers and shims/polyfills are important tools to keep you on the forefront of where the language is headed.

If there's any narrative important to understand about the new reality for JavaScript, it's that all JS developers are strongly implored to move from the trailing edge of the curve to the leading edge. And learning ES6 is where that all starts!

You Don't Know JS: ES6 & Beyond

Chapter 2: Syntax

If you've been writing JS for any length of time, odds are the syntax is pretty familiar to you. There are certainly many quirks, but overall it's a fairly reasonable and straightforward syntax that draws many similarities from other languages.

However, ES6 adds quite a few new syntactic forms that take some getting used to. In this chapter, we'll tour through them to find out what's in store.

Tip: At the time of this writing, some of the features discussed in this book have been implemented in various browsers (Firefox, Chrome, etc.), but some have only been partially implemented and many others have not been implemented at all. Your experience may be mixed trying these examples directly. If so, try them out with transpilers, as most of these features are covered by those tools. ES6Fiddle (<http://www.es6fiddle.net/>) is a great, easy-to-use playground for trying out ES6, as is the online REPL for the Babel transpiler (<http://babeljs.io/repl/>).

Block-Scoped Declarations

You're probably aware that the fundamental unit of variable scoping in JavaScript has always been the `function`. If you needed to create a block of scope, the most prevalent way to do so other than a regular function declaration was the immediately invoked function expression (IIFE). For example:

```
var a = 2;

(function IIFE(){
  var a = 3;
  console.log( a );    // 3
})();

console.log( a );      // 2
```

let Declarations

However, we can now create declarations that are bound to any block, called (unsurprisingly) *block scoping*. This means all we need is a pair of `{ ... }` to create a scope. Instead of using `var`, which always declares variables attached to the enclosing function (or global, if top level) scope, use `let`:

```
var a = 2;

{
  let a = 3;
```

```
    console.log( a );    // 3
}

console.log( a );      // 2
```

It's not very common or idiomatic thus far in JS to use a standalone `{ ... }` block, but it's always been valid. And developers from other languages that have *block scoping* will readily recognize that pattern.

I believe this is the best way to create block-scoped variables, with a dedicated `{ ... }` block. Moreover, you should always put the `let` declaration(s) at the very top of that block. If you have more than one to declare, I'd recommend using just one `let`.

Stylistically, I even prefer to put the `let` on the same line as the opening `{`, to make it clearer that this block is only for the purpose of declaring the scope for those variables.

```
{  let a = 2, b, c;
  // ...
}
```

Now, that's going to look strange and it's not likely going to match the recommendations given in most other ES6 literature. But I have reasons for my madness.

There's another experimental (not standardized) form of the `let` declaration called the `let-block`, which looks like:

```
let (a = 2, b, c) {
  // ...
}
```

That form is what I call *explicit* block scoping, whereas the `let ...` declaration form that mirrors `var` is more *implicit*, as it kind of hijacks whatever `{ ... }` pair it's found in. Generally developers find *explicit* mechanisms a bit more preferable than *implicit* mechanisms, and I claim this is one of those cases.

If you compare the previous two snippet forms, they're very similar, and in my opinion both qualify stylistically as *explicit* block scoping. Unfortunately, the `let (...) { ... }` form, the most *explicit* of the options, was not adopted in ES6. That may be revisited post-ES6, but for now the former option is our best bet, I think.

To reinforce the *implicit* nature of `let ..` declarations, consider these usages:

```
let a = 2;

if (a > 1) {
  let b = a * 3;
  console.log( b );      // 6
```

```

for (let i = a; i <= b; i++) {
  let j = i + 10;
  console.log( j );
}
// 12 13 14 15 16

let c = a + b;
console.log( c );           // 8
}

```

Quick quiz without looking back at that snippet: which variable(s) exist only inside the `if` statement, and which variable(s) exist only inside the `for` loop?

The answers: the `if` statement contains `b` and `c` block-scoped variables, and the `for` loop contains `i` and `j` block-scoped variables.

Did you have to think about it for a moment? Does it surprise you that `i` isn't added to the enclosing `if` statement scope? That mental pause and questioning – I call it a “mental tax” – comes from the fact that this `let` mechanism is not only new to us, but it's also *implicit*.

There's also hazard in the `let c = ...` declaration appearing so far down in the scope. Unlike traditional `var`-declared variables, which are attached to the entire enclosing function scope regardless of where they appear, `let` declarations attach to the block scope but are not initialized until they appear in the block.

Accessing a `let`-declared variable earlier than its `let ...` declaration/initialization causes an error, whereas with `var` declarations the ordering doesn't matter (except stylistically).

Consider:

```
{
  console.log( a );    // undefined
  console.log( b );    // ReferenceError!

  var a;
  let b;
}
```

Warning: This `ReferenceError` from accessing too-early `let`-declared references is technically called a *Temporal Dead Zone (TDZ)* error – you're accessing a variable that's been declared but not yet initialized. This will not be the only time we see TDZ errors – they crop up in several places in ES6. Also, note that “initialized” doesn't require explicitly assigning a value in your code, as `let b;` is totally valid. A variable that's not given an assignment at declaration time is assumed to have been assigned the `undefined` value, so `let b;` is the same as `let b = undefined;`. Explicit assignment or not, you cannot access `b` until the `let b` statement is run.

One last gotcha: `typeof` behaves differently with TDZ variables than it does with undeclared (or declared!) variables. For example:

```
{  
  // `a` is not declared  
  if (typeof a === "undefined") {  
    console.log( "cool" );  
  }  
  
  // `b` is declared, but in its TDZ  
  if (typeof b === "undefined") {    // ReferenceError!  
    // ..  
  }  
  
  // ..  
  
  let b;  
}
```

The `a` is not declared, so `typeof` is the only safe way to check for its existence or not. But `typeof b` throws the TDZ error because farther down in the code there happens to be a `let b` declaration. Oops.

Now it should be clearer why I insist that `let` declarations should all be at the top of their scope. That totally avoids the accidental errors of accessing too early. It also makes it more *explicit* when you look at the start of a block, any block, what variables it contains.

Your blocks (`if` statements, `while` loops, etc.) don't have to share their original behavior with scoping behavior.

This explicitness on your part, which is up to you to maintain with discipline, will save you lots of refactor headaches and footguns down the line.

Note: For more information on `let` and block scoping, see Chapter 3 of the *Scope & Closures* title of this series.

let + for The only exception I'd make to the preference for the *explicit* form of `let` declaration blocking is a `let` that appears in the header of a `for` loop. The reason may seem nuanced, but I believe it to be one of the more important ES6 features.

Consider:

```
var funcs = [];  
  
for (let i = 0; i < 5; i++) {  
  funcs.push( function(){  
    console.log( i );  
  } );
```

```

        } );
}

funcs[3]();      // 3

```

The `let i` in the `for` header declares an `i` not just for the `for` loop itself, but it redeclares a new `i` for each iteration of the loop. That means that closures created inside the loop iteration close over those per-iteration variables the way you'd expect.

If you tried that same snippet but with `var i` in the `for` loop header, you'd get 5 instead of 3, because there'd only be one `i` in the outer scope that was closed over, instead of a new `i` for each iteration's function to close over.

You could also have accomplished the same thing slightly more verbosely:

```

var funcs = [];

for (var i = 0; i < 5; i++) {
  let j = i;
  funcs.push( function(){
    console.log( j );
  });
}

funcs[3]();      // 3

```

Here, we forcibly create a new `j` for each iteration, and then the closure works the same way. I prefer the former approach; that extra special capability is why I endorse the `for (let ...) ...` form. It could be argued it's somewhat more *implicit*, but it's *explicit* enough, and useful enough, for my tastes.

`let` also works the same way with `for..in` and `for..of` loops (see “`for..of` Loops”).

const Declarations

There's one other form of block-scoped declaration to consider: the `const`, which creates *constants*.

What exactly is a constant? It's a variable that's read-only after its initial value is set. Consider:

```

{
  const a = 2;
  console.log( a );    // 2

  a = 3;              // TypeError!
}

```

You are not allowed to change the value the variable holds once it's been set, at declaration time. A `const` declaration must have an explicit initialization. If you wanted a *constant* with the `undefined` value, you'd have to declare `const a = undefined` to get it.

Constants are not a restriction on the value itself, but on the variable's assignment of that value. In other words, the value is not frozen or immutable because of `const`, just the assignment of it. If the value is complex, such as an object or array, the contents of the value can still be modified:

```
{  
  const a = [1,2,3];  
  a.push( 4 );  
  console.log( a );      // [1,2,3,4]  
  
  a = 42;              // TypeError!  
}
```

The `a` variable doesn't actually hold a constant array; rather, it holds a constant reference to the array. The array itself is freely mutable.

Warning: Assigning an object or array as a constant means that value will not be able to be garbage collected until that constant's lexical scope goes away, as the reference to the value can never be unset. That may be desirable, but be careful if it's not your intent!

Essentially, `const` declarations enforce what we've stylistically signaled with our code for years, where we declared a variable name of all uppercase letters and assigned it some literal value that we took care never to change. There's no enforcement on a `var` assignment, but there is now with a `const` assignment, which can help you catch unintended changes.

`const` can be used with variable declarations of `for`, `for..in`, and `for..of` loops (see “`for..of` Loops”). However, an error will be thrown if there's any attempt to reassign, such as the typical `i++` clause of a `for` loop.

const Or Not There's some rumored assumptions that a `const` could be more optimizable by the JS engine in certain scenarios than a `let` or `var` would be. Theoretically, the engine more easily knows the variable's value/type will never change, so it can eliminate some possible tracking.

Whether `const` really helps here or this is just our own fantasies and intuitions, the much more important decision to make is if you intend constant behavior or not. Remember: one of the most important roles for source code is to communicate clearly, not only to you, but your future self and other code collaborators, what your intent is.

Some developers prefer to start out every variable declaration as a `const` and then relax a declaration back to a `let` if it becomes necessary for its value to

change in the code. This is an interesting perspective, but it's not clear that it genuinely improves the readability or reason-ability of code.

It's not really a *protection*, as many believe, because any later developer who wants to change a value of a `const` can just blindly change `const` to `let` on the declaration. At best, it protects accidental change. But again, other than our intuitions and sensibilities, there doesn't appear to be objective and clear measure of what constitutes "accidents" or prevention thereof. Similar mindsets exist around type enforcement.

My advice: to avoid potentially confusing code, only use `const` for variables that you're intentionally and obviously signaling will not change. In other words, don't *rely on* `const` for code behavior, but instead use it as a tool for signaling intent, when intent can be signaled clearly.

Block-scoped Functions

Starting with ES6, function declarations that occur inside of blocks are now specified to be scoped to that block. Prior to ES6, the specification did not call for this, but many implementations did it anyway. So now the specification meets reality.

Consider:

```
{  
  foo();           // works!  
  
  function foo() {  
    // ..  
  }  
}  
  
foo();           // ReferenceError
```

The `foo()` function is declared inside the `{ ... }` block, and as of ES6 is block-scoped there. So it's not available outside that block. But also note that it is "hoisted" within the block, as opposed to `let` declarations, which suffer the TDZ error trap mentioned earlier.

Block-scoping of function declarations could be a problem if you've ever written code like this before, and relied on the old legacy non-block-scoped behavior:

```
if (something) {  
  function foo() {  
    console.log( "1" );  
  }  
}  
else {  
  function foo() {
```

```

        console.log( "2" );
    }
}

foo();      // ??

```

In pre-ES6 environments, `foo()` would print "2" regardless of the value of `something`, because both function declarations were hoisted out of the blocks, and the second one always wins.

In ES6, that last line throws a `ReferenceError`.

Spread/Rest

ES6 introduces a new `...` operator that's typically referred to as the *spread* or *rest* operator, depending on where/how it's used. Let's take a look:

```

function foo(x,y,z) {
    console.log( x, y, z );
}

foo( ...[1,2,3] );           // 1 2 3

```

When `...` is used in front of an array (actually, any *iterable*, which we cover in Chapter 3), it acts to “spread” it out into its individual values.

You'll typically see that usage as is shown in that previous snippet, when spreading out an array as a set of arguments to a function call. In this usage, `...` acts to give us a simpler syntactic replacement for the `apply(..)` method, which we would typically have used pre-ES6 as:

```
foo.apply( null, [1,2,3] );    // 1 2 3
```

But `...` can be used to spread out/expand a value in other contexts as well, such as inside another array declaration:

```

var a = [2,3,4];
var b = [ 1, ...a, 5 ];

console.log( b );           // [1,2,3,4,5]

```

In this usage, `...` is basically replacing `concat(..)`, as it behaves like `[1].concat(a, [5])` here.

The other common usage of `...` can be seen as essentially the opposite; instead of spreading a value out, the `...` *gathers* a set of values together into an array. Consider:

```

function foo(x, y, ...z) {
    console.log( x, y, z );
}

```

```
foo( 1, 2, 3, 4, 5 );           // 1 2 [3,4,5]
```

The `...z` in this snippet is essentially saying: “gather the *rest* of the arguments (if any) into an array called `z`.” Because `x` was assigned 1, and `y` was assigned 2, the rest of the arguments 3, 4, and 5 were gathered into `z`.

Of course, if you don’t have any named parameters, the `...` gathers all arguments:

```
function foo(...args) {
  console.log( args );
}

foo( 1, 2, 3, 4, 5 );           // [1,2,3,4,5]
```

Note: The `...args` in the `foo(..)` function declaration is usually called “rest parameters,” because you’re collecting the rest of the parameters. I prefer “gather,” because it’s more descriptive of what it does rather than what it contains.

The best part about this usage is that it provides a very solid alternative to using the long-since-deprecated `arguments` array – actually, it’s not really an array, but an array-like object. Because `args` (or whatever you call it – a lot of people prefer `r` or `rest`) is a real array, we can get rid of lots of silly pre-ES6 tricks we jumped through to make `arguments` into something we can treat as an array.

Consider:

```
// doing things the new ES6 way
function foo(...args) {
  // `args` is already a real array

  // discard first element in `args`
  args.shift();

  // pass along all of `args` as arguments
  // to `console.log(..)`
  console.log( ...args );
}

// doing things the old-school pre-ES6 way
function bar() {
  // turn `arguments` into a real array
  var args = Array.prototype.slice.call( arguments );

  // add some elements on the end
  args.push( 4, 5 );
```

```

// filter out odd numbers
args = args.filter( function(v){
  return v % 2 == 0;
} );

// pass along all of `args` as arguments
// to `foo(..)`
foo.apply( null, args );
}

bar( 0, 1, 2, 3 );           // 24

```

The `...args` in the `foo(..)` function declaration gathers arguments, and the `...args` in the `console.log(..)` call spreads them out. That's a good illustration of the symmetric but opposite uses of the `...` operator.

Besides the `...` usage in a function declaration, there's another case where `...` is used for gathering values, and we'll look at it in the "Too Many, Too Few, Just Enough" section later in this chapter.

Default Parameter Values

Perhaps one of the most common idioms in JavaScript relates to setting a default value for a function parameter. The way we've done this for years should look quite familiar:

```

function foo(x,y) {
  x = x || 11;
  y = y || 31;

  console.log( x + y );
}

foo();           // 42
foo( 5, 6 );    // 11
foo( 5 );       // 36
foo( null, 6 ); // 17

```

Of course, if you've used this pattern before, you know that it's both helpful and a little bit dangerous, if for example you need to be able to pass in what would otherwise be considered a falsy value for one of the parameters. Consider:

```
foo( 0, 42 );    // 53 <-- Oops, not 42
```

Why? Because the `0` is falsy, and so the `x || 11` results in `11`, not the directly passed in `0`.

To fix this gotcha, some people will instead write the check more verbosely like

this:

```
function foo(x,y) {
  x = (x !== undefined) ? x : 11;
  y = (y !== undefined) ? y : 31;

  console.log( x + y );
}

foo( 0, 42 );           // 42
foo( undefined, 6 );    // 17
```

Of course, that means that any value except `undefined` can be directly passed in. However, `undefined` will be assumed to signal, “I didn’t pass this in.” That works great unless you actually need to be able to pass `undefined` in.

In that case, you could test to see if the argument is actually omitted, by it actually not being present in the `arguments` array, perhaps like this:

```
function foo(x,y) {
  x = (0 in arguments) ? x : 11;
  y = (1 in arguments) ? y : 31;

  console.log( x + y );
}

foo( 5 );           // 36
foo( 5, undefined );    // NaN
```

But how would you omit the first `x` argument without the ability to pass in any kind of value (not even `undefined`) that signals “I’m omitting this argument”?

`foo(,5)` is tempting, but it’s invalid syntax. `foo.apply(null,[,5])` seems like it should do the trick, but `apply(..)`’s quirks here mean that the arguments are treated as `[undefined,5]`, which of course doesn’t omit.

If you investigate further, you’ll find you can only omit arguments on the end (i.e., righthand side) by simply passing fewer arguments than “expected,” but you cannot omit arguments in the middle or at the beginning of the arguments list. It’s just not possible.

There’s a principle applied to JavaScript’s design here that is important to remember: `undefined` means *missing*. That is, there’s no difference between `undefined` and *missing*, at least as far as function arguments go.

Note: There are, confusingly, other places in JS where this particular design principle doesn’t apply, such as for arrays with empty slots. See the *Types & Grammar* title of this series for more information.

With all this in mind, we can now examine a nice helpful syntax added as of ES6 to streamline the assignment of default values to missing arguments:

```

function foo(x = 11, y = 31) {
  console.log( x + y );
}

foo();                      // 42
foo( 5, 6 );                // 11
foo( 0, 42 );                // 42

foo( 5 );                    // 36
foo( 5, undefined );        // 36 <-- `undefined` is missing
foo( 5, null );              // 5 <-- null coerces to `0`

foo( undefined, 6 );         // 17 <-- `undefined` is missing
foo( null, 6 );              // 6 <-- null coerces to `0`

```

Notice the results and how they imply both subtle differences and similarities to the earlier approaches.

`x = 11` in a function declaration is more like `x !== undefined ? x : 11` than the much more common idiom `x || 11`, so you'll need to be careful in converting your pre-ES6 code to this ES6 default parameter value syntax.

Note: A rest/gather parameter (see “Spread/Rest”) cannot have a default value. So, while `function foo(...vals=[1,2,3]) {` might seem an intriguing capability, it’s not valid syntax. You’ll need to continue to apply that sort of logic manually if necessary.

Default Value Expressions

Function default values can be more than just simple values like `31`; they can be any valid expression, even a function call:

```

function bar(val) {
  console.log( "bar called!" );
  return y + val;
}

function foo(x = y + 3, z = bar( x )) {
  console.log( x, z );
}

var y = 5;
foo();                      // "bar called"
                           // 8 13
foo( 10 );                  // "bar called"
                           // 10 15
y = 6;
foo( undefined, 10 );        // 9 10

```

As you can see, the default value expressions are lazily evaluated, meaning they're only run if and when they're needed – that is, when a parameter's argument is omitted or is `undefined`.

It's a subtle detail, but the formal parameters in a function declaration are in their own scope (think of it as a scope bubble wrapped around just the `(..)` of the function declaration), not in the function body's scope. That means a reference to an identifier in a default value expression first matches the formal parameters' scope before looking to an outer scope. See the *Scope & Closures* title of this series for more information.

Consider:

```
var w = 1, z = 2;

function foo( x = w + 1, y = x + 1, z = z + 1 ) {
    console.log( x, y, z );
}

foo();           // ReferenceError
```

The `w` in the `w + 1` default value expression looks for `w` in the formal parameters' scope, but does not find it, so the outer scope's `w` is used. Next, The `x` in the `x + 1` default value expression finds `x` in the formal parameters' scope, and luckily `x` has already been initialized, so the assignment to `y` works fine.

However, the `z` in `z + 1` finds `z` as a not-yet-initialized-at-that-moment parameter variable, so it never tries to find the `z` from the outer scope.

As we mentioned in the “`let` Declarations” section earlier in this chapter, ES6 has a TDZ, which prevents a variable from being accessed in its uninitialized state. As such, the `z + 1` default value expression throws a TDZ `ReferenceError` error.

Though it's not necessarily a good idea for code clarity, a default value expression can even be an inline function expression call – commonly referred to as an immediately invoked function expression (IIFE):

```
function foo( x =
    (function(v){ return v + 11; })( 31 )
) {
    console.log( x );
}

foo();           // 42
```

There will very rarely be any cases where an IIFE (or any other executed inline function expression) will be appropriate for default value expressions. If you find yourself tempted to do this, take a step back and reevaluate!

Warning: If the IIFE had tried to access the `x` identifier and had not declared its own `x`, this would also have been a TDZ error, just as discussed before.

The default value expression in the previous snippet is an IIFE in that in the sense that it's a function that's executed right inline, via (31). If we had left that part off, the default value assigned to `x` would have just been a function reference itself, perhaps like a default callback. There will probably be cases where that pattern will be quite useful, such as:

```
function ajax(url, cb = function(){}) {
  // ...
}

ajax( "http://some.url.1" );
```

In this case, we essentially want to default `cb` to be a no-op empty function call if not otherwise specified. The function expression is just a function reference, not a function call itself (no invoking `()` on the end of it), which accomplishes that goal.

Since the early days of JS, there's been a little-known but useful quirk available to us: `Function.prototype` is itself an empty no-op function. So, the declaration could have been `cb = Function.prototype` and saved the inline function expression creation.

Destructuring

ES6 introduces a new syntactic feature called *destructuring*, which may be a little less confusing if you instead think of it as *structured assignment*. To understand this meaning, consider:

```
function foo() {
  return [1,2,3];
}

var tmp = foo(),
  a = tmp[0], b = tmp[1], c = tmp[2];

console.log( a, b, c );           // 1 2 3
```

As you can see, we created a manual assignment of the values in the array that `foo()` returns to individual variables `a`, `b`, and `c`, and to do so we (unfortunately) needed the `tmp` variable.

Similarly, we can do the following with objects:

```
function bar() {
  return {
    x: 4,
    y: 5,
  }
}
```

```

        z: 6
    };
}

var tmp = bar(),
    x = tmp.x, y = tmp.y, z = tmp.z;

console.log( x, y, z );           // 4 5 6

```

The `tmp.x` property value is assigned to the `x` variable, and likewise for `tmp.y` to `y` and `tmp.z` to `z`.

Manually assigning indexed values from an array or properties from an object can be thought of as *structured assignment*. ES6 adds a dedicated syntax for *destructuring*, specifically *array destructuring* and *object destructuring*. This syntax eliminates the need for the `tmp` variable in the previous snippets, making them much cleaner. Consider:

```

var [ a, b, c ] = foo();
var { x: x, y: y, z: z } = bar();

console.log( a, b, c );           // 1 2 3
console.log( x, y, z );           // 4 5 6

```

You're likely more accustomed to seeing syntax like `[a,b,c]` on the righthand side of an `=` assignment, as the value being assigned.

Destructuring symmetrically flips that pattern, so that `[a,b,c]` on the lefthand side of the `=` assignment is treated as a kind of "pattern" for decomposing the righthand side array value into separate variable assignments.

Similarly, `{ x: x, y: y, z: z }` specifies a "pattern" to decompose the object value from `bar()` into separate variable assignments.

Object Property Assignment Pattern

Let's dig into that `{ x: x, ... }` syntax from the previous snippet. If the property name being matched is the same as the variable you want to declare, you can actually shorten the syntax:

```

var { x, y, z } = bar();

console.log( x, y, z );           // 4 5 6

```

Pretty cool, right?

But is `{ x, ... }` leaving off the `x: part` or leaving off the `: x part`? We're actually leaving off the `x: part` when we use the shorter syntax. That may not seem like an important detail, but you'll understand its importance in just a moment.

If you can write the shorter form, why would you ever write out the longer form? Because that longer form actually allows you to assign a property to a different variable name, which can sometimes be quite useful:

```
var { x: bam, y: baz, z: bap } = bar();  
  
console.log( bam, baz, bap );           // 4 5 6  
console.log( x, y, z );                // ReferenceError
```

There's a subtle but super-important quirk to understand about this variation of the object destructuring form. To illustrate why it can be a gotcha you need to be careful of, let's consider the "pattern" of how normal object literals are specified:

```
var X = 10, Y = 20;  
  
var o = { a: X, b: Y };  
  
console.log( o.a, o.b );           // 10 20
```

In `{ a: X, b: Y }`, we know that `a` is the object property, and `X` is the source value that gets assigned to it. In other words, the syntactic pattern is `target: source`, or more obviously, `property-alias: value`. We intuitively understand this because it's the same as `= assignment`, where the pattern is `target = source`.

However, when you use object destructuring assignment – that is, putting the `{ ... }` object literal-looking syntax on the lefthand side of the `=` operator – you invert that `target: source` pattern.

Recall:

```
var { x: bam, y: baz, z: bap } = bar();
```

The syntactic pattern here is `source: target` (or `value: variable-alias`). `x: bam` means the `x` property is the source value and `bam` is the target variable to assign to. In other words, object literals are `target <-- source`, and object destructuring assignments are `source --> target`. See how that's flipped?

There's another way to think about this syntax though, which may help ease the confusion. Consider:

```
var aa = 10, bb = 20;  
  
var o = { x: aa, y: bb };  
var { x: AA, y: BB } = o;  
  
console.log( AA, BB );           // 10 20
```

In the `{ x: aa, y: bb }` line, the `x` and `y` represent the object properties. In the `{ x: AA, y: BB }` line, the `x` and the `y` *also* represent the object properties.

Recall how earlier I asserted that `{ x, ... }` was leaving off the `x:` part? In those two lines, if you erase the `x:` and `y:` parts in that snippet, you're left only with `aa`, `bb` and `AA`, `BB`, which in effect – only conceptually, not actually – are assignments from `aa` to `AA` and from `bb` to `BB`.

So, that symmetry may help to explain why the syntactic pattern was intentionally flipped for this ES6 feature.

Note: I would have preferred the syntax to be `{ AA: x, BB: y }` for the destructuring assignment, as that would have preserved consistency of the more familiar `target: source` pattern for both usages. Alas, I'm having to train my brain for the inversion, as some readers may also have to do.

Not Just Declarations

So far, we've used destructuring assignment with `var` declarations (of course, they could also use `let` and `const`), but destructuring is a general assignment operation, not just a declaration.

Consider:

```
var a, b, c, x, y, z;  
  
[a,b,c] = foo();  
( { x, y, z } = bar() );  
  
console.log( a, b, c );           // 1 2 3  
console.log( x, y, z );           // 4 5 6
```

The variables can already be declared, and then the destructuring only does assignments, exactly as we've already seen.

Note: For the object destructuring form specifically, when leaving off a `var/let/const` declarator, we had to surround the whole assignment expression in `()`, because otherwise the `{ ... }` on the lefthand side as the first element in the statement is taken to be a block statement instead of an object.

In fact, the assignment expressions (`a`, `y`, etc.) don't actually need to be just variable identifiers. Anything that's a valid assignment expression is allowed. For example:

```
var o = {};  
  
[o.a, o.b, o.c] = foo();  
( { x: o.x, y: o.y, z: o.z } = bar() );  
  
console.log( o.a, o.b, o.c );           // 1 2 3  
console.log( o.x, o.y, o.z );           // 4 5 6
```

You can even use computed property expressions in the destructuring. Consider:

```

var which = "x",
o = {};

( { [which]: o[which] } = bar() );

console.log( o.x );           // 4

```

The `[which]:` part is the computed property, which results in `x` – the property to destructure from the object in question as the source of the assignment. The `o[which]` part is just a normal object key reference, which equates to `o.x` as the target of the assignment.

You can use the general assignments to create object mappings/transformations, such as:

```

var o1 = { a: 1, b: 2, c: 3 },
o2 = {};

( { a: o2.x, b: o2.y, c: o2.z } = o1 );

console.log( o2.x, o2.y, o2.z );    // 1 2 3

```

Or you can map an object to an array, such as:

```

var o1 = { a: 1, b: 2, c: 3 },
a2 = [];

( { a: a2[0], b: a2[1], c: a2[2] } = o1 );

console.log( a2 );           // [1,2,3]

```

Or the other way around:

```

var a1 = [ 1, 2, 3 ],
o2 = {};

[ o2.a, o2.b, o2.c ] = a1;

console.log( o2.a, o2.b, o2.c );    // 1 2 3

```

Or you could reorder one array to another:

```

var a1 = [ 1, 2, 3 ],
a2 = [];

[ a2[2], a2[0], a2[1] ] = a1;

console.log( a2 );           // [2,3,1]

```

You can even solve the traditional “swap two variables” task without a temporary variable:

```

var x = 10, y = 20;
[ y, x ] = [ x, y ];
console.log( x, y );           // 20 10

```

Warning: Be careful: you shouldn't mix in declaration with assignment unless you want all of the assignment expressions *also* to be treated as declarations. Otherwise, you'll get syntax errors. That's why in the earlier example I had to do **var a2 = []** separately from the **[a2[0], ...] = ...** destructuring assignment. It wouldn't make any sense to try **var [a2[0], ...] = ...**, because **a2[0]** isn't a valid declaration identifier; it also obviously couldn't implicitly create a **var a2 = []** declaration to use.

Repeated Assignments

The object destructuring form allows a source property (holding any value type) to be listed multiple times. For example:

```

var { a: X, a: Y } = { a: 1 };

X; // 1
Y; // 1

```

That also means you can both destructure a sub-object/array property and also capture the sub-object/array's value itself. Consider:

```
var { a: { x: X, x: Y }, a } = { a: { x: 1 } };
```

```

X; // 1
Y; // 1
a; // { x: 1 }

```

```
( { a: X, a: Y, a: [ Z ] } = { a: [ 1 ] } );
```

```

X.push( 2 );
Y[0] = 10;

```

```

X; // [10,2]
Y; // [10,2]
Z; // 1

```

A word of caution about destructuring: it may be tempting to list destructuring assignments all on a single line as has been done thus far in our discussion. However, it's a much better idea to spread destructuring assignment patterns over multiple lines, using proper indentation – much like you would in JSON or with an object literal value – for readability sake.

```
// harder to read:
```

```

var { a: { b: [ c, d ], e: { f } }, g } = obj;

// better:
var {
  a: {
    b: [ c, d ],
    e: { f }
  },
  g
} = obj;

```

Remember: the purpose of destructuring is not just less typing, but more declarative readability.

Destructuring Assignment Expressions The assignment expression with object or array destructuring has as its completion value the full righthand object/array value. Consider:

```

var o = { a:1, b:2, c:3 },
  a, b, c, p;

p = { a, b, c } = o;

console.log( a, b, c );           // 1 2 3
p === o;                          // true

```

In the previous snippet, p was assigned the o object reference, not one of the a, b, or c values. The same is true of array destructuring:

```

var o = [1,2,3],
  a, b, c, p;

p = [ a, b, c ] = o;

console.log( a, b, c );           // 1 2 3
p === o;                          // true

```

By carrying the object/array value through as the completion, you can chain destructuring assignment expressions together:

```

var o = { a:1, b:2, c:3 },
  p = [4,5,6],
  a, b, c, x, y, z;

( {a} = {b,c} = o );
[x,y] = [z] = p;

console.log( a, b, c );           // 1 2 3

```

```
console.log( x, y, z );           // 4 5 4
```

Too Many, Too Few, Just Enough

With both array destructuring assignment and object destructuring assignment, you do not have to assign all the values that are present. For example:

```
var [ ,b] = foo();  
var { x, z } = bar();
```

```
console.log( b, x, z );           // 2 4 6
```

The 1 and 3 values that came back from `foo()` are discarded, as is the 5 value from `bar()`.

Similarly, if you try to assign more values than are present in the value you're destructuring/decomposing, you get graceful fallback to `undefined`, as you'd expect:

```
var [ , ,c,d] = foo();  
var { w, z } = bar();
```

```
console.log( c, z );           // 3 6  
console.log( d, w );           // undefined undefined
```

This behavior follows symmetrically from the earlier stated “`undefined` is missing” principle.

We examined the `...` operator earlier in this chapter, and saw that it can sometimes be used to spread an array value out into its separate values, and sometimes it can be used to do the opposite: to gather a set of values together into an array.

In addition to the gather/rest usage in function declarations, `...` can perform the same behavior in destructuring assignments. To illustrate, let's recall a snippet from earlier in this chapter:

```
var a = [2,3,4];  
var b = [ 1, ...a, 5 ];
```

```
console.log( b );           // [1,2,3,4,5]
```

Here we see that `...a` is spreading `a` out, because it appears in the array `[..]` value position. If `...a` appears in an array destructuring position, it performs the gather behavior:

```
var a = [2,3,4];  
var [ b, ...c ] = a;  
  
console.log( b, c );           // 2 [3,4]
```

The `var [...] = a` destructuring assignment spreads `a` out to be assigned to the pattern described inside the `[...]`. The first part names `b` for the first value in `a` (2). But then `...c` gathers the rest of the values (3 and 4) into an array and calls it `c`.

Note: We've seen how `...` works with arrays, but what about with objects? It's not an ES6 feature, but see Chapter 8 for discussion of a possible "beyond ES6" feature where `...` works with spreading or gathering objects.

Default Value Assignment

Both forms of destructuring can offer a default value option for an assignment, using the `=` syntax similar to the default function argument values discussed earlier.

Consider:

```
var [ a = 3, b = 6, c = 9, d = 12 ] = foo();
var { x = 5, y = 10, z = 15, w = 20 } = bar();

console.log( a, b, c, d );           // 1 2 3 12
console.log( x, y, z, w );           // 4 5 6 20
```

You can combine the default value assignment with the alternative assignment expression syntax covered earlier. For example:

```
var { x, y, z, w: WW = 20 } = bar();

console.log( x, y, z, WW );           // 4 5 6 20
```

Be careful about confusing yourself (or other developers who read your code) if you use an object or array as the default value in a destructuring. You can create some really hard to understand code:

```
var x = 200, y = 300, z = 100;
var o1 = { x: { y: 42 }, z: { y: z } };

( { y: x = { y: y } } = o1 );
( { z: y = { y: z } } = o1 );
( { x: z = { y: x } } = o1 );
```

Can you tell from that snippet what values `x`, `y`, and `z` have at the end? Takes a moment of pondering, I would imagine. I'll end the suspense:

```
console.log( x.y, y.y, z.y );           // 300 100 42
```

The takeaway here: destructuring is great and can be very useful, but it's also a sharp sword that can cause injury (to someone's brain) if used unwisely.

Nested Destructuring

If the values you're destructuring have nested objects or arrays, you can destructure those nested values as well:

```
var a1 = [ 1, [2, 3, 4], 5 ];
var o1 = { x: { y: { z: 6 } } };

var [ a, [ b, c, d ], e ] = a1;
var { x: { y: { z: w } } } = o1;

console.log( a, b, c, d, e );           // 1 2 3 4 5
console.log( w );                      // 6
```

Nested destructuring can be a simple way to flatten out object namespaces. For example:

```
var App = {
  model: {
    User: function(){ ... }
  }
};

// instead of:
// var User = App.model.User;

var { model: { User } } = App;
```

Destructuring Parameters

In the following snippet, can you spot the assignment?

```
function foo(x) {
  console.log( x );
}

foo( 42 );
```

The assignment is kinda hidden: `42` (the argument) is assigned to `x` (the parameter) when `foo(42)` is executed. If parameter/argument pairing is an assignment, then it stands to reason that it's an assignment that could be destructured, right? Of course!

Consider array destructuring for parameters:

```
function foo( [ x, y ] ) {
  console.log( x, y );
}

foo( [ 1, 2 ] );           // 1 2
```

```
foo( [ 1 ] );                                // 1 undefined
foo( [] );                                    // undefined undefined
```

Object destructuring for parameters works, too:

```
function foo( { x, y } ) {
  console.log( x, y );
}

foo( { y: 1, x: 2 } );                      // 2 1
foo( { y: 42 } );                           // undefined 42
foo( {} );                                  // undefined undefined
```

This technique is an approximation of named arguments (a long requested feature for JS!), in that the properties on the object map to the destructured parameters of the same names. That also means that we get optional parameters (in any position) for free, as you can see leaving off the `x` “parameter” worked as we’d expect.

Of course, all the previously discussed variations of destructuring are available to us with parameter destructuring, including nested destructuring, default values, and more. Destructuring also mixes fine with other ES6 function parameter capabilities, like default parameter values and rest/gather parameters.

Consider these quick illustrations (certainly not exhaustive of the possible variations):

```
function f1([ x=2, y=3, z ]){ ... }
function f2([ x, y, ...z ], w){ ... }
function f3([ x, y, ...z ], ...w){ ... }

function f4({ x: X, y }){ ... }
function f5({ x: X = 10, y = 20 }){ ... }
function f6({ x = 10 } = {}, { y } = { y: 10 }){ ... }
```

Let’s take one example from this snippet and examine it, for illustration purposes:

```
function f3([ x, y, ...z ], ...w) {
  console.log( x, y, z, w );
}

f3();                                         // undefined undefined []
f3( [1,2,3,4], 5, 6 );                      // 1 2 [3,4] [5,6]
```

There are two `...` operators in use here, and they’re both gathering values in arrays (`z` and `w`), though `...z` gathers from the rest of the values left over in the first array argument, while `...w` gathers from the rest of the main arguments left over after the first.

Destructuring Defaults + Parameter Defaults There's one subtle point you should be particularly careful to notice – the difference in behavior between a destructuring default value and a function parameter default value. For example:

```
function f6({ x = 10 } = {}, { y } = { y: 10 }) {
  console.log( x, y );
}

f6(); // 10 10
```

At first, it would seem that we've declared a default value of 10 for both the `x` and `y` parameters, but in two different ways. However, these two different approaches will behave differently in certain cases, and the difference is awfully subtle.

Consider:

```
f6( {}, {} ); // 10 undefined
```

Wait, why did that happen? It's pretty clear that named parameter `x` is defaulting to 10 if not passed as a property of that same name in the first argument's object.

But what about `y` being `undefined`? The `{ y: 10 }` value is an object as a function parameter default value, not a destructuring default value. As such, it only applies if the second argument is not passed at all, or is passed as `undefined`.

In the previous snippet, we *are* passing a second argument (`{}`), so the default `{ y: 10 }` value is not used, and the `{ y }` destructuring occurs against the passed in `{}` empty object value.

Now, compare `{ y } = { y: 10 }` to `{ x = 10 } = {}`.

For the `x`'s form usage, if the first function argument is omitted or `undefined`, the `{}` empty object default applies. Then, whatever value is in the first argument position – either the default `{}` or whatever you passed in – is destructured with the `{ x = 10 }`, which checks to see if an `x` property is found, and if not found (or `undefined`), the 10 default value is applied to the `x` named parameter.

Deep breath. Read back over those last few paragraphs a couple of times. Let's review via code:

```
function f6({ x = 10 } = {}, { y } = { y: 10 }) {
  console.log( x, y );
}

f6(); // 10 10
f6( undefined, undefined ); // 10 10
f6( {}, undefined ); // 10 10
```

```

f6( {}, {} );
// 10 undefined
f6( undefined, {} );
// 10 undefined

f6( { x: 2 }, { y: 3 } );
// 2 3

```

It would generally seem that the defaulting behavior of the `x` parameter is probably the more desirable and sensible case compared to that of `y`. As such, it's important to understand why and how `{ x = 10 } = {}` form is different from `{ y } = { y: 10 }` form.

If that's still a bit fuzzy, go back and read it again, and play with this yourself. Your future self will thank you for taking the time to get this very subtle gotcha nuance detail straight.

Nested Defaults: Destructured and Restructured Although it may at first be difficult to grasp, an interesting idiom emerges for setting defaults for a nested object's properties: using object destructuring along with what I'd call *restructuring*.

Consider a set of defaults in a nested object structure, like the following:

```
// taken from: http://es-discourse.com/t/partial-default-arguments/120/7

var defaults = {
  options: {
    remove: true,
    enable: false,
    instance: {}
  },
  log: {
    warn: true,
    error: true
  }
};

```

Now, let's say that you have an object called `config`, which has some of these applied, but perhaps not all, and you'd like to set all the defaults into this object in the missing spots, but not override specific settings already present:

```

var config = {
  options: {
    remove: false,
    instance: null
  }
};

```

You can of course do so manually, as you might have done in the past:

```

config.options = config.options || {};
config.options.remove = (config.options.remove !== undefined) ?
  config.options.remove : defaults.options.remove;
config.options.enable = (config.options.enable !== undefined) ?
  config.options.enable : defaults.options.enable;
...

```

Yuck.

Others may prefer the assign-overwrite approach to this task. You might be tempted by the ES6 `Object.assign(..)` utility (see Chapter 6) to clone the properties first from `defaults` and then overwritten with the cloned properties from `config`, as so:

```
config = Object.assign( {}, defaults, config );
```

That looks way nicer, huh? But there's a major problem! `Object.assign(..)` is shallow, which means when it copies `defaults.options`, it just copies that object reference, not deep cloning that object's properties to a `config.options` object. `Object.assign(..)` would need to be applied (sort of "recursively") at all levels of your object's tree to get the deep cloning you're expecting.

Note: Many JS utility libraries/frameworks provide their own option for deep cloning of an object, but those approaches and their gotchas are beyond our scope to discuss here.

So let's examine if ES6 object destructuring with defaults can help at all:

```

config.options = config.options || {};
config.log = config.log || {};
({
  options: {
    remove: config.options.remove = defaults.options.remove,
    enable: config.options.enable = defaults.options.enable,
    instance: config.options.instance = defaults.options.instance
  } = {},
  log: {
    warn: config.log.warn = defaults.log.warn,
    error: config.log.error = defaults.log.error
  } = {}
} = config);

```

Not as nice as the false promise of `Object.assign(..)` (being that it's shallow only), but it's better than the manual approach by a fair bit, I think. It is still unfortunately verbose and repetitive, though.

The previous snippet's approach works because I'm hacking the destructuring and defaults mechanism to do the property `==` `undefined` checks and assignment decisions for me. It's a trick in that I'm destructuring `config` (see the `= config` at the end of the snippet), but I'm reassigning all the destructured

values right back into `config`, with the `config.options.enable` assignment references.

Still too much, though. Let's see if we can make anything better.

The following trick works best if you know that all the various properties you're destructuring are uniquely named. You can still do it even if that's not the case, but it's not as nice – you'll have to do the destructuring in stages, or create unique local variables as temporary aliases.

If we fully destructure all the properties into top-level variables, we can then immediately restructure to reconstitute the original nested object structure.

But all those temporary variables hanging around would pollute scope. So, let's use block scoping (see “Block-Scoped Declarations” earlier in this chapter) with a general `{ }` enclosing block:

```
// merge `defaults` into `config`
{
  // destructure (with default value assignments)
  let {
    options: {
      remove = defaults.options.remove,
      enable = defaults.options.enable,
      instance = defaults.options.instance
    } = {},
    log: {
      warn = defaults.log.warn,
      error = defaults.log.error
    } = {}
  } = config;

  // restructure
  config = {
    options: { remove, enable, instance },
    log: { warn, error }
  };
}
```

That seems a fair bit nicer, huh?

Note: You could also accomplish the scope enclosure with an arrow IIFE instead of the general `{ }` block and `let` declarations. Your destructuring assignments/defaults would be in the parameter list and your restructuring would be the `return` statement in the function body.

The `{ warn, error }` syntax in the restructuring part may look new to you; that's called “concise properties” and we cover it in the next section!

Object Literal Extensions

ES6 adds a number of important convenience extensions to the humble `{ ... }` object literal.

Concise Properties

You're certainly familiar with declaring object literals in this form:

```
var x = 2, y = 3,
  o = {
    x: x,
    y: y
 };
```

If it's always felt redundant to say `x: x` all over, there's good news. If you need to define a property that is the same name as a lexical identifier, you can shorten it from `x: x` to `x`. Consider:

```
var x = 2, y = 3,
  o = {
    x,
    y
 };
```

Concise Methods

In a similar spirit to concise properties we just examined, functions attached to properties in object literals also have a concise form, for convenience.

The old way:

```
var o = {
  x: function(){
    // ..
  },
  y: function(){
    // ..
  }
}
```

And as of ES6:

```
var o = {
  x() {
    // ..
  },
  y() {
    // ..
  }
}
```

```
    }
}
```

Warning: While `x() { ... }` seems to just be shorthand for `x: function() { ... }`, concise methods have special behaviors that their older counterparts don't; specifically, the allowance for `super` (see "Object `super`" later in this chapter).

Generators (see Chapter 4) also have a concise method form:

```
var o = {
  *foo() { ... }
};
```

Concisely Unnamed While that convenience shorthand is quite attractive, there's a subtle gotcha to be aware of. To illustrate, let's examine pre-ES6 code like the following, which you might try to refactor to use concise methods:

```
function runSomething(o) {
  var x = Math.random(),
  y = Math.random();

  return o.something( x, y );
}

runSomething( {
  something: function something(x,y) {
    if (x > y) {
      // recursively call with `x`
      // and `y` swapped
      return something( y, x );
    }

    return y - x;
  }
});
```

This obviously silly code just generates two random numbers and subtracts the smaller from the bigger. But what's important here isn't what it does, but rather how it's defined. Let's focus on the object literal and function definition, as we see here:

```
runSomething( {
  something: function something(x,y) {
    // ...
  }
});
```

Why do we say both `something:` and `function something?` Isn't that redundant? Actually, no, both are needed for different purposes. The property

`something` is how we can call `o.something(..)`, sort of like its public name. But the second `something` is a lexical name to refer to the function from inside itself, for recursion purposes.

Can you see why the line `return something(y,x)` needs the name `something` to refer to the function? There's no lexical name for the object, such that it could have said `return o.something(y,x)` or something of that sort.

That's actually a pretty common practice when the object literal does have an identifying name, such as:

```
var controller = {
  makeRequest: function(..){
    // ...
    controller.makeRequest(..);
  }
};
```

Is this a good idea? Perhaps, perhaps not. You're assuming that the name `controller` will always point to the object in question. But it very well may not – the `makeRequest(..)` function doesn't control the outer code and so can't force that to be the case. This could come back to bite you.

Others prefer to use `this` to define such things:

```
var controller = {
  makeRequest: function(..){
    // ...
    this.makeRequest(..);
  }
};
```

That looks fine, and should work if you always invoke the method as `controller.makeRequest(..)`. But you now have a `this` binding gotcha if you do something like:

```
btn.addEventListener( "click", controller.makeRequest, false );
```

Of course, you can solve that by passing `controller.makeRequest.bind(controller)` as the handler reference to bind the event to. But yuck – it isn't very appealing.

Or what if your inner `this.makeRequest(..)` call needs to be made from a nested function? You'll have another `this` binding hazard, which people will often solve with the hacky `var self = this`, such as:

```
var controller = {
  makeRequest: function(..){
    var self = this;

    btn.addEventListener( "click", function(){
      // ...
    });
  }
};
```

```

        self.makeRequest(..);
    }, false );
}
};


```

More yuck.

Note: For more information on `this` binding rules and gotchas, see Chapters 1-2 of the *this & Object Prototypes* title of this series.

OK, what does all this have to do with concise methods? Recall our `something(..)` method definition:

```

runSomething( {
    something: function something(x,y) {
        // ...
    }
} );

```

The second `something` here provides a super convenient lexical identifier that will always point to the function itself, giving us the perfect reference for recursion, event binding/unbinding, and so on – no messing around with `this` or trying to use an untrustable object reference.

Great!

So, now we try to refactor that function reference to this ES6 concise method form:

```

runSomething( {
    something(x,y) {
        if (x > y) {
            return something( y, x );
        }

        return y - x;
    }
} );

```

Seems fine at first glance, except this code will break. The `return something(..)` call will not find a `something` identifier, so you'll get a `ReferenceError`. Oops. But why?

The above ES6 snippet is interpreted as meaning:

```

runSomething( {
    something: function(x,y){
        if (x > y) {
            return something( y, x );
        }
    }
} );

```

```

        return y - x;
    }
} );

```

Look closely. Do you see the problem? The concise method definition implies `something: function(x,y)`. See how the second `something` we were relying on has been omitted? In other words, concise methods imply anonymous function expressions.

Yeah, yuck.

Note: You may be tempted to think that `=>` arrow functions are a good solution here, but they're equally insufficient, as they're also anonymous function expressions. We'll cover them in "Arrow Functions" later in this chapter.

The partially redeeming news is that our `something(x,y)` concise method won't be totally anonymous. See "Function Names" in Chapter 7 for information about ES6 function name inference rules. That won't help us for our recursion, but it helps with debugging at least.

So what are we left to conclude about concise methods? They're short and sweet, and a nice convenience. But you should only use them if you're never going to need them to do recursion or event binding/unbinding. Otherwise, stick to your old-school `something: function something(..)` method definitions.

A lot of your methods are probably going to benefit from concise method definitions, so that's great news! Just be careful of the few where there's an un-naming hazard.

ES5 Getter/Setter Technically, ES5 defined getter/setter literals forms, but they didn't seem to get used much, mostly due to the lack of transpilers to handle that new syntax (the only major new syntax added in ES5, really). So while it's not a new ES6 feature, we'll briefly refresh on that form, as it's probably going to be much more useful with ES6 going forward.

Consider:

```

var o = {
    __id: 10,
    get id() { return this.__id++; },
    set id(v) { this.__id = v; }
}

o.id;           // 10
o.id;           // 11
o.id = 20;
o.id;           // 20

// and:

```

```
o.___id;           // 21
o.___id;           // 21 -- still!
```

These getter and setter literal forms are also present in classes; see Chapter 3.

Warning: It may not be obvious, but the setter literal must have exactly one declared parameter; omitting it or listing others is illegal syntax. The single required parameter *can* use destructuring and defaults (e.g., `set id({ id: v = 0 }) { ... }`), but the gather/rest `...` is not allowed (`set id(...v) { ... }`).

Computed Property Names

You've probably been in a situation like the following snippet, where you have one or more property names that come from some sort of expression and thus can't be put into the object literal:

```
var prefix = "user_";

var o = {
  baz: function(..){ ... }
};

o[ prefix + "foo" ] = function(..){ ... };
o[ prefix + "bar" ] = function(..){ ... };
..
```

ES6 adds a syntax to the object literal definition which allows you to specify an expression that should be computed, whose result is the property name assigned. Consider:

```
var prefix = "user_";

var o = {
  baz: function(..){ ... },
  [ prefix + "foo" ]: function(..){ ... },
  [ prefix + "bar" ]: function(..){ ... }
};
..
```

Any valid expression can appear inside the `[...]` that sits in the property name position of the object literal definition.

Probably the most common use of computed property names will be with `Symbols` (which we cover in “Symbols” later in this chapter), such as:

```
var o = {
  [Symbol.toStringTag]: "really cool thing",
  ..
};
```

`Symbol.toStringTag` is a special built-in value, which we evaluate with the `[..]` syntax, so we can assign the "really cool thing" value to the special property name.

Computed property names can also appear as the name of a concise method or a concise generator:

```
var o = {
  ["f" + "oo"]() { .. }  // computed concise method
  *[["b" + "ar"]() { .. } // computed concise generator
};
```

Setting `[[Prototype]]`

We won't cover prototypes in detail here, so for more information, see the *this & Object Prototypes* title of this series.

Sometimes it will be helpful to assign the `[[Prototype]]` of an object at the same time you're declaring its object literal. The following has been a non-standard extension in many JS engines for a while, but is standardized as of ES6:

```
var o1 = {
  // ..
};

var o2 = {
  __proto__: o1,
  // ..
};
```

`o2` is declared with a normal object literal, but it's also `[[Prototype]]`-linked to `o1`. The `__proto__` property name here can also be a string "`__proto__`", but note that it *cannot* be the result of a computed property name (see the previous section).

`__proto__` is controversial, to say the least. It's a decades-old proprietary extension to JS that is finally standardized, somewhat begrudgingly it seems, in ES6. Many developers feel it shouldn't ever be used. In fact, it's in "Annex B" of ES6, which is the section that lists things JS feels it has to standardize for compatibility reasons only.

Warning: Though I'm narrowly endorsing `__proto__` as a key in an object literal definition, I definitely do not endorse using it in its object property form, like `o.__proto__`. That form is both a getter and setter (again for compatibility reasons), but there are definitely better options. See the *this & Object Prototypes* title of this series for more information.

For setting the `[[Prototype]]` of an existing object, you can use the ES6 utility `Object.setPrototypeOf(..)`. Consider:

```

var o1 = {
  // ..
};

var o2 = {
  // ..
};

Object.setPrototypeOf( o2, o1 );

```

Note: We'll discuss `Object` again in Chapter 6. “`Object.setPrototypeOf(...)` Static Function” provides additional details on `Object.setPrototypeOf(...)`. Also see “`Object.assign(..)` Static Function” for another form that relates `o2` prototypically to `o1`.

Object super

`super` is typically thought of as being only related to classes. However, due to JS's classless-objects-with-prototypes nature, `super` is equally effective, and nearly the same in behavior, with plain objects' concise methods.

Consider:

```

var o1 = {
  foo() {
    console.log( "o1:foo" );
  }
};

var o2 = {
  foo() {
    super.foo();
    console.log( "o2:foo" );
  }
};

Object.setPrototypeOf( o2, o1 );

o2.foo();      // o1:foo
               // o2:foo

```

Warning: `super` is only allowed in concise methods, not regular function expression properties. It also is only allowed in `super.XXX` form (for property/method access), not in `super()` form.

The `super` reference in the `o2.foo()` method is locked statically to `o2`, and specifically to the `[[Prototype]]` of `o2`. `super` here would basically be `Object.getPrototypeOf(o2)` – resolves to `o1` of course – which is how it finds

and calls `o1.foo()`.

For complete details on `super`, see “Classes” in Chapter 3.

Template Literals

At the very outset of this section, I’m going to have to call out the name of this ES6 feature as being awfully... misleading, depending on your experiences with what the word *template* means.

Many developers think of templates as being reusable renderable pieces of text, such as the capability provided by most template engines (Mustache, Handlebars, etc.). ES6’s use of the word *template* would imply something similar, like a way to declare inline template literals that can be re-rendered. However, that’s not at all the right way to think about this feature.

So, before we go on, I’m renaming to what it should have been called: *interpolated string literals* (or *interpoliterals* for short).

You’re already well aware of declaring string literals with " or ' delimiters, and you also know that these are not *smart strings* (as some languages have), where the contents would be parsed for interpolation expressions.

However, ES6 introduces a new type of string literal, using the ` backtick as the delimiter. These string literals allow basic string interpolation expressions to be embedded, which are then automatically parsed and evaluated.

Here’s the old pre-ES6 way:

```
var name = "Kyle";  
  
var greeting = "Hello " + name + "!";  
  
console.log( greeting );           // "Hello Kyle!"  
console.log( typeof greeting );    // "string"
```

Now, consider the new ES6 way:

```
var name = "Kyle";  
  
var greeting = `Hello ${name}!`;  
  
console.log( greeting );           // "Hello Kyle!"  
console.log( typeof greeting );    // "string"
```

As you can see, we used the `...` around a series of characters, which are interpreted as a string literal, but any expressions of the form \${...} are parsed and evaluated inline immediately. The fancy term for such parsing and evaluating is *interpolation* (much more accurate than templating).

The result of the interpolated string literal expression is just a plain old normal string, assigned to the `greeting` variable.

Warning: `typeof greeting == "string"` illustrates why it's important not to think of these entities as special template values, as you cannot assign the unevaluated form of the literal to something and reuse it. The ``..`` string literal is more like an IIFE in the sense that it's automatically evaluated inline. The result of a ``..`` string literal is, simply, just a string.

One really nice benefit of interpolated string literals is they are allowed to split across multiple lines:

```
var text =
`Now is the time for all good men
to come to the aid of their
country!`;

console.log( text );
// Now is the time for all good men
// to come to the aid of their
// country!
```

The line breaks (newlines) in the interpolated string literal were preserved in the string value.

Unless appearing as explicit escape sequences in the literal value, the value of the `\r` carriage return character (code point U+000D) or the value of the `\r\n` carriage return + line feed sequence (code points U+000D and U+000A) are both normalized to a `\n` line feed character (code point U+000A). Don't worry though; this normalization is rare and would likely only happen if copy-pasting text into your JS file.

Interpolated Expressions

Any valid expression is allowed to appear inside `${..}` in an interpolated string literal, including function calls, inline function expression calls, and even other interpolated string literals!

Consider:

```
function upper(s) {
    return s.toUpperCase();
}

var who = "reader";

var text =
`A very ${upper( "warm" )} welcome
to all of you ${upper( `${who}s` )}!`;
```

```
console.log( text );
// A very WARM welcome
// to all of you READERS!
```

Here, the inner `\${who}s` interpolated string literal was a little bit nicer convenience for us when combining the `who` variable with the `"s"` string, as opposed to `who + "s"`. There will be cases that nesting interpolated string literals is helpful, but be wary if you find yourself doing that kind of thing often, or if you find yourself nesting several levels deep.

If that's the case, the odds are good that your string value production could benefit from some abstractions.

Warning: As a word of caution, be very careful about the readability of your code with such new found power. Just like with default value expressions and de-structuring assignment expressions, just because you *can* do something doesn't mean you *should* do it. Never go so overboard with new ES6 tricks that your code becomes more clever than you or your other team members.

Expression Scope One quick note about the scope that is used to resolve variables in expressions. I mentioned earlier that an interpolated string literal is kind of like an IIFE, and it turns out thinking about it like that explains the scoping behavior as well.

Consider:

```
function foo(str) {
  var name = "foo";
  console.log( str );
}

function bar() {
  var name = "bar";
  foo( `Hello from ${name}!` );
}

var name = "global";

bar();           // "Hello from bar!"
```

At the moment the `...` string literal is expressed, inside the `bar()` function, the scope available to it finds `bar()`'s `name` variable with value `"bar"`. Neither the global `name` nor `foo(..)`'s `name` matter. In other words, an interpolated string literal is just lexically scoped where it appears, not dynamically scoped in any way.

Tagged Template Literals

Again, renaming the feature for sanity sake: *tagged string literals*.

To be honest, this is one of the cooler tricks that ES6 offers. It may seem a little strange, and perhaps not all that generally practical at first. But once you've spent some time with it, tagged string literals may just surprise you in their usefulness.

For example:

```
function foo(strings, ...values) {
  console.log( strings );
  console.log( values );
}

var desc = "awesome";

foo`Everything is ${desc}!`;
// [ "Everything is ", "!" ]
// [ "awesome" ]
```

Let's take a moment to consider what's happening in the previous snippet. First, the most jarring thing that jumps out is `foo`Everything...``. That doesn't look like anything we've seen before. What is it?

It's essentially a special kind of function call that doesn't need the `(...)`. The *tag* – the `foo` part before the ``...`` string literal – is a function value that should be called. Actually, it can be any expression that results in a function, even a function call that returns another function, like:

```
function bar() {
  return function foo(strings, ...values) {
    console.log( strings );
    console.log( values );
  }
}

var desc = "awesome";

bar()`Everything is ${desc}!`;
// [ "Everything is ", "!" ]
// [ "awesome" ]
```

But what gets passed to the `foo(..)` function when invoked as a tag for a string literal?

The first argument – we called it `strings` – is an array of all the plain strings (the stuff between any interpolated expressions). We get two values in the `strings` array: `"Everything is "` and `"!"`.

For convenience sake in our example, we then gather up all subsequent arguments into an array called `values` using the `... gather/rest` operator (see the “Spread/Rest” section earlier in this chapter), though you could of course have left them as individual named parameters following the `strings` parameter.

The argument(s) gathered into our `values` array are the results of the already-evaluated interpolation expressions found in the string literal. So obviously the only element in `values` in our example is “`awesome`”.

You can think of these two arrays as: the values in `values` are the separators if you were to splice them in between the values in `strings`, and then if you joined everything together, you’d get the complete interpolated string value.

A tagged string literal is like a processing step after the interpolation expressions are evaluated but before the final string value is compiled, allowing you more control over generating the string from the literal.

Typically, the string literal tag function (`foo(..)` in the previous snippets) should compute an appropriate string value and return it, so that you can use the tagged string literal as a value just like untagged string literals:

```
function tag(strings, ...values) {
  return strings.reduce( function(s,v,idx){
    return s + (idx > 0 ? values[idx-1] : "") + v;
  }, "" );
}

var desc = "awesome";

var text = tag`Everything is ${desc}!`;

console.log( text );           // Everything is awesome!
```

In this snippet, `tag(..)` is a pass-through operation, in that it doesn’t perform any special modifications, but just uses `reduce(..)` to loop over and splice/interleave `strings` and `values` together the same way an untagged string literal would have done.

So what are some practical uses? There are many advanced ones that are beyond our scope to discuss here. But here’s a simple idea that formats numbers as U.S. dollars (sort of like basic localization):

```
function dollabillsyall(strings, ...values) {
  return strings.reduce( function(s,v,idx){
    if (idx > 0) {
      if (typeof values[idx-1] == "number") {
        // look, also using interpolated
        // string literals!
        s += `$$ ${values[idx-1].toFixed( 2 )}`;
      }
    }
  });
}
```

```

        else {
            s += values[idx-1];
        }
    }

    return s + v;
}, "") );
}

var amt1 = 11.99,
amt2 = amt1 * 1.08,
name = "Kyle";

var text = dollabillsyall
`Thanks for your purchase, ${name}! Your
product cost was ${amt1}, which with tax
comes out to ${amt2}.`

console.log( text );
// Thanks for your purchase, Kyle! Your
// product cost was $11.99, which with tax
// comes out to $12.95.

```

If a `number` value is encountered in the `values` array, we put "\$" in front of it and format it to two decimal places with `toFixed(2)`. Otherwise, we let the value pass-through untouched.

Raw Strings In the previous snippets, our tag functions receive the first argument we called `strings`, which is an array. But there's an additional bit of data included: the raw unprocessed versions of all the strings. You can access those raw string values using the `.raw` property, like this:

```

function showraw(strings, ...values) {
    console.log( strings );
    console.log( strings.raw );
}

showraw`Hello\nWorld`;
// [ "Hello
// World" ]
// [ "Hello\nWorld" ]

```

The raw version of the value preserves the raw escaped `\n` sequence (the `\` and the `n` are separate characters), while the processed version considers it a single newline character. However, the earlier mentioned line-ending normalization is applied to both values.

ES6 comes with a built-in function that can be used as a string literal tag: `String.raw(..)`. It simply passes through the raw versions of the `strings` values:

```
console.log(`Hello\nWorld`);  
// Hello  
// World  
  
console.log(String.raw`Hello\nWorld`);  
// Hello\nWorld  
  
String.raw`Hello\nWorld`.length;  
// 12
```

Other uses for string literal tags included special processing for internationalization, localization, and more!

Arrow Functions

We've touched on `this` binding complications with functions earlier in this chapter, and they're covered at length in the *this & Object Prototypes* title of this series. It's important to understand the frustrations that `this`-based programming with normal functions brings, because that is the primary motivation for the new ES6 `=>` arrow function feature.

Let's first illustrate what an arrow function looks like, as compared to normal functions:

```
function foo(x,y) {  
  return x + y;  
}  
  
// versus  
  
var foo = (x,y) => x + y;
```

The arrow function definition consists of a parameter list (of zero or more parameters, and surrounding `(..)` if there's not exactly one parameter), followed by the `=>` marker, followed by a function body.

So, in the previous snippet, the arrow function is just the `(x,y) => x + y` part, and that function reference happens to be assigned to the variable `foo`.

The body only needs to be enclosed by `{ .. }` if there's more than one expression, or if the body consists of a non-expression statement. If there's only one expression, and you omit the surrounding `{ .. }`, there's an implied `return` in front of the expression, as illustrated in the previous snippet.

Here's some other arrow function variations to consider:

```

var f1 = () => 12;
var f2 = x => x * 2;
var f3 = (x,y) => {
  var z = x * 2 + y;
  y++;
  x *= 3;
  return (x + y + z) / 2;
};

```

Arrow functions are *always* function expressions; there is no arrow function declaration. It also should be clear that they are anonymous function expressions – they have no named reference for the purposes of recursion or event binding/unbinding – though “Function Names” in Chapter 7 will describe ES6’s function name inference rules for debugging purposes.

Note: All the capabilities of normal function parameters are available to arrow functions, including default values, destructuring, rest parameters, and so on.

Arrow functions have a nice, shorter syntax, which makes them on the surface very attractive for writing terser code. Indeed, nearly all literature on ES6 (other than the titles in this series) seems to immediately and exclusively adopt the arrow function as “the new function.”

It is telling that nearly all examples in discussion of arrow functions are short single statement utilities, such as those passed as callbacks to various utilities. For example:

```

var a = [1,2,3,4,5];

a = a.map( v => v * 2 );

console.log( a );           // [2,4,6,8,10]

```

In those cases, where you have such inline function expressions, and they fit the pattern of computing a quick calculation in a single statement and returning that result, arrow functions indeed look to be an attractive and lightweight alternative to the more verbose **function** keyword and syntax.

Most people tend to *ooh and aah* at nice terse examples like that, as I imagine you just did!

However, I would caution you that it would seem to me somewhat a misapplication of this feature to use arrow function syntax with otherwise normal, multistatement functions, especially those that would otherwise be naturally expressed as function declarations.

Recall the **dollabillsyall(..)** string literal tag function from earlier in this chapter – let’s change it to use **=>** syntax:

```

var dollabillsyall = (strings, ...values) =>
  strings.reduce( (s,v,idx) => {

```

```

        if (idx > 0) {
            if (typeof values[idx-1] == "number") {
                // look, also using interpolated
                // string literals!
                s += `$$values[idx-1].toFixed( 2 )`;
            }
            else {
                s += values[idx-1];
            }
        }

        return s + v;
}, "" );

```

In this example, the only modifications I made were the removal of `function`, `return`, and some `{ ... }`, and then the insertion of `=>` and a `var`. Is this a significant improvement in the readability of the code? Meh.

I'd actually argue that the lack of `return` and outer `{ ... }` partially obscures the fact that the `reduce(..)` call is the only statement in the `dollabillsyall(..)` function and that its result is the intended result of the call. Also, the trained eye that is so used to hunting for the word `function` in code to find scope boundaries now needs to look for the `=>` marker, which can definitely be harder to find in the thick of the code.

While not a hard-and-fast rule, I'd say that the readability gains from `=>` arrow function conversion are inversely proportional to the length of the function being converted. The longer the function, the less `=>` helps; the shorter the function, the more `=>` can shine.

I think it's probably more sensible and reasonable to adopt `=>` for the places in code where you do need short inline function expressions, but leave your normal-length main functions as is.

Not Just Shorter Syntax, But this

Most of the popular attention toward `=>` has been on saving those precious keystrokes by dropping `function`, `return`, and `{ ... }` from your code.

But there's a big detail we've skipped over so far. I said at the beginning of the section that `=>` functions are closely related to `this` binding behavior. In fact, `=>` arrow functions are *primarily designed* to alter `this` behavior in a specific way, solving a particular and common pain point with `this`-aware coding.

The saving of keystrokes is a red herring, a misleading sideshow at best.

Let's revisit another example from earlier in this chapter:

```

var controller = {
  makeRequest: function(...){

```

```

var self = this;

btn.addEventListener( "click", function(){
    // ...
    self.makeRequest(..);
}, false );
}
;

```

We used the `var self = this` hack, and then referenced `self.makeRequest(..)`, because inside the callback function we're passing to `addEventListener(..)`, the `this` binding will not be the same as it is in `makeRequest(..)` itself. In other words, because `this` bindings are dynamic, we fall back to the predictability of lexical scope via the `self` variable.

Herein we finally can see the primary design characteristic of `=>` arrow functions. Inside arrow functions, the `this` binding is not dynamic, but is instead lexical. In the previous snippet, if we used an arrow function for the callback, `this` will be predictably what we wanted it to be.

Consider:

```

var controller = {
    makeRequest: function(..){
        btn.addEventListener( "click", () => {
            // ...
            this.makeRequest(..);
        }, false );
    }
};

```

Lexical `this` in the arrow function callback in the previous snippet now points to the same value as in the enclosing `makeRequest(..)` function. In other words, `=>` is a syntactic stand-in for `var self = this`.

In cases where `var self = this` (or, alternatively, a function `.bind(this)` call) would normally be helpful, `=>` arrow functions are a nicer alternative operating on the same principle. Sounds great, right?

Not quite so simple.

If `=>` replaces `var self = this` or `.bind(this)` and it helps, guess what happens if you use `=>` with a `this`-aware function that *doesn't* need `var self = this` to work? You might be able to guess that it's going to mess things up. Yeah.

Consider:

```

var controller = {
    makeRequest: (...) => {
        // ...
    }
};

```

```

        this.helper(..);
},
helper: (...) => {
  // ...
}
};

controller.makeRequest(..);

```

Although we invoke as `controller.makeRequest(..)`, the `this.helper` reference fails, because `this` here doesn't point to `controller` as it normally would. Where does it point? It lexically inherits `this` from the surrounding scope. In this previous snippet, that's the global scope, where `this` points to the global object. Ugh.

In addition to lexical `this`, arrow functions also have lexical `arguments` – they don't have their own `arguments` array but instead inherit from their parent – as well as lexical `super` and `new.target` (see “Classes” in Chapter 3).

So now we can conclude a more nuanced set of rules for when `=>` is appropriate and not:

- If you have a short, single-statement inline function expression, where the only statement is a `return` of some computed value, *and* that function doesn't already make a `this` reference inside it, *and* there's no self-reference (recursion, event binding/unbinding), *and* you don't reasonably expect the function to ever be that way, you can probably safely refactor it to be an `=>` arrow function.
- If you have an inner function expression that's relying on a `var self = this` hack or a `.bind(this)` call on it in the enclosing function to ensure proper `this` binding, that inner function expression can probably safely become an `=>` arrow function.
- If you have an inner function expression that's relying on something like `var args = Array.prototype.slice.call(arguments)` in the enclosing function to make a lexical copy of `arguments`, that inner function expression can probably safely become an `=>` arrow function.
- For everything else – normal function declarations, longer multistatement function expressions, functions that need a lexical name identifier self-reference (recursion, etc.), and any other function that doesn't fit the previous characteristics – you should probably avoid `=>` function syntax.

Bottom line: `=>` is about lexical binding of `this`, `arguments`, and `super`. These are intentional features designed to fix some common problems, not bugs, quirks, or mistakes in ES6.

Don't believe any hype that `=>` is primarily, or even mostly, about fewer keystrokes. Whether you save keystrokes or waste them, you should know exactly what you are intentionally doing with every character typed.

Tip: If you have a function that for any of these articulated reasons is not a good match for an `=>` arrow function, but it's being declared as part of an object literal, recall from “Concise Methods” earlier in this chapter that there's another option for shorter function syntax.

If you prefer a visual decision chart for how/why to pick an arrow function:

for..of Loops

Joining the `for` and `for..in` loops from the JavaScript we're all familiar with, ES6 adds a `for..of` loop, which loops over the set of values produced by an *iterator*.

The value you loop over with `for..of` must be an *iterable*, or it must be a value which can be coerced/boxed to an object (see the *Types & Grammar* title of this series) that is an iterable. An iterable is simply an object that is able to produce an iterator, which the loop then uses.

Let's compare `for..of` to `for..in` to illustrate the difference:

```
var a = ["a", "b", "c", "d", "e"];

for (var idx in a) {
  console.log( idx );
}

// 0 1 2 3 4

for (var val of a) {
  console.log( val );
}

// "a" "b" "c" "d" "e"
```

As you can see, `for..in` loops over the keys/indexes in the `a` array, while `for..of` loops over the values in `a`.

Here's the pre-ES6 version of the `for..of` from that previous snippet:

```
var a = ["a", "b", "c", "d", "e"],
  k = Object.keys( a );

for (var val, i = 0; i < k.length; i++) {
  val = a[ k[i] ];
  console.log( val );
}

// "a" "b" "c" "d" "e"
```

And here's the ES6 but non-`for..of` equivalent, which also gives a glimpse at manually iterating an iterator (see “Iterators” in Chapter 3):

```

var a = ["a", "b", "c", "d", "e"];

for (var val, ret, it = a[Symbol.iterator]());  

  (ret = it.next()) && !ret.done;  

) {  

  val = ret.value;  

  console.log( val );  

}  

// "a" "b" "c" "d" "e"

```

Under the covers, the `for..of` loop asks the iterable for an iterator (using the built-in `Symbol.iterator`; see “Well-Known Symbols” in Chapter 7), then it repeatedly calls the iterator and assigns its produced value to the loop iteration variable.

Standard built-in values in JavaScript that are by default iterables (or provide them) include:

- Arrays
- Strings
- Generators (see Chapter 3)
- Collections / TypedArrays (see Chapter 5)

Warning: Plain objects are not by default suitable for `for..of` looping. That’s because they don’t have a default iterator, which is intentional, not a mistake. However, we won’t go any further into those nuanced reasonings here. In “Iterators” in Chapter 3, we’ll see how to define iterators for our own objects, which lets `for..of` loop over any object to get a set of values we define.

Here’s how to loop over the characters in a primitive string:

```

for (var c of "hello") {  

  console.log( c );  

}  

// "h" "e" "l" "l" "o"

```

The “hello” primitive string value is coerced/boxed to the `String` object wrapper equivalent, which is an iterable by default.

In `for (XYZ of ABC)...`, the `XYZ` clause can either be an assignment expression or a declaration, identical to that same clause in `for` and `for..in` loops. So you can do stuff like this:

```

var o = {};  

for (o.a of [1,2,3]) {  

  console.log( o.a );  

}  

// 1 2 3

```

```

for ({x: o.a} of [ {x: 1}, {x: 2}, {x: 3} ]) {
  console.log( o.a );
}
// 1 2 3

```

`for..of` loops can be prematurely stopped, just like other loops, with `break`, `continue`, `return` (if in a function), and thrown exceptions. In any of these cases, the iterator's `return(..)` function is automatically called (if one exists) to let the iterator perform cleanup tasks, if necessary.

Note: See “Iterators” in Chapter 3 for more complete coverage on iterables and iterators.

Regular Expressions

Let's face it: regular expressions haven't changed much in JS in a long time. So it's a great thing that they've finally learned a couple of new tricks in ES6. We'll briefly cover the additions here, but the overall topic of regular expressions is so dense that you'll need to turn to chapters/books dedicated to it (of which there are many!) if you need a refresher.

Unicode Flag

We'll cover the topic of Unicode in more detail in “Unicode” later in this chapter. Here, we'll just look briefly at the new `u` flag for ES6+ regular expressions, which turns on Unicode matching for that expression.

JavaScript strings are typically interpreted as sequences of 16-bit characters, which correspond to the characters in the *Basic Multilingual Plane (BMP)* (http://en.wikipedia.org/wiki/Plane_%28Unicode%29). But there are many UTF-16 characters that fall outside this range, and so strings may have these multibyte characters in them.

Prior to ES6, regular expressions could only match based on BMP characters, which means that those extended characters were treated as two separate characters for matching purposes. This is often not ideal.

So, as of ES6, the `u` flag tells a regular expression to process a string with the interpretation of Unicode (UTF-16) characters, such that such an extended character will be matched as a single entity.

Warning: Despite the name implication, “UTF-16” doesn't strictly mean 16 bits. Modern Unicode uses 21 bits, and standards like UTF-8 and UTF-16 refer roughly to how many bits are used in the representation of a character.

An example (straight from the ES6 specification): (the musical symbol G-clef) is Unicode point U+1D11E (0x1D11E).

If this character appears in a regular expression pattern (like `/ /`), the standard BMP interpretation would be that it's two separate characters (0xD834 and

0xDD1E) to match with. But the new ES6 Unicode-aware mode means that `/u` (or the escaped Unicode form `/u{1D11E}/u`) will match " " in a string as a single matched character.

You might be wondering why this matters? In non-Unicode BMP mode, the pattern is treated as two separate characters, but would still find the match in a string with the " " character in it, as you can see if you try:

```
/ .test( "-clef" );           // true
```

The length of the match is what matters. For example:

```
/^.-clef/ .test( "-clef" );           // false
/^.-clef/u.test( "-clef" );           // true
```

The `^.-clef` in the pattern says to match only a single character at the beginning before the normal `"-clef"` text. In standard BMP mode, the match fails (two characters), but with `u` Unicode mode flagged on, the match succeeds (one character).

It's also important to note that `u` makes quantifiers like `+` and `*` apply to the entire Unicode code point as a single character, not just the *lower surrogate* (aka rightmost half of the symbol) of the character. The same goes for Unicode characters appearing in character classes, like `/[-]/u`.

Note: There's plenty more nitty-gritty details about `u` behavior in regular expressions, which Mathias Bynens (<https://twitter.com/mathias>) has written extensively about (<https://mathiasbynens.be/notes/es6-unicode-regex>).

Sticky Flag

Another flag mode added to ES6 regular expressions is `y`, which is often called "sticky mode." *Sticky* essentially means the regular expression has a virtual anchor at its beginning that keeps it rooted to matching at only the position indicated by the regular expression's `lastIndex` property.

To illustrate, let's consider two regular expressions, the first without sticky mode and the second with:

```
var re1 = /foo/,
    str = "++foo++";

re1.lastIndex;           // 0
re1.test( str );        // true
re1.lastIndex;           // 0 -- not updated

re1.lastIndex = 4;
re1.test( str );        // true -- ignored `lastIndex`
re1.lastIndex;           // 4 -- not updated
```

Three things to observe about this snippet:

- `test(..)` doesn't pay any attention to `lastIndex`'s value, and always just performs its match from the beginning of the input string.
- Because our pattern does not have a `^` start-of-input anchor, the search for "foo" is free to move ahead through the whole string looking for a match.
- `lastIndex` is not updated by `test(..)`.

Now, let's try a sticky mode regular expression:

```
var re2 = /foo/y,           // <-- notice the `y` sticky flag
str = "++foo++";

re2.lastIndex;             // 0
re2.test( str );          // false -- "foo" not found at `0`
re2.lastIndex;             // 0

re2.lastIndex = 2;
re2.test( str );          // true
re2.lastIndex;             // 5 -- updated to after previous match

re2.test( str );          // false
re2.lastIndex;             // 0 -- reset after previous match failure
```

And so our new observations about sticky mode:

- `test(..)` uses `lastIndex` as the exact and only position in `str` to look to make a match. There is no moving ahead to look for the match – it's either there at the `lastIndex` position or not.
- If a match is made, `test(..)` updates `lastIndex` to point to the character immediately following the match. If a match fails, `test(..)` resets `lastIndex` back to 0.

Normal non-sticky patterns that aren't otherwise `^`-rooted to the start-of-input are free to move ahead in the input string looking for a match. But sticky mode restricts the pattern to matching just at the position of `lastIndex`.

As I suggested at the beginning of this section, another way of looking at this is that `y` implies a virtual anchor at the beginning of the pattern that is relative (aka constrains the start of the match) to exactly the `lastIndex` position.

Warning: In previous literature on the topic, it has alternatively been asserted that this behavior is like `y` implying a `^` (start-of-input) anchor in the pattern. This is inaccurate. We'll explain in further detail in "Anchored Sticky" later.

Sticky Positioning It may seem strangely limiting that to use `y` for repeated matches, you have to manually ensure `lastIndex` is in the exact right position, as it has no move-ahead capability for matching.

Here's one possible scenario: if you know that the match you care about is always

going to be at a position that's a multiple of a number (e.g., 0, 10, 20, etc.), you can just construct a limited pattern matching what you care about, but then manually set `lastIndex` each time before match to those fixed positions.

Consider:

```
var re = /f..y,
      str = "foo      far      fad";

str.match( re );           // ["foo"]

re.lastIndex = 10;
str.match( re );           // ["far"]

re.lastIndex = 20;
str.match( re );           // ["fad"]
```

However, if you're parsing a string that isn't formatted in fixed positions like that, figuring out what to set `lastIndex` to before each match is likely going to be untenable.

There's a saving nuance to consider here. `y` requires that `lastIndex` be in the exact position for a match to occur. But it doesn't strictly require that *you* manually set `lastIndex`.

Instead, you can construct your expressions in such a way that they capture in each main match everything before and after the thing you care about, up to right before the next thing you'll care to match.

Because `lastIndex` will set to the next character beyond the end of a match, if you've matched everything up to that point, `lastIndex` will always be in the correct position for the `y` pattern to start from the next time.

Warning: If you can't predict the structure of the input string in a sufficiently patterned way like that, this technique may not be suitable and you may not be able to use `y`.

Having structured string input is likely the most practical scenario where `y` will be capable of performing repeated matching throughout a string. Consider:

```
var re = /\d+\.\s(.*?)\s(?:\s|$)/y
      str = "1. foo 2. bar 3. baz";

str.match( re );           // [ "1. foo ", "foo" ]

re.lastIndex;              // 7 -- correct position!
str.match( re );           // [ "2. bar ", "bar" ]

re.lastIndex;              // 14 -- correct position!
str.match( re );           // [ "3. baz", "baz" ]
```

This works because I knew something ahead of time about the structure of the input string: there is always a numeral prefix like "1. " before the desired match ("foo", etc.), and either a space after it, or the end of the string (\$ anchor). So the regular expression I constructed captures all of that in each main match, and then I use a matching group () so that the stuff I really care about is separated out for convenience.

After the first match ("1. foo "), the `lastIndex` is 7, which is already the position needed to start the next match, for "2. bar ", and so on.

If you're going to use `y` sticky mode for repeated matches, you'll probably want to look for opportunities to have `lastIndex` automatically positioned as we've just demonstrated.

Sticky Versus Global Some readers may be aware that you can emulate something like this `lastIndex`-relative matching with the `g` global match flag and the `exec(..)` method, as so:

```
var re = /o+./g,           // <-- look, `g`!
      str = "foot book more";

re.exec( str );           // ["oot"]
re.lastIndex;             // 4

re.exec( str );           // ["ook"]
re.lastIndex;             // 9

re.exec( str );           // ["or"]
re.lastIndex;             // 13

re.exec( str );           // null -- no more matches!
re.lastIndex;             // 0 -- starts over now!
```

While it's true that `g` pattern matches with `exec(..)` start their matching from `lastIndex`'s current value, and also update `lastIndex` after each match (or failure), this is not the same thing as `y`'s behavior.

Notice in the previous snippet that "ook", located at position 6, was matched and found by the second `exec(..)` call, even though at the time, `lastIndex` was 4 (from the end of the previous match). Why? Because as we said earlier, non-sticky matches are free to move ahead in their matching. A sticky mode expression would have failed here, because it would not be allowed to move ahead.

In addition to perhaps undesired move-ahead matching behavior, another downside to just using `g` instead of `y` is that `g` changes the behavior of some matching methods, like `str.match(re)`.

Consider:

```

var re = /o+./g,           // <-- look, `g`!
str = "foot book more";

str.match( re );           // ["oot", "ook", "or"]

```

See how all the matches were returned at once? Sometimes that's OK, but sometimes that's not what you want.

The `y` sticky flag will give you one-at-a-time progressive matching with utilities like `test(..)` and `match(..)`. Just make sure the `lastIndex` is always in the right position for each match!

Anchored Sticky As we warned earlier, it's inaccurate to think of sticky mode as implying a pattern starts with `^`. The `^` anchor has a distinct meaning in regular expressions, which is *not altered* by sticky mode. `^` is an anchor that *always* refers to the beginning of the input, and *is not* in any way relative to `lastIndex`.

Besides poor/inaccurate documentation on this topic, the confusion is unfortunately strengthened further because an older pre-ES6 experiment with sticky mode in Firefox *did* make `^` relative to `lastIndex`, so that behavior has been around for years.

ES6 elected not to do it that way. `^` in a pattern means start-of-input absolutely and only.

As a consequence, a pattern like `/^foo/y` will always and only find a "foo" match at the beginning of a string, *if it's allowed to match there*. If `lastIndex` is not 0, the match will fail. Consider:

```

var re = /^fooy,
str = "foo";

re.test( str );           // true
re.test( str );           // false
re.lastIndex;           // 0 -- reset after failure

re.lastIndex = 1;
re.test( str );           // false -- failed for positioning
re.lastIndex;           // 0 -- reset after failure

```

Bottom line: `y` plus `^` plus `lastIndex > 0` is an incompatible combination that will always cause a failed match.

Note: While `y` does not alter the meaning of `^` in any way, the `m` multiline mode *does*, such that `^` means start-of-input *or* start of text after a newline. So, if you combine `y` and `m` flags together for a pattern, you can find multiple `^`-rooted matches in a string. But remember: because it's `y` sticky, you'll have to make sure `lastIndex` is pointing at the correct new line position (likely by matching

to the end of the line) each subsequent time, or no subsequent matches will be made.

Regular Expression flags

Prior to ES6, if you wanted to examine a regular expression object to see what flags it had applied, you needed to parse them out – ironically, probably with another regular expression – from the content of the `source` property, such as:

```
var re = /foo/ig;

re.toString();           // "/foo/ig"

var flags = re.toString().match( /\/(.*?)([gim]*$)/ )[1];

flags;                  // "ig"
```

As of ES6, you can now get these values directly, with the new `flags` property:

```
var re = /foo/ig;

re.flags;               // "gi"
```

It's a small nuance, but the ES6 specification calls for the expression's flags to be listed in this order: "gimuy", regardless of what order the original pattern was specified with. That's the reason for the difference between `/ig` and "gi".

No, the order of flags specified or listed doesn't matter.

Another tweak from ES6 is that the `RegExp(..)` constructor is now flags-aware if you pass it an existing regular expression:

```
var re1 = /foo*/y;
re1.source;           // "foo*"
re1.flags;            // "y"

var re2 = new RegExp( re1 );
re2.source;           // "foo*"
re2.flags;            // "y"

var re3 = new RegExp( re1, "ig" );
re3.source;           // "foo*"
re3.flags;            // "gi"
```

Prior to ES6, the `re3` construction would throw an error, but as of ES6 you can override the flags when duplicating.

Number Literal Extensions

Prior to ES5, number literals looked like the following – the octal form was not officially specified, only allowed as an extension that browsers had come to de facto agreement on:

```
var dec = 42,  
    oct = 052,  
    hex = 0x2a;
```

Note: Though you are specifying a number in different bases, the number's mathematic value is what is stored, and the default output interpretation is always base-10. The three variables in the previous snippet all have the 42 value stored in them.

To further illustrate that 052 was a nonstandard form extension, consider:

```
Number( "42" );           // 42  
Number( "052" );          // 52  
Number( "0x2a" );         // 42
```

ES5 continued to permit the browser-extended octal form (including such inconsistencies), except that in strict mode, the octal literal (052) form is disallowed. This restriction was done mainly because many developers had the habit (from other languages) of seemingly innocuously prefixing otherwise base-10 numbers with 0's for code alignment purposes, and then running into the accidental fact that they'd changed the number value entirely!

ES6 continues the legacy of changes/variations to how number literals outside base-10 numbers can be represented. There's now an official octal form, an amended hexadecimal form, and a brand-new binary form. For web compatibility reasons, the old octal 052 form will continue to be legal (though unspecified) in non-strict mode, but should really never be used anymore.

Here are the new ES6 number literal forms:

```
var dec = 42,  
    oct = 0o52,           // or `0052` :/  
    hex = 0x2a,           // or `0X2a` :/  
    bin = 0b101010;       // or `0B101010` :/
```

The only decimal form allowed is base-10. Octal, hexadecimal, and binary are all integer forms.

And the string representations of these forms are all able to be coerced/converted to their number equivalent:

```
Number( "42" );           // 42  
Number( "0o52" );          // 42  
Number( "0x2a" );          // 42  
Number( "0b101010" );       // 42
```

Though not strictly new to ES6, it's a little-known fact that you can actually go the opposite direction of conversion (well, sort of):

```
var a = 42;

a.toString();           // "42" -- also `a.toString( 10 )`-
a.toString( 8 );       // "52"
a.toString( 16 );      // "2a"
a.toString( 2 );       // "101010"
```

In fact, you can represent a number this way in any base from 2 to 36, though it'd be rare that you'd go outside the standard bases: 2, 8, 10, and 16.

Unicode

Let me just say that this section is not an exhaustive everything-you-ever-wanted-to-know-about-Unicode resource. I want to cover what you need to know that's *changing* for Unicode in ES6, but we won't go much deeper than that. Mathias Bynens (<http://twitter.com/mathias>) has written/spoken extensively and brilliantly about JS and Unicode (see <https://mathiasbynens.be/notes/javascript-unicode> and <http://fluentconf.com/javascript-html-2015/public/content/2015/02/18-javascript-loves-unicode>).

The Unicode characters that range from 0x0000 to 0xFFFF contain all the standard printed characters (in various languages) that you're likely to have seen or interacted with. This group of characters is called the *Basic Multilingual Plane (BMP)*. The BMP even contains fun symbols like this cool snowman: (U+2603).

There are lots of other extended Unicode characters beyond this BMP set, which range up to 0x10FFFF. These symbols are often referred to as *astral* symbols, as that's the name given to the set of 16 *planes* (e.g., layers/groupings) of characters beyond the BMP. Examples of astral symbols include (U+1D11E) and (U+1F4A9).

Prior to ES6, JavaScript strings could specify Unicode characters using Unicode escaping, such as:

```
var snowman = "\u2603";
console.log( snowman );           // "
```

However, the \uXXXX Unicode escaping only supports four hexadecimal characters, so you can only represent the BMP set of characters in this way. To represent an astral character using Unicode escaping prior to ES6, you need to use a *surrogate pair* – basically two specially calculated Unicode-escaped characters side by side, which JS interprets together as a single astral character:

```
var gclef = "\uD834\uDD1E";
console.log( gclef );           // "
```

As of ES6, we now have a new form for Unicode escaping (in strings and regular expressions), called Unicode *code point escaping*:

```
var gclef = "\u{1D11E}";
console.log( gclef ); // "
```

As you can see, the difference is the presence of the `{ }` in the escape sequence, which allows it to contain any number of hexadecimal characters. Because you only need six to represent the highest possible code point value in Unicode (i.e., `0x10FFFF`), this is sufficient.

Unicode-Aware String Operations

By default, JavaScript string operations and methods are not sensitive to astral symbols in string values. So, they treat each BMP character individually, even the two surrogate halves that make up an otherwise single astral character. Consider:

```
var snowman = " ";
snowman.length; // 1

var gclef = " ";
gclef.length; // 2
```

So, how do we accurately calculate the length of such a string? In this scenario, the following trick will work:

```
var gclef = " ";

[...gclef].length; // 1
Array.from( gclef ).length; // 1
```

Recall from the “`for..of` Loops” section earlier in this chapter that ES6 strings have built-in iterators. This iterator happens to be Unicode-aware, meaning it will automatically output an astral symbol as a single value. We take advantage of that using the `...` spread operator in an array literal, which creates an array of the string’s symbols. Then we just inspect the length of that resultant array. ES6’s `Array.from(..)` does basically the same thing as `[...XYZ]`, but we’ll cover that utility in detail in Chapter 6.

Warning: It should be noted that constructing and exhausting an iterator just to get the length of a string is quite expensive on performance, relatively speaking, compared to what a theoretically optimized native utility/property would do.

Unfortunately, the full answer is not as simple or straightforward. In addition to the surrogate pairs (which the string iterator takes care of), there are special Unicode code points that behave in other special ways, which is much harder to account for. For example, there’s a set of code points that modify the previous adjacent character, known as *Combining Diacritical Marks*.

Consider these two string outputs:

```
console.log( s1 );           // "é"
console.log( s2 );           // "é"
```

They look the same, but they're not! Here's how we created `s1` and `s2`:

```
var s1 = "\xE9",
s2 = "e\u0301";
```

As you can probably guess, our previous `length` trick doesn't work with `s2`:

```
[...s1].length;           // 1
[...s2].length;           // 2
```

So what can we do? In this case, we can perform a *Unicode normalization* on the value before inquiring about its length, using the ES6 `String#normalize(..)` utility (which we'll cover more in Chapter 6):

```
var s1 = "\xE9",
s2 = "e\u0301";

s1.normalize().length;      // 1
s2.normalize().length;      // 1

s1 === s2;                // false
s1 === s2.normalize();     // true
```

Essentially, `normalize(..)` takes a sequence like `"e\u0301"` and normalizes it to `"\xE9"`. Normalization can even combine multiple adjacent combining marks if there's a suitable Unicode character they combine to:

```
var s1 = "o\u0302\u0300",
s2 = s1.normalize(),
s3 = "ò";

s1.length;                // 3
s2.length;                // 1
s3.length;                // 1

s2 === s3;                // true
```

Unfortunately, normalization isn't fully perfect here, either. If you have multiple combining marks modifying a single character, you may not get the length count you'd expect, because there may not be a single defined normalized character that represents the combination of all the marks. For example:

```
var s1 = "e\u0301\u0330";

console.log( s1 );         // "é"
```

```
s1.normalize().length;           // 2
```

The further you go down this rabbit hole, the more you realize that it's difficult to get one precise definition for "length." What we see visually rendered as a single character – more precisely called a *grapheme* – doesn't always strictly relate to a single "character" in the program processing sense.

Tip: If you want to see just how deep this rabbit hole goes, check out the "Grapheme Cluster Boundaries" algorithm (http://www.unicode.org/reports/tr29/#Grapheme_Cluster_Boundaries)

Character Positioning

Similar to length complications, what does it actually mean to ask, "what is the character at position 2?" The naive pre-ES6 answer comes from `charAt(..)`, which will not respect the atomicity of an astral character, nor will it take into account combining marks.

Consider:

```
var s1 = "abc\u0301d",
  s2 = "ab\u0107d",
  s3 = "ab\u{1d49e}d";

console.log( s1 );           // "ab d"
console.log( s2 );           // "ab d"
console.log( s3 );           // "ab d"

s1.charAt( 2 );              // "c"
s2.charAt( 2 );              // " "
s3.charAt( 2 );              // "" <-- unprintable surrogate
s3.charAt( 3 );              // "" <-- unprintable surrogate
```

So, is ES6 giving us a Unicode-aware version of `charAt(..)`? Unfortunately, no. At the time of this writing, there's a proposal for such a utility that's under consideration for post-ES6.

But with what we explored in the previous section (and of course with the limitations noted thereof!), we can hack an ES6 answer:

```
var s1 = "abc\u0301d",
  s2 = "ab\u0107d",
  s3 = "ab\u{1d49e}d";

[...s1.normalize()][2];        // " "
[...s2.normalize()][2];        // " "
[...s3.normalize()][2];        // ""
```

Warning: Reminder of an earlier warning: constructing and exhausting an iterator each time you want to get at a single character is... not very ideal,

performance wise. Let's hope we get a built-in and optimized utility for this soon, post-ES6.

What about a Unicode-aware version of the `charCodeAt(..)` utility? ES6 gives us `codePointAt(..)`:

```
var s1 = "abc\u0301d",
    s2 = "ab\u0107d",
    s3 = "ab\u{1d49e}d";

s1.normalize().codePointAt( 2 ).toString( 16 );
// "107"

s2.normalize().codePointAt( 2 ).toString( 16 );
// "107"

s3.normalize().codePointAt( 2 ).toString( 16 );
// "1d49e"
```

What about the other direction? A Unicode-aware version of `String.fromCharCode(..)` is ES6's `String.fromCodePoint(..)`:

```
String.fromCodePoint( 0x107 );      // "ć"

String.fromCodePoint( 0x1d49e );    // "
```

So wait, can we just combine `String.fromCodePoint(..)` and `codePointAt(..)` to get a better version of a Unicode-aware `charAt(..)` from earlier? Yep!

```
var s1 = "abc\u0301d",
    s2 = "ab\u0107d",
    s3 = "ab\u{1d49e}d";

String.fromCodePoint( s1.normalize().codePointAt( 2 ) );
// "ć"

String.fromCodePoint( s2.normalize().codePointAt( 2 ) );
// "ć"

String.fromCodePoint( s3.normalize().codePointAt( 2 ) );
// "
```

There's quite a few other string methods we haven't addressed here, including `toUpperCase()`, `toLowerCase()`, `substring(..)`, `indexOf(..)`, `slice(..)`, and a dozen others. None of these have been changed or augmented for full Unicode awareness, so you should be very careful – probably just avoid them! – when working with strings containing astral symbols.

There are also several string methods that use regular expressions for their behavior, like `replace(..)` and `match(..)`. Thankfully, ES6 brings Unicode

awareness to regular expressions, as we covered in “Unicode Flag” earlier in this chapter.

OK, there we have it! JavaScript’s Unicode string support is significantly better over pre-ES6 (though still not perfect) with the various additions we’ve just covered.

Unicode Identifier Names

Unicode can also be used in identifier names (variables, properties, etc.). Prior to ES6, you could do this with Unicode-escapes, like:

```
var \u03A9 = 42;  
// same as: var Ω = 42;
```

As of ES6, you can also use the earlier explained code point escape syntax:

```
var \u{2B400} = 42;  
// same as: var = 42;
```

There’s a complex set of rules around exactly which Unicode characters are allowed. Furthermore, some are allowed only if they’re not the first character of the identifier name.

Note: Mathias Bynens has a great post (<https://mathiasbynens.be/notes/javascript-identifiers-es6>) on all the nitty-gritty details.

The reasons for using such unusual characters in identifier names are rather rare and academic. You typically won’t be best served by writing code that relies on these esoteric capabilities.

Symbols

With ES6, for the first time in quite a while, a new primitive type has been added to JavaScript: the `symbol`. Unlike the other primitive types, however, symbols don’t have a literal form.

Here’s how you create a symbol:

```
var sym = Symbol( "some optional description" );  
  
typeof sym; // "symbol"
```

Some things to note:

- You cannot and should not use `new` with `Symbol(..)`. It’s not a constructor, nor are you producing an object.
- The parameter passed to `Symbol(..)` is optional. If passed, it should be a string that gives a friendly description for the symbol’s purpose.

- The `typeof` output is a new value ("symbol") that is the primary way to identify a symbol.

The description, if provided, is solely used for the stringification representation of the symbol:

```
sym.toString();      // "Symbol(some optional description)"
```

Similar to how primitive string values are not instances of `String`, symbols are also not instances of `Symbol`. If, for some reason, you want to construct a boxed wrapper object form of a symbol value, you can do the following:

```
sym instanceof Symbol;      // false

var symObj = Object( sym );
symObj instanceof Symbol;  // true

symObj.valueOf() === sym;  // true
```

Note: `symObj` in this snippet is interchangeable with `sym`; either form can be used in all places symbols are utilized. There's not much reason to use the boxed wrapper object form (`symObj`) instead of the primitive form (`sym`). Keeping with similar advice for other primitives, it's probably best to prefer `sym` over `symObj`.

The internal value of a symbol itself – referred to as its `name` – is hidden from the code and cannot be obtained. You can think of this symbol value as an automatically generated, unique (within your application) string value.

But if the value is hidden and unobtainable, what's the point of having a symbol at all?

The main point of a symbol is to create a string-like value that can't collide with any other value. So, for example, consider using a symbol as a constant representing an event name:

```
const EVT_LOGIN = Symbol( "event.login" );
```

You'd then use `EVT_LOGIN` in place of a generic string literal like `"event.login"`:

```
evthub.listen( EVT_LOGIN, function(data){
  // ...
} );
```

The benefit here is that `EVT_LOGIN` holds a value that cannot be duplicated (accidentally or otherwise) by any other value, so it is impossible for there to be any confusion of which event is being dispatched or handled.

Note: Under the covers, the `evthub` utility assumed in the previous snippet would almost certainly be using the symbol value from the `EVT_LOGIN` argument directly as the property/key in some internal object (hash) that tracks event handlers. If `evthub` instead needed to use the symbol value as a real string, it

would need to explicitly coerce with `String(..)` or `toString()`, as implicit string coercion of symbols is not allowed.

You may use a symbol directly as a property name/key in an object, such as a special property that you want to treat as hidden or meta in usage. It's important to know that although you intend to treat it as such, it is not *actually* a hidden or untouchable property.

Consider this module that implements the *singleton* pattern behavior – that is, it only allows itself to be created once:

```
const INSTANCE = Symbol( "instance" );

function HappyFace() {
  if (HappyFace[INSTANCE]) return HappyFace[INSTANCE];

  function smile() { ... }

  return HappyFace[INSTANCE] = {
    smile: smile
  };
}

var me = HappyFace(),
  you = HappyFace();

me === you; // true
```

The `INSTANCE` symbol value here is a special, almost hidden, meta-like property stored statically on the `HappyFace()` function object.

It could alternatively have been a plain old property like `__instance`, and the behavior would have been identical. The usage of a symbol simply improves the metaprogramming style, keeping this `INSTANCE` property set apart from any other normal properties.

Symbol Registry

One mild downside to using symbols as in the last few examples is that the `EVT_LOGIN` and `INSTANCE` variables had to be stored in an outer scope (perhaps even the global scope), or otherwise somehow stored in a publicly available location, so that all parts of the code that need to use the symbols can access them.

To aid in organizing code with access to these symbols, you can create symbol values with the *global symbol registry*. For example:

```
const EVT_LOGIN = Symbol.for( "event.login" );
```

```
console.log( EVT_LOGIN );           // Symbol(event.login)
```

And:

```
function HappyFace() {
  const INSTANCE = Symbol.for( "instance" );

  if (HappyFace[INSTANCE]) return HappyFace[INSTANCE];

  // ...

  return HappyFace[INSTANCE] = { ... };
}
```

`Symbol.for(..)` looks in the global symbol registry to see if a symbol is already stored with the provided description text, and returns it if so. If not, it creates one to return. In other words, the global symbol registry treats symbol values, by description text, as singletons themselves.

But that also means that any part of your application can retrieve the symbol from the registry using `Symbol.for(..)`, as long as the matching description name is used.

Ironically, symbols are basically intended to replace the use of *magic strings* (arbitrary string values given special meaning) in your application. But you precisely use *magic* description string values to uniquely identify/locate them in the global symbol registry!

To avoid accidental collisions, you'll probably want to make your symbol descriptions quite unique. One easy way of doing that is to include prefix/context/namespacing information in them.

For example, consider a utility such as the following:

```
function extractValues(str) {
  var key = Symbol.for( "extractValues.parse" ),
      re = extractValues[key] ||
            /[^=&]+=?([^\&]+?)(&|=|$)/g,
      values = [], match;

  while (match = re.exec( str )) {
    values.push( match[1] );
  }

  return values;
}
```

We use the magic string value `"extractValues.parse"` because it's quite unlikely that any other symbol in the registry would ever collide with that description.

If a user of this utility wants to override the parsing regular expression, they can also use the symbol registry:

```
extractValues[Symbol.for( "extractValues.parse" )] =  
  /..some pattern../g;  
  
extractValues( "..some string.." );
```

Aside from the assistance the symbol registry provides in globally storing these values, everything we're seeing here could have been done by just actually using the magic string `"extractValues.parse"` as the key, rather than the symbol. The improvements exist at the metaprogramming level more than the functional level.

You may have occasion to use a symbol value that has been stored in the registry to look up what description text (key) it's stored under. For example, you may need to signal to another part of your application how to locate a symbol in the registry because you cannot pass the symbol value itself.

You can retrieve a registered symbol's description text (key) using `Symbol.keyFor(..)`:

```
var s = Symbol.for( "something cool" );  
  
var desc = Symbol.keyFor( s );  
console.log( desc );           // "something cool"  
  
// get the symbol from the registry again  
var s2 = Symbol.for( desc );  
  
s2 === s;                   // true
```

Symbols as Object Properties

If a symbol is used as a property/key of an object, it's stored in a special way so that the property will not show up in a normal enumeration of the object's properties:

```
var o = {  
  foo: 42,  
  [ Symbol( "bar" ) ]: "hello world",  
  baz: true  
};  
  
Object.getOwnPropertyNames( o );    // [ "foo", "baz" ]
```

To retrieve an object's symbol properties:

```
Object.getOwnPropertySymbols( o ); // [ Symbol(bar) ]
```

This makes it clear that a property symbol is not actually hidden or inaccessible, as you can always see it in the `Object.getOwnPropertySymbols(...)` list.

Built-In Symbols ES6 comes with a number of predefined built-in symbols that expose various meta behaviors on JavaScript object values. However, these symbols are *not* registered in the global symbol registry, as one might expect.

Instead, they're stored as properties on the `Symbol` function object. For example, in the “`for..of`” section earlier in this chapter, we introduced the `Symbol.iterator` value:

```
var a = [1,2,3];  
  
a[Symbol.iterator];           // native function
```

The specification uses the `@@` prefix notation to refer to the built-in symbols, the most common ones being: `@@iterator`, `@@toStringTag`, `@@toPrimitive`. Several others are defined as well, though they probably won't be used as often.

Note: See “Well Known Symbols” in Chapter 7 for detailed information about how these built-in symbols are used for meta programming purposes.

Review

ES6 adds a heap of new syntax forms to JavaScript, so there's plenty to learn!

Most of these are designed to ease the pain points of common programming idioms, such as setting default values to function parameters and gathering the “rest” of the parameters into an array. Destructuring is a powerful tool for more concisely expressing assignments of values from arrays and nested objects.

While features like `=>` arrow functions appear to also be all about shorter and nicer-looking syntax, they actually have very specific behaviors that you should intentionally use only in appropriate situations.

Expanded Unicode support, new tricks for regular expressions, and even a new primitive `symbol` type round out the syntactic evolution of ES6.

You Don't Know JS: ES6 & Beyond

Chapter 3: Organization

It's one thing to write JS code, but it's another to properly organize it. Utilizing common patterns for organization and reuse goes a long way to improving the readability and understandability of your code. Remember: code is at least as much about communicating to other developers as it is about feeding the computer instructions.

ES6 has several important features that help significantly improve these patterns, including: iterators, generators, modules, and classes.

Iterators

An *iterator* is a structured pattern for pulling information from a source in one-at-a-time fashion. This pattern has been around programming for a long time. And to be sure, JS developers have been ad hoc designing and implementing iterators in JS programs since before anyone can remember, so it's not at all a new topic.

What ES6 has done is introduce an implicit standardized interface for iterators. Many of the built-in data structures in JavaScript will now expose an iterator implementing this standard. And you can also construct your own iterators adhering to the same standard, for maximal interoperability.

Iterators are a way of organizing ordered, sequential, pull-based consumption of data.

For example, you may implement a utility that produces a new unique identifier each time it's requested. Or you may produce an infinite series of values that rotate through a fixed list, in round-robin fashion. Or you could attach an iterator to a database query result to pull out new rows one at a time.

Although they have not commonly been used in JS in such a manner, iterators can also be thought of as controlling behavior one step at a time. This can be illustrated quite clearly when considering generators (see "Generators" later in this chapter), though you can certainly do the same without generators.

Interfaces

At the time of this writing, ES6 section 25.1.1.2 (<https://people.mozilla.org/~jorendorff/es6-draft.html#sec-iterator-interface>) details the `Iterator` interface as having the following requirement:

```
Iterator [required]
  next() {method}: retrieves next IteratorResult
```

There are two optional members that some iterators are extended with:

```
Iterator [optional]
  return() {method}: stops iterator and returns IteratorResult
  throw() {method}: signals error and returns IteratorResult
```

The `IteratorResult` interface is specified as:

```
IteratorResult
  value {property}: current iteration value or final return value
    (optional if `undefined`)
  done {property}: boolean, indicates completion status
```

Note: I call these interfaces implicit not because they're not explicitly called out in the specification – they are! – but because they're not exposed as direct objects accessible to code. JavaScript does not, in ES6, support any notion of “interfaces,” so adherence for your own code is purely conventional. However, wherever JS expects an iterator – a `for..of` loop, for instance – what you provide must adhere to these interfaces or the code will fail.

There's also an `Iterable` interface, which describes objects that must be able to produce iterators:

```
Iterable
  @@iterator() {method}: produces an Iterator
```

If you recall from “Built-In Symbols” in Chapter 2, `@@iterator` is the special built-in symbol representing the method that can produce iterator(s) for the object.

IteratorResult The `IteratorResult` interface specifies that the return value from any iterator operation will be an object of the form:

```
{ value: ... , done: true / false }
```

Built-in iterators will always return values of this form, but more properties are, of course, allowed to be present on the return value, as necessary.

For example, a custom iterator may add additional metadata to the result object (e.g., where the data came from, how long it took to retrieve, cache expiration length, frequency for the appropriate next request, etc.).

Note: Technically, `value` is optional if it would otherwise be considered absent or unset, such as in the case of the value `undefined`. Because accessing `res.value` will produce `undefined` whether it's present with that value or absent entirely, the presence/absence of the property is more an implementation detail or an optimization (or both), rather than a functional issue.

`next()` Iteration

Let's look at an array, which is an iterable, and the iterator it can produce to consume its values:

```

var arr = [1,2,3];

var it = arr[Symbol.iterator]();

it.next();      // { value: 1, done: false }
it.next();      // { value: 2, done: false }
it.next();      // { value: 3, done: false }

it.next();      // { value: undefined, done: true }

```

Each time the method located at `Symbol.iterator` (see Chapters 2 and 7) is invoked on this `arr` value, it will produce a new fresh iterator. Most structures will do the same, including all the built-in data structures in JS.

However, a structure like an event queue consumer might only ever produce a single iterator (singleton pattern). Or a structure might only allow one unique iterator at a time, requiring the current one to be completed before a new one can be created.

The `it` iterator in the previous snippet doesn't report `done: true` when you receive the 3 value. You have to call `next()` again, in essence going beyond the end of the array's values, to get the complete signal `done: true`. It may not be clear why until later in this section, but that design decision will typically be considered a best practice.

Primitive string values are also iterables by default:

```

var greeting = "hello world";

var it = greeting[Symbol.iterator]();

it.next();      // { value: "h", done: false }
it.next();      // { value: "e", done: false }
..

```

Note: Technically, the primitive value itself isn't iterable, but thanks to "boxing", "hello world" is coerced/converted to its `String` object wrapper form, which *is* an iterable. See the *Types & Grammar* title of this series for more information.

ES6 also includes several new data structures, called collections (see Chapter 5). These collections are not only iterables themselves, but they also provide API method(s) to generate an iterator, such as:

```

var m = new Map();
m.set( "foo", 42 );
m.set( { cool: true }, "hello world" );

var it1 = m[Symbol.iterator]();
var it2 = m.entries();

```

```
it1.next();      // { value: [ "foo", 42 ], done: false }
it2.next();      // { value: [ "foo", 42 ], done: false }
..
```

The `next(..)` method of an iterator can optionally take one or more arguments. The built-in iterators mostly do not exercise this capability, though a generator's iterator definitely does (see “Generators” later in this chapter).

By general convention, including all the built-in iterators, calling `next(..)` on an iterator that's already been exhausted is not an error, but will simply continue to return the result `{ value: undefined, done: true }`.

Optional: `return(..)` and `throw(..)`

The optional methods on the iterator interface – `return(..)` and `throw(..)` – are not implemented on most of the built-in iterators. However, they definitely do mean something in the context of generators, so see “Generators” for more specific information.

`return(..)` is defined as sending a signal to an iterator that the consuming code is complete and will not be pulling any more values from it. This signal can be used to notify the producer (the iterator responding to `next(..)` calls) to perform any cleanup it may need to do, such as releasing/closing network, database, or file handle resources.

If an iterator has a `return(..)` present and any condition occurs that can automatically be interpreted as abnormal or early termination of consuming the iterator, `return(..)` will automatically be called. You can call `return(..)` manually as well.

`return(..)` will return an `IteratorResult` object just like `next(..)` does. In general, the optional value you send to `return(..)` would be sent back as `value` in this `IteratorResult`, though there are nuanced cases where that might not be true.

`throw(..)` is used to signal an exception/error to an iterator, which possibly may be used differently by the iterator than the completion signal implied by `return(..)`. It does not necessarily imply a complete stop of the iterator as `return(..)` generally does.

For example, with generator iterators, `throw(..)` actually injects a thrown exception into the generator's paused execution context, which can be caught with a `try..catch`. An uncaught `throw(..)` exception would end up abnormally aborting the generator's iterator.

Note: By general convention, an iterator should not produce any more results after having called `return(..)` or `throw(..)`.

Iterator Loop

As we covered in the “`for..of`” section in Chapter 2, the ES6 `for..of` loop directly consumes a conforming iterable.

If an iterator is also an iterable, it can be used directly with the `for..of` loop. You make an iterator an iterable by giving it a `Symbol.iterator` method that simply returns the iterator itself:

```
var it = {
  // make the `it` iterator an iterable
  [Symbol.iterator]() { return this; },

  next() { .. },
  ..
};

it[Symbol.iterator]() === it; // true
```

Now we can consume the `it` iterator with a `for..of` loop:

```
for (var v of it) {
  console.log( v );
}
```

To fully understand how such a loop works, recall the `for` equivalent of a `for..of` loop from Chapter 2:

```
for (var v, res; (res = it.next()) && !res.done; ) {
  v = res.value;
  console.log( v );
}
```

If you look closely, you’ll see that `it.next()` is called before each iteration, and then `res.done` is consulted. If `res.done` is `true`, the expression evaluates to `false` and the iteration doesn’t occur.

Recall earlier that we suggested iterators should in general not return `done: true` along with the final intended value from the iterator. Now you can see why.

If an iterator returned `{ done: true, value: 42 }`, the `for..of` loop would completely discard the 42 value and it’d be lost. For this reason, assuming that your iterator may be consumed by patterns like the `for..of` loop or its manual `for` equivalent, you should probably wait to return `done: true` for signaling completion until after you’ve already returned all relevant iteration values.

Warning: You can, of course, intentionally design your iterator to return some relevant `value` at the same time as returning `done: true`. But don’t do this unless you’ve documented that as the case, and thus implicitly forced consumers

of your iterator to use a different pattern for iteration than is implied by `for..of` or its manual equivalent we depicted.

Custom Iterators

In addition to the standard built-in iterators, you can make your own! All it takes to make them interoperate with ES6's consumption facilities (e.g., the `for..of` loop and the `... operator`) is to adhere to the proper interface(s).

Let's try constructing an iterator that produces the infinite series of numbers in the Fibonacci sequence:

```
var Fib = {
  [Symbol.iterator]() {
    var n1 = 1, n2 = 1;

    return {
      // make the iterator an iterable
      [Symbol.iterator]() { return this; },

      next() {
        var current = n2;
        n2 = n1;
        n1 = n1 + current;
        return { value: current, done: false };
      },
      return(v) {
        console.log(
          "Fibonacci sequence abandoned."
        );
        return { value: v, done: true };
      }
    };
  };

  for (var v of Fib) {
    console.log( v );

    if (v > 50) break;
  }
  // 1 1 2 3 5 8 13 21 34 55
  // Fibonacci sequence abandoned.
```

Warning: If we hadn't inserted the `break` condition, this `for..of` loop would have run forever, which is probably not the desired result in terms of breaking

your program!

The `Fib[Symbol.iterator]()` method when called returns the iterator object with `next()` and `return(..)` methods on it. State is maintained via `n1` and `n2` variables, which are kept by the closure.

Let's *next* consider an iterator that is designed to run through a series (aka a queue) of actions, one item at a time:

```
var tasks = {
  [Symbol.iterator]() {
    var steps = this.actions.slice();

    return {
      // make the iterator an iterable
      [Symbol.iterator]() { return this; },

      next(...args) {
        if (steps.length > 0) {
          let res = steps.shift()( ...args );
          return { value: res, done: false };
        }
        else {
          return { done: true }
        }
      },
      return(v) {
        steps.length = 0;
        return { value: v, done: true };
      }
    };
  },
  actions: []
};
```

The iterator on `tasks` steps through functions found in the `actions` array property, if any, and executes them one at a time, passing in whatever arguments you pass to `next(..)`, and returning any return value to you in the standard `IteratorResult` object.

Here's how we could use this `tasks` queue:

```
tasks.actions.push(
  function step1(x){
    console.log( "step 1:", x );
    return x * 2;
  },
  function step2(x,y){
```

```

        console.log( "step 2:", x, y );
        return x + (y * 2);
    },
    function step3(x,y,z){
        console.log( "step 3:", x, y, z );
        return (x * y) + z;
    }
);

var it = tasks[Symbol.iterator]();

it.next( 10 );           // step 1: 10
// { value: 20, done: false }

it.next( 20, 50 );       // step 2: 20 50
// { value: 120, done: false }

it.next( 20, 50, 120 ); // step 3: 20 50 120
// { value: 1120, done: false }

it.next();               // { done: true }

```

This particular usage reinforces that iterators can be a pattern for organizing functionality, not just data. It's also reminiscent of what we'll see with generators in the next section.

You could even get creative and define an iterator that represents meta operations on a single piece of data. For example, we could define an iterator for numbers that by default ranges from 0 up to (or down to, for negative numbers) the number in question.

Consider:

```

if (!Number.prototype[Symbol.iterator]) {
    Object.defineProperty(
        Number.prototype,
        Symbol.iterator,
        {
            writable: true,
            configurable: true,
            enumerable: false,
            value: function iterator(){
                var i, inc, done = false, top = +this;

                // iterate positively or negatively?
                inc = 1 * (top < 0 ? -1 : 1);

                return {

```

Now, what tricks does this creativity afford us?

```
for (var i of [3]) {
    console.log( i );
}
// 0 1 2 3

[...-3]; // [0,-1,-2,-3]
```

Those are some fun tricks, though the practical utility is somewhat debatable. But then again, one might wonder why ES6 didn't just ship with such a minor but delightful feature easter egg!?

I'd be remiss if I didn't at least remind you that extending native prototypes as I'm doing in the previous snippet is something you should only do with caution

and awareness of potential hazards.

In this case, the chances that you'll have a collision with other code or even a future JS feature is probably exceedingly low. But just beware of the slight possibility. And document what you're doing verbosely for posterity's sake.

Note: I've expounded on this particular technique in this blog post (<http://blog.getify.com/iterating-es6-numbers/>) if you want more details. And this comment (<http://blog.getify.com/iterating-es6-numbers/comment-page-1/#comment-535294>) even suggests a similar trick but for making string character ranges.

Iterator Consumption

We've already shown consuming an iterator item by item with the `for..of` loop. But there are other ES6 structures that can consume iterators.

Let's consider the iterator attached to this array (though any iterator we choose would have the following behaviors):

```
var a = [1,2,3,4,5];
```

The `...` spread operator fully exhausts an iterator. Consider:

```
function foo(x,y,z,w,p) {
  console.log( x + y + z + w + p );
}
```

```
foo( ...a );           // 15
```

... can also spread an iterator inside an array:

```
var b = [ 0, ...a, 6 ];
b;                      // [0,1,2,3,4,5,6]
```

Array destructuring (see “Destructuring” in Chapter 2) can partially or completely (if paired with a `...` rest/gather operator) consume an iterator:

```
var it = a[Symbol.iterator]();
var [x,y] = it;           // take just the first two elements from `it`
var [z, ...w] = it;       // take the third, then the rest all at once

// is `it` fully exhausted? Yep.
it.next();                // { value: undefined, done: true }

x;                         // 1
y;                         // 2
z;                         // 3
w;                         // [4,5]
```

Generators

All functions run to completion, right? In other words, once a function starts running, it finishes before anything else can interrupt.

At least that's how it's been for the whole history of JavaScript up to this point. As of ES6, a new somewhat exotic form of function is being introduced, called a generator. A generator can pause itself in mid-execution, and can be resumed either right away or at a later time. So it clearly does not hold the run-to-completion guarantee that normal functions do.

Moreover, each pause/resume cycle in mid-execution is an opportunity for two-way message passing, where the generator can return a value, and the controlling code that resumes it can send a value back in.

As with iterators in the previous section, there are multiple ways to think about what a generator is, or rather what it's most useful for. There's no one right answer, but we'll try to consider several angles.

Note: See the *Async & Performance* title of this series for more information about generators, and also see Chapter 4 of this current title.

Syntax

The generator function is declared with this new syntax:

```
function *foo() {  
  // ...  
}
```

The position of the * is not functionally relevant. The same declaration could be written as any of the following:

```
function *foo() { ... }  
function* foo() { ... }  
function * foo() { ... }  
function*foo() { ... }  
...
```

The *only* difference here is stylistic preference. Most other literature seems to prefer `function* foo(..) { ... }`. I prefer `function *foo(..) { ... }`, so that's how I'll present them for the rest of this title.

My reason is purely didactic in nature. In this text, when referring to a generator function, I will use `*foo(..)`, as opposed to `foo(..)` for a normal function. I observe that `*foo(..)` more closely matches the * positioning of `function *foo(..) { ... }`.

Moreover, as we saw in Chapter 2 with concise methods, there's a concise generator form in object literals:

```
var a = {
  *foo() { .. }
};
```

I would say that with concise generators, `*foo() { .. }` is rather more natural than `* foo() { .. }`. So that further argues for matching the consistency with `*foo()`.

Consistency eases understanding and learning.

Executing a Generator Though a generator is declared with `*`, you still execute it like a normal function:

```
foo();
```

You can still pass it arguments, as in:

```
function *foo(x,y) {
  // ..
}

foo( 5, 10 );
```

The major difference is that executing a generator, like `foo(5,10)` doesn't actually run the code in the generator. Instead, it produces an iterator that will control the generator to execute its code.

We'll come back to this later in "Iterator Control," but briefly:

```
function *foo() {
  // ..
}

var it = foo();

// to start/advanced `*foo()`, call
// `it.next(..)`
```

yield Generators also have a new keyword you can use inside them, to signal the pause point: `yield`. Consider:

```
function *foo() {
  var x = 10;
  var y = 20;

  yield;

  var z = x + y;
}
```

In this `*foo()` generator, the operations on the first two lines would run at the beginning, then `yield` would pause the generator. If and when resumed, the last line of `*foo()` would run. `yield` can appear any number of times (or not at all, technically!) in a generator.

You can even put `yield` inside a loop, and it can represent a repeated pause point. In fact, a loop that never completes just means a generator that never completes, which is completely valid, and sometimes entirely what you need.

`yield` is not just a pause point. It's an expression that sends out a value when pausing the generator. Here's a `while..true` loop in a generator that for each iteration `yields` a new random number:

```
function *foo() {
  while (true) {
    yield Math.random();
  }
}
```

The `yield ..` expression not only sends a value – `yield` without a value is the same as `yield undefined` – but also receives (e.g., is replaced by) the eventual resumption value. Consider:

```
function *foo() {
  var x = yield 10;
  console.log( x );
}
```

This generator will first `yield` out the value 10 when pausing itself. When you resume the generator – using the `it.next(..)` we referred to earlier – whatever value (if any) you resume with will replace/complete the whole `yield 10` expression, meaning that value will be assigned to the `x` variable.

A `yield ..` expression can appear anywhere a normal expression can. For example:

```
function *foo() {
  var arr = [ yield 1, yield 2, yield 3 ];
  console.log( arr, yield 4 );
}
```

`*foo()` here has four `yield ..` expressions. Each `yield` results in the generator pausing to wait for a resumption value that's then used in the various expression contexts.

`yield` is not technically an operator, though when used like `yield 1` it sure looks like it. Because `yield` can be used all by itself as in `var x = yield;`, thinking of it as an operator can sometimes be confusing.

Technically, `yield ..` is of the same “expression precedence” – similar conceptually to operator precedence – as an assignment expression like `a = 3`. That

means `yield ..` can basically appear anywhere `a = 3` can validly appear.

Let's illustrate the symmetry:

```
var a, b;

a = 3;           // valid
b = 2 + a = 3;  // invalid
b = 2 + (a = 3); // valid

yield 3;          // valid
a = 2 + yield 3; // invalid
a = 2 + (yield 3); // valid
```

Note: If you think about it, it makes a sort of conceptual sense that a `yield ..` expression would behave similar to an assignment expression. When a paused `yield` expression is resumed, it's completed/replaced by the resumption value in a way that's not terribly dissimilar from being "assigned" that value.

The takeaway: if you need `yield ..` to appear in a position where an assignment like `a = 3` would not itself be allowed, it needs to be wrapped in a `()`.

Because of the low precedence of the `yield` keyword, almost any expression after a `yield ..` will be computed first before being sent with `yield`. Only the `...` spread operator and the `,` comma operator have lower precedence, meaning they'd bind after the `yield` has been evaluated.

So just like with multiple operators in normal statements, another case where `()` might be needed is to override (elevate) the low precedence of `yield`, such as the difference between these expressions:

```
yield 2 + 3;           // same as `yield (2 + 3)`

(yield 2) + 3;         // `yield 2` first, then `+ 3`
```

Just like `=` assignment, `yield` is also "right-associative," which means that multiple `yield` expressions in succession are treated as having been `(...)` grouped from right to left. So, `yield yield yield 3` is treated as `yield (yield (yield 3))`. A "left-associative" interpretation like `((yield) yield) yield 3` would make no sense.

Just like with operators, it's a good idea to use `(...)` grouping, even if not strictly required, to disambiguate your intent if `yield` is combined with other operators or `yields`.

Note: See the *Types & Grammar* title of this series for more information about operator precedence and associativity.

`yield *` In the same way that the `*` makes a `function` declaration into `function *` generator declaration, a `*` makes `yield` into `yield *`, which is a very different mechanism, called *yield delegation*. Grammatically, `yield *..` will behave the same as a `yield ..`, as discussed in the previous section.

`yield * ..` requires an iterable; it then invokes that iterable's iterator, and delegates its own host generator's control to that iterator until it's exhausted. Consider:

```
function *foo() {
  yield *[1,2,3];
}
```

Note: As with the `*` position in a generator's declaration (discussed earlier), the `*` positioning in `yield *` expressions is stylistically up to you. Most other literature prefers `yield* ..`, but I prefer `yield *..`, for very symmetrical reasons as already discussed.

The `[1,2,3]` value produces an iterator that will step through its values, so the `*foo()` generator will yield those values out as it's consumed. Another way to illustrate the behavior is in `yield` delegating to another generator:

```
function *foo() {
  yield 1;
  yield 2;
  yield 3;
}

function *bar() {
  yield *foo();
}
```

The iterator produced when `*bar()` calls `*foo()` is delegated to via `yield *`, meaning whatever value(s) `*foo()` produces will be produced by `*bar()`.

Whereas with `yield ..` the completion value of the expression comes from resuming the generator with `it.next(..)`, the completion value of the `yield *..` expression comes from the return value (if any) from the delegated-to iterator.

Built-in iterators generally don't have return values, as we covered at the end of the "Iterator Loop" section earlier in this chapter. But if you define your own custom iterator (or generator), you can design it to `return` a value, which `yield *..` would capture:

```
function *foo() {
  yield 1;
  yield 2;
  yield 3;
  return 4;
```

```

}

function *bar() {
  var x = yield *foo();
  console.log( "x:", x );
}

for (var v of bar()) {
  console.log( v );
}
// 1 2 3
// x: 4

```

While the 1, 2, and 3 values are yielded out of `*foo()` and then out of `*bar()`, the 4 value returned from `*foo()` is the completion value of the `yield *foo()` expression, which then gets assigned to `x`.

Because `yield *` can call another generator (by way of delegating to its iterator), it can also perform a sort of generator recursion by calling itself:

```

function *foo(x) {
  if (x < 3) {
    x = yield *foo( x + 1 );
  }
  return x * 2;
}

foo( 1 );

```

The result from `foo(1)` and then calling the iterator's `next()` to run it through its recursive steps will be 24. The first `*foo(...)` run has `x` at value 1, which is `x < 3`. `x + 1` is passed recursively to `*foo(...)`, so `x` is then 2. One more recursive call results in `x` of 3.

Now, because `x < 3` fails, the recursion stops, and `return 3 * 2` gives 6 back to the previous call's `yield ...` expression, which is then assigned to `x`. Another `return 6 * 2` returns 12 back to the previous call's `x`. Finally `12 * 2`, or 24, is returned from the completed run of the `*foo(..)` generator.

Iterator Control

Earlier, we briefly introduced the concept that generators are controlled by iterators. Let's fully dig into that now.

Recall the recursive `*foo(..)` from the previous section. Here's how we'd run it:

```

function *foo(x) {
  if (x < 3) {

```

```

        x = yield *foo( x + 1 );
    }
    return x * 2;
}

var it = foo( 1 );
it.next();           // { value: 24, done: true }

```

In this case, the generator doesn't really ever pause, as there's no `yield` .. expression. Instead, `yield *` just keeps the current iteration step going via the recursive call. So, just one call to the iterator's `next()` function fully runs the generator.

Now let's consider a generator that will have multiple steps and thus multiple produced values:

```

function *foo() {
    yield 1;
    yield 2;
    yield 3;
}

```

We already know we can consume an iterator, even one attached to a generator like `*foo()`, with a `for..of` loop:

```

for (var v of foo()) {
    console.log( v );
}
// 1 2 3

```

Note: The `for..of` loop requires an iterable. A generator function reference (like `foo`) by itself is not an iterable; you must execute it with `foo()` to get the iterator (which is also an iterable, as we explained earlier in this chapter). You could theoretically extend the `GeneratorPrototype` (the prototype of all generator functions) with a `Symbol.iterator` function that essentially just does `return this()`. That would make the `foo` reference itself an iterable, which means `for (var v of foo) { .. }` (notice no `()` on `foo`) will work.

Let's instead iterate the generator manually:

```

function *foo() {
    yield 1;
    yield 2;
    yield 3;
}

var it = foo();

it.next();           // { value: 1, done: false }
it.next();           // { value: 2, done: false }

```

```
it.next();           // { value: 3, done: false }

it.next();           // { value: undefined, done: true }
```

If you look closely, there are three `yield` statements and four `next()` calls. That may seem like a strange mismatch. In fact, there will always be one more `next()` call than `yield` expression, assuming all are evaluated and the generator is fully run to completion.

But if you look at it from the opposite perspective (inside-out instead of outside-in), the matching between `yield` and `next()` makes more sense.

Recall that the `yield ..` expression will be completed by the value you resume the generator with. That means the argument you pass to `next(..)` completes whatever `yield ..` expression is currently paused waiting for a completion.

Let's illustrate this perspective this way:

```
function *foo() {
  var x = yield 1;
  var y = yield 2;
  var z = yield 3;
  console.log( x, y, z );
}
```

In this snippet, each `yield ..` is sending a value out (1, 2, 3), but more directly, it's pausing the generator to wait for a value. In other words, it's almost like asking the question, "What value should I use here? I'll wait to hear back."

Now, here's how we control `*foo()` to start it up:

```
var it = foo();

it.next();           // { value: 1, done: false }
```

That first `next()` call is starting up the generator from its initial paused state, and running it to the first `yield`. At the moment you call that first `next()`, there's no `yield ..` expression waiting for a completion. If you passed a value to that first `next()` call, it would currently just be thrown away, because no `yield` is waiting to receive such a value.

Note: An early proposal for the "beyond ES6" timeframe *would* let you access a value passed to an initial `next(..)` call via a separate meta property (see Chapter 7) inside the generator.

Now, let's answer the currently pending question, "What value should I assign to `x`?" We'll answer it by sending a value to the `next` `next(..)` call:

```
it.next( "foo" );           // { value: 2, done: false }
```

Now, the `x` will have the value "foo", but we've also asked a new question, "What value should I assign to `y`?" And we answer:

```
it.next( "bar" );           // { value: 3, done: false }
```

Answer given, another question asked. Final answer:

```
it.next( "baz" );           // "foo" "bar" "baz"
                           // { value: undefined, done: true }
```

Now it should be clearer how each `yield` .. “question” is answered by the `next` `next(..)` call, and so the “extra” `next()` call we observed is always just the initial one that starts everything going.

Let’s put all those steps together:

```
var it = foo();

// start up the generator
it.next();                 // { value: 1, done: false }

// answer first question
it.next( "foo" );          // { value: 2, done: false }

// answer second question
it.next( "bar" );          // { value: 3, done: false }

// answer third question
it.next( "baz" );          // "foo" "bar" "baz"
                           // { value: undefined, done: true }
```

You can think of a generator as a producer of values, in which case each iteration is simply producing a value to be consumed.

But in a more general sense, perhaps it’s appropriate to think of generators as controlled, progressive code execution, much like the `tasks` queue example from the earlier “Custom Iterators” section.

Note: That perspective is exactly the motivation for how we’ll revisit generators in Chapter 4. Specifically, there’s no reason that `next(..)` has to be called right away after the previous `next(..)` finishes. While the generator’s inner execution context is paused, the rest of the program continues unblocked, including the ability for asynchronous actions to control when the generator is resumed.

Early Completion

As we covered earlier in this chapter, the iterator attached to a generator supports the optional `return(..)` and `throw(..)` methods. Both of them have the effect of aborting a paused generator immediately.

Consider:

```

function *foo() {
  yield 1;
  yield 2;
  yield 3;
}

var it = foo();

it.next();           // { value: 1, done: false }

it.return( 42 );    // { value: 42, done: true }

it.next();           // { value: undefined, done: true }

return(x) is kind of like forcing a return x to be processed at exactly that
moment, such that you get the specified value right back. Once a generator is
completed, either normally or early as shown, it no longer processes any code
or returns any values.

```

In addition to `return(..)` being callable manually, it's also called automatically at the end of iteration by any of the ES6 constructs that consume iterators, such as the `for..of` loop and the `...` spread operator.

The purpose for this capability is so the generator can be notified if the controlling code is no longer going to iterate over it anymore, so that it can perhaps do any cleanup tasks (freeing up resources, resetting status, etc.). Identical to a normal function cleanup pattern, the main way to accomplish this is to use a `finally` clause:

```

function *foo() {
  try {
    yield 1;
    yield 2;
    yield 3;
  }
  finally {
    console.log( "cleanup!" );
  }
}

for (var v of foo()) {
  console.log( v );
}

// 1 2 3
// cleanup!

var it = foo();

```

```

it.next();           // { value: 1, done: false }
it.return( 42 );    // cleanup!
// { value: 42, done: true }

```

Warning: Do not put a `yield` statement inside the `finally` clause! It's valid and legal, but it's a really terrible idea. It acts in a sense as deferring the completion of the `return(..)` call you made, as any `yield ..` expressions in the `finally` clause are respected to pause and send messages; you don't immediately get a completed generator as expected. There's basically no good reason to opt in to that crazy *bad part*, so avoid doing so!

In addition to the previous snippet showing how `return(..)` aborts the generator while still triggering the `finally` clause, it also demonstrates that a generator produces a whole new iterator each time it's called. In fact, you can use multiple iterators attached to the same generator concurrently:

```

function *foo() {
  yield 1;
  yield 2;
  yield 3;
}

var it1 = foo();
it1.next();           // { value: 1, done: false }
it1.next();           // { value: 2, done: false }

var it2 = foo();
it2.next();           // { value: 1, done: false }

it1.next();           // { value: 3, done: false }

it2.next();           // { value: 2, done: false }
it2.next();           // { value: 3, done: false }

it2.next();           // { value: undefined, done: true }
it1.next();           // { value: undefined, done: true }

```

Early Abort Instead of calling `return(..)`, you can call `throw(..)`. Just like `return(x)` is essentially injecting a `return x` into the generator at its current pause point, calling `throw(x)` is essentially like injecting a `throw x` at the pause point.

Other than the exception behavior (we cover what that means to `try` clauses in the next section), `throw(..)` produces the same sort of early completion that aborts the generator's run at its current pause point. For example:

```

function *foo() {
  yield 1;

```

```

        yield 2;
        yield 3;
    }

var it = foo();

it.next();           // { value: 1, done: false }

try {
    it.throw( "Oops!" );
}
catch (err) {
    console.log( err ); // Exception: Oops!
}

it.next();           // { value: undefined, done: true }

```

Because `throw(..)` basically injects a `throw ..` in replacement of the `yield 1` line of the generator, and nothing handles this exception, it immediately propagates back out to the calling code, which handles it with a `try..catch`.

Unlike `return(..)`, the iterator's `throw(..)` method is never called automatically.

Of course, though not shown in the previous snippet, if a `try..finally` clause was waiting inside the generator when you call `throw(..)`, the `finally` clause would be given a chance to complete before the exception is propagated back to the calling code.

Error Handling

As we've already hinted, error handling with generators can be expressed with `try..catch`, which works in both inbound and outbound directions:

```

function *foo() {
    try {
        yield 1;
    }
    catch (err) {
        console.log( err );
    }

    yield 2;

    throw "Hello!";
}

var it = foo();

```

```

it.next();           // { value: 1, done: false }

try {
  it.throw( "Hi!" ); // Hi!
  // { value: 2, done: false }
  it.next();

  console.log( "never gets here" );
}
catch (err) {
  console.log( err ); // Hello!
}

```

Errors can also propagate in both directions through `yield *` delegation:

```

function *foo() {
  try {
    yield 1;
  }
  catch (err) {
    console.log( err );
  }

  yield 2;

  throw "foo: e2";
}

function *bar() {
  try {
    yield *foo();

    console.log( "never gets here" );
  }
  catch (err) {
    console.log( err );
  }
}

var it = bar();

try {
  it.next();           // { value: 1, done: false }

  it.throw( "e1" ); // e1
  // { value: 2, done: false }
}

```

```

    it.next();           // foo: e2
    // { value: undefined, done: true }
}
catch (err) {
    console.log( "never gets here" );
}

it.next();           // { value: undefined, done: true }

```

When `*foo()` calls `yield 1`, the `1` value passes through `*bar()` untouched, as we've already seen.

But what's most interesting about this snippet is that when `*foo()` calls `throw "foo: e2"`, this error propagates to `*bar()` and is immediately caught by `*bar()`'s `try..catch` block. The error doesn't pass through `*bar()` like the `1` value did.

`*bar()`'s `catch` then does a normal output of `err ("foo: e2")` and then `*bar()` finishes normally, which is why the `{ value: undefined, done: true }` iterator result comes back from `it.next()`.

If `*bar()` didn't have a `try..catch` around the `yield ...` expression, the error would of course propagate all the way out, and on the way through it still would complete (abort) `*bar()`.

Transpiling a Generator

Is it possible to represent a generator's capabilities prior to ES6? It turns out it is, and there are several great tools that do so, including most notably Facebook's Regenerator tool (<https://facebook.github.io/regenerator/>).

But just to better understand generators, let's try our hand at manually converting. Basically, we're going to create a simple closure-based state machine.

We'll keep our source generator really simple:

```

function *foo() {
    var x = yield 42;
    console.log( x );
}

```

To start, we'll need a function called `foo()` that we can execute, which needs to return an iterator:

```

function foo() {
    // ..

    return {
        next: function(v) {

```

```

        // ...
    }

    // we'll skip `return(..)` and `throw(..)`
};

}

```

Now, we need some inner variable to keep track of where we are in the steps of our “generator”’s logic. We’ll call it `state`. There will be three states: 0 initially, 1 while waiting to fulfill the `yield` expression, and 2 once the generator is complete.

Each time `next(..)` is called, we need to process the next step, and then increment `state`. For convenience, we’ll put each step into a `case` clause of a `switch` statement, and we’ll hold that in an inner function called `nextState(..)` that `next(..)` can call. Also, because `x` is a variable across the overall scope of the “generator,” it needs to live outside the `nextState(..)` function.

Here it is all together (obviously somewhat simplified, to keep the conceptual illustration clearer):

```

function foo() {
    function nextState(v) {
        switch (state) {
            case 0:
                state++;

                // the `yield` expression
                return 42;
            case 1:
                state++;

                // `yield` expression fulfilled
                x = v;
                console.log( x );

                // the implicit `return`
                return undefined;

                // no need to handle state `2`
        }
    }

    var state = 0, x;

    return {
        next: function(v) {
            var ret = nextState( v );

```

```

        return { value: ret, done: (state == 2) };
    }

    // we'll skip `return(..)` and `throw(..)`
};

}

```

And finally, let's test our pre-ES6 "generator":

```

var it = foo();

it.next();           // { value: 42, done: false }

it.next( 10 );      // 10
                    // { value: undefined, done: true }

```

Not bad, huh? Hopefully this exercise solidifies in your mind that generators are actually just simple syntax for state machine logic. That makes them widely applicable.

Generator Uses

So, now that we much more deeply understand how generators work, what are they useful for?

We've seen two major patterns:

- *Producing a series of values:* This usage can be simple (e.g., random strings or incremented numbers), or it can represent more structured data access (e.g., iterating over rows returned from a database query).

Either way, we use the iterator to control a generator so that some logic can be invoked for each call to `next(..)`. Normal iterators on data structures merely pull values without any controlling logic.

- *Queue of tasks to perform serially:* This usage often represents flow control for the steps in an algorithm, where each step requires retrieval of data from some external source. The fulfillment of each piece of data may be immediate, or may be asynchronously delayed.

From the perspective of the code inside the generator, the details of sync or async at a `yield` point are entirely opaque. Moreover, these details are intentionally abstracted away, such as not to obscure the natural sequential expression of steps with such implementation complications. Abstraction also means the implementations can be swapped/refactored often without touching the code in the generator at all.

When generators are viewed in light of these uses, they become a lot more than just a different or nicer syntax for a manual state machine. They are a

powerful abstraction tool for organizing and controlling orderly production and consumption of data.

Modules

I don't think it's an exaggeration to suggest that the single most important code organization pattern in all of JavaScript is, and always has been, the module. For myself, and I think for a large cross-section of the community, the module pattern drives the vast majority of code.

The Old Way

The traditional module pattern is based on an outer function with inner variables and functions, and a returned “public API” with methods that have closure over the inner data and capabilities. It's often expressed like this:

```
function Hello(name) {
  function greeting() {
    console.log( "Hello " + name + "!" );
  }

  // public API
  return {
    greeting: greeting
  };
}

var me = Hello( "Kyle" );
me.greeting();           // Hello Kyle!
```

This `Hello(..)` module can produce multiple instances by being called subsequent times. Sometimes, a module is only called for as a singleton (i.e., it just needs one instance), in which case a slight variation on the previous snippet, using an IIFE, is common:

```
var me = (function Hello(name){
  function greeting() {
    console.log( "Hello " + name + "!" );
  }

  // public API
  return {
    greeting: greeting
  };
})( "Kyle" );

me.greeting();           // Hello Kyle!
```

This pattern is tried and tested. It's also flexible enough to have a wide assortment of variations for a number of different scenarios.

One of the most common is the Asynchronous Module Definition (AMD), and another is the Universal Module Definition (UMD). We won't cover the particulars of these patterns and techniques here, but they're explained extensively in many places online.

Moving Forward

As of ES6, we no longer need to rely on the enclosing function and closure to provide us with module support. ES6 modules have first class syntactic and functional support.

Before we get into the specific syntax, it's important to understand some fairly significant conceptual differences with ES6 modules compared to how you may have dealt with modules in the past:

- ES6 uses file-based modules, meaning one module per file. At this time, there is no standardized way of combining multiple modules into a single file.

That means that if you are going to load ES6 modules directly into a browser web application, you will be loading them individually, not as a large bundle in a single file as has been common in performance optimization efforts.

It's expected that the contemporaneous advent of HTTP/2 will significantly mitigate any such performance concerns, as it operates on a persistent socket connection and thus can very efficiently load many smaller files in parallel and interleaved with one another.

- The API of an ES6 module is static. That is, you define statically what all the top-level exports are on your module's public API, and those cannot be amended later.

Some uses are accustomed to being able to provide dynamic API definitions, where methods can be added/removed/replaced in response to runtime conditions. Either these uses will have to change to fit with ES6 static APIs, or they will have to restrain the dynamic changes to properties/methods of a second-level object.

- ES6 modules are singletons. That is, there's only one instance of the module, which maintains its state. Every time you import that module into another module, you get a reference to the one centralized instance. If you want to be able to produce multiple module instances, your module will need to provide some sort of factory to do it.
- The properties and methods you expose on a module's public API are not just normal assignments of values or references. They are actual bindings

(almost like pointers) to the identifiers in your inner module definition.

In pre-ES6 modules, if you put a property on your public API that holds a primitive value like a number or string, that property assignment was by value-copy, and any internal update of a corresponding variable would be separate and not affect the public copy on the API object.

With ES6, exporting a local private variable, even if it currently holds a primitive string/number/etc, exports a binding to the variable. If the module changes the variable's value, the external import binding now resolves to that new value.

- Importing a module is the same thing as statically requesting it to load (if it hasn't already). If you're in a browser, that implies a blocking load over the network. If you're on a server (i.e., Node.js), it's a blocking load from the filesystem.

However, don't panic about the performance implications. Because ES6 modules have static definitions, the import requirements can be statically scanned, and loads will happen preemptively, even before you've used the module.

ES6 doesn't actually specify or handle the mechanics of how these load requests work. There's a separate notion of a Module Loader, where each hosting environment (browser, Node.js, etc.) provides a default Loader appropriate to the environment. The importing of a module uses a string value to represent where to get the module (URL, file path, etc.), but this value is opaque in your program and only meaningful to the Loader itself.

You can define your own custom Loader if you want more fine-grained control than the default Loader affords – which is basically none, as it's totally hidden from your program's code.

As you can see, ES6 modules will serve the overall use case of organizing code with encapsulation, controlling public APIs, and referencing dependency imports. But they have a very particular way of doing so, and that may or may not fit very closely with how you've already been doing modules for years.

CommonJS There's a similar, but not fully compatible, module syntax called CommonJS, which is familiar to those in the Node.js ecosystem.

For lack of a more tactful way to say this, in the long run, ES6 modules essentially are bound to supersede all previous formats and standards for modules, even CommonJS, as they are built on syntactic support in the language. This will, in time, inevitably win out as the superior approach, if for no other reason than ubiquity.

We face a fairly long road to get to that point, though. There are literally hundreds of thousands of CommonJS style modules in the server-side JavaScript world, and 10 times that many modules of varying format standards (UMD,

AMD, ad hoc) in the browser world. It will take many years for the transitions to make any significant progress.

In the interim, module transpilers/converters will be an absolute necessity. You might as well just get used to that new reality. Whether you author in regular modules, AMD, UMD, CommonJS, or ES6, these tools will have to parse and convert to a format that is suitable for whatever environment your code will run in.

For Node.js, that probably means (for now) that the target is CommonJS. For the browser, it's probably UMD or AMD. Expect lots of flux on this over the next few years as these tools mature and best practices emerge.

From here on out, my best advice on modules is this: whatever format you've been religiously attached to with strong affinity, also develop an appreciation for and understanding of ES6 modules, such as they are, and let your other module tendencies fade. They *are* the future of modules in JS, even if that reality is a bit of a ways off.

The New Way

The two main new keywords that enable ES6 modules are `import` and `export`. There's lots of nuance to the syntax, so let's take a deeper look.

Warning: An important detail that's easy to overlook: both `import` and `export` must always appear in the top-level scope of their respective usage. For example, you cannot put either an `import` or `export` inside an `if` conditional; they must appear outside of all blocks and functions.

exporting API Members The `export` keyword is either put in front of a declaration, or used as an operator (of sorts) with a special list of bindings to export. Consider:

```
export function foo() {  
  // ..  
}  
  
export var awesome = 42;  
  
var bar = [1,2,3];  
export { bar };
```

Another way of expressing the same exports:

```
function foo() {  
  // ..  
}  
  
var awesome = 42;
```

```
var bar = [1,2,3];

export { foo, awesome, bar };
```

These are all called *named exports*, as you are in effect exporting the name bindings of the variables/functions/etc.

Anything you don't *label* with `export` stays private inside the scope of the module. That is, although something like `var bar = ..` looks like it's declaring at the top-level global scope, the top-level scope is actually the module itself; there is no global scope in modules.

Note: Modules *do* still have access to `window` and all the "globals" that hang off it, just not as lexical top-level scope. However, you really should stay away from the globals in your modules if at all possible.

You can also "rename" (aka alias) a module member during named export:

```
function foo() { .. }

export { foo as bar };
```

When this module is imported, only the `bar` member name is available to import; `foo` stays hidden inside the module.

Module exports are not just normal assignments of values or references, as you're accustomed to with the `=` assignment operator. Actually, when you export something, you're exporting a binding (kinda like a pointer) to that thing (variable, etc.).

Within your module, if you change the value of a variable you already exported a binding to, even if it's already been imported (see the next section), the imported binding will resolve to the current (updated) value.

Consider:

```
var awesome = 42;
export { awesome };

// later
awesome = 100;
```

When this module is imported, regardless of whether that's before or after the `awesome = 100` setting, once that assignment has happened, the imported binding resolves to the 100 value, not 42.

That's because the binding is, in essence, a reference to, or a pointer to, the `awesome` variable itself, rather than a copy of its value. This is a mostly unprecedented concept for JS introduced with ES6 module bindings.

Though you can clearly use `export` multiple times inside a module's definition, ES6 definitely prefers the approach that a module has a single export, which is

known as a *default export*. In the words of some members of the TC39 committee, you’re “rewarded with simpler `import` syntax” if you follow that pattern, and conversely “penalized” with more verbose syntax if you don’t.

A default export sets a particular exported binding to be the default when importing the module. The name of the binding is literally `default`. As you’ll see later, when importing module bindings you can also rename them, as you commonly will with a default export.

There can only be one `default` per module definition. We’ll cover `import` in the next section, and you’ll see how the `import` syntax is more concise if the module has a default export.

There’s a subtle nuance to default export syntax that you should pay close attention to. Compare these two snippets:

```
function foo(..) {  
  // ..  
}  
  
export default foo;
```

And this one:

```
function foo(..) {  
  // ..  
}  
  
export { foo as default };
```

In the first snippet, you are exporting a binding to the function expression value at that moment, *not* to the identifier `foo`. In other words, `export default ..` takes an expression. If you later assign `foo` to a different value inside your module, the module import still reveals the function originally exported, not the new value.

By the way, the first snippet could also have been written as:

```
export default function foo(..) {  
  // ..  
}
```

Warning: Although the `function foo..` part here is technically a function expression, for the purposes of the internal scope of the module, it’s treated like a function declaration, in that the `foo` name is bound in the module’s top-level scope (often called “hoisting”). The same is true for `export default class Foo...`. However, while you *can* do `export var foo = ...`, you currently cannot do `export default var foo = ...` (or `let` or `const`), in a frustrating case of inconsistency. At the time of this writing, there’s already discussion of adding that capability in soon, post-ES6, for consistency sake.

Recall the second snippet again:

```
function foo(..) {  
  // ..  
}  
  
export { foo as default };
```

In this version of the module export, the default export binding is actually to the `foo` identifier rather than its value, so you get the previously described binding behavior (i.e., if you later change `foo`'s value, the value seen on the import side will also be updated).

Be very careful of this subtle gotcha in default export syntax, especially if your logic calls for export values to be updated. If you never plan to update a default export's value, `export default ..` is fine. If you do plan to update the value, you must use `export { .. as default }`. Either way, make sure to comment your code to explain your intent!

Because there can only be one `default` per module, you may be tempted to design your module with one default export of an object with all your API methods on it, such as:

```
export default {  
  foo() { .. },  
  bar() { .. },  
  ..  
};
```

That pattern seems to map closely to how a lot of developers have already structured their pre-ES6 modules, so it seems like a natural approach. Unfortunately, it has some downsides and is officially discouraged.

In particular, the JS engine cannot statically analyze the contents of a plain object, which means it cannot do some optimizations for static `import` performance. The advantage of having each member individually and explicitly exported is that the engine *can* do the static analysis and optimization.

If your API has more than one member already, it seems like these principles – one default export per module, and all API members as named exports – are in conflict, doesn't it? But you *can* have a single default export as well as other named exports; they are not mutually exclusive.

So, instead of this (discouraged) pattern:

```
export default function foo() { .. }  
  
foo.bar = function() { .. };  
foo.baz = function() { .. };
```

You can do:

```
export default function foo() { ... }

export function bar() { ... }
export function baz() { ... }
```

Note: In this previous snippet, I used the name `foo` for the function that `default` labels. That `foo` name, however, is ignored for the purposes of export – `default` is actually the exported name. When you import this default binding, you can give it whatever name you want, as you'll see in the next section.

Alternatively, some will prefer:

```
function foo() { ... }
function bar() { ... }
function baz() { ... }

export { foo as default, bar, baz, ... };
```

The effects of mixing default and named exports will be more clear when we cover `import` shortly. But essentially it means that the most concise default import form would only retrieve the `foo()` function. The user could additionally manually list `bar` and `baz` as named imports, if they want them.

You can probably imagine how tedious that's going to be for consumers of your module if you have lots of named export bindings. There is a wildcard import form where you import all of a module's exports within a single namespace object, but there's no way to wildcard import to top-level bindings.

Again, the ES6 module mechanism is intentionally designed to discourage modules with lots of exports; relatively speaking, it's desired that such approaches be a little more difficult, as a sort of social engineering to encourage simple module design in favor of large/complex module design.

I would probably recommend you not mix default export with named exports, especially if you have a large API and refactoring to separate modules isn't practical or desired. In that case, just use all named exports, and document that consumers of your module should probably use the `import * as ..` (namespace import, discussed in the next section) approach to bring the whole API in at once on a single namespace.

We mentioned this earlier, but let's come back to it in more detail. Other than the `export default ...` form that exports an expression value binding, all other export forms are exporting bindings to local identifiers. For those bindings, if you change the value of a variable inside a module after exporting, the external imported binding will access the updated value:

```
var foo = 42;
export { foo as default };

export var bar = "hello world";
```

```
foo = 10;
bar = "cool";
```

When you import this module, the `default` and `bar` exports will be bound to the local variables `foo` and `bar`, meaning they will reveal the updated 10 and "cool" values. The values at time of export are irrelevant. The values at time of import are irrelevant. The bindings are live links, so all that matters is what the current value is when you access the binding.

Warning: Two-way bindings are not allowed. If you import a `foo` from a module, and try to change the value of your imported `foo` variable, an error will be thrown! We'll revisit that in the next section.

You can also re-export another module's exports, such as:

```
export { foo, bar } from "baz";
export { foo as FOO, bar as BAR } from "baz";
export * from "baz";
```

Those forms are similar to just first importing from the "baz" module then listing its members explicitly for export from your module. However, in these forms, the members of the "baz" module are never imported to your module's local scope; they sort of pass through untouched.

importing API Members To import a module, unsurprisingly you use the `import` statement. Just as `export` has several nuanced variations, so does `import`, so spend plenty of time considering the following issues and experimenting with your options.

If you want to import certain specific named members of a module's API into your top-level scope, you use this syntax:

```
import { foo, bar, baz } from "foo";
```

Warning: The `{ ... }` syntax here may look like an object literal, or even an object destructuring syntax. However, its form is special just for modules, so be careful not to confuse it with other `{ ... }` patterns elsewhere.

The "foo" string is called a *module specifier*. Because the whole goal is statically analyzable syntax, the module specifier must be a string literal; it cannot be a variable holding the string value.

From the perspective of your ES6 code and the JS engine itself, the contents of this string literal are completely opaque and meaningless. The module loader will interpret this string as an instruction of where to find the desired module, either as a URL path or a local filesystem path.

The `foo`, `bar`, and `baz` identifiers listed must match named exports on the module's API (static analysis and error assertion apply). They are bound as top-level identifiers in your current scope:

```
import { foo } from "foo";
```

```
foo();
```

You can rename the bound identifiers imported, as:

```
import { foo as theFooFunc } from "foo";
```

```
theFooFunc();
```

If the module has just a default export that you want to import and bind to an identifier, you can opt to skip the `{ .. }` surrounding syntax for that binding. The `import` in this preferred case gets the nicest and most concise of the `import` syntax forms:

```
import foo from "foo";
```

// or:

```
import { default as foo } from "foo";
```

Note: As explained in the previous section, the `default` keyword in a module's `export` specifies a named export where the name is actually `default`, as is illustrated by the second more verbose syntax option. The renaming from `default` to, in this case, `foo`, is explicit in the latter syntax and is identical yet implicit in the former syntax.

You can also import a default export along with other named exports, if the module has such a definition. Recall this module definition from earlier:

```
export default function foo() { .. }
```

```
export function bar() { .. }
```

```
export function baz() { .. }
```

To import that module's default export and its two named exports:

```
import FOOFN, { bar, baz as BAZ } from "foo";
```

```
FOOFN();
```

```
bar();
```

```
BAZ();
```

The strongly suggested approach from ES6's module philosophy is that you only import the specific bindings from a module that you need. If a module provides 10 API methods, but you only need two of them, some believe it wasteful to bring in the entire set of API bindings.

One benefit, besides code being more explicit, is that narrow imports make static analysis and error detection (accidentally using the wrong binding name, for instance) more robust.

Of course, that's just the standard position influenced by ES6 design philosophy; there's nothing that requires adherence to that approach.

Many developers would be quick to point out that such approaches can be more tedious, requiring you to regularly revisit and update your `import` statement(s) each time you realize you need something else from a module. The trade-off is in exchange for convenience.

In that light, the preference might be to import everything from the module into a single namespace, rather than importing individual members, each directly into the scope. Fortunately, the `import` statement has a syntax variation that can support this style of module consumption, called *namespace import*.

Consider a "foo" module exported as:

```
export function bar() { ... }
export var x = 42;
export function baz() { ... }
```

You can import that entire API to a single module namespace binding:

```
import * as foo from "foo";

foo.bar();
foo.x;           // 42
foo.baz();
```

Note: The `* as ..` clause requires the `*` wildcard. In other words, you cannot do something like `import { bar, x } as foo from "foo"` to bring in only part of the API but still bind to the `foo` namespace. I would have liked something like that, but for ES6 it's all or nothing with the namespace import.

If the module you're importing with `* as ..` has a default export, it is named `default` in the namespace specified. You can additionally name the default import outside of the namespace binding, as a top-level identifier. Consider a "world" module exported as:

```
export default function foo() { ... }
export function bar() { ... }
export function baz() { ... }
```

And this `import`:

```
import foofn, * as hello from "world";

foofn();
hello.default();
hello.bar();
hello.baz();
```

While this syntax is valid, it can be rather confusing that one method of the module (the default export) is bound at the top-level of your scope, whereas the

rest of the named exports (and one called `default`) are bound as properties on a differently named (`hello`) identifier namespace.

As I mentioned earlier, my suggestion would be to avoid designing your module exports in this way, to reduce the chances that your module's users will suffer these strange quirks.

All imported bindings are immutable and/or read-only. Consider the previous import; all of these subsequent assignment attempts will throw `TypeError`s:

```
import foofn, * as hello from "world";

foofn = 42;          // (runtime) TypeError!
hello.default = 42; // (runtime) TypeError!
hello.bar = 42;     // (runtime) TypeError!
hello.baz = 42;     // (runtime) TypeError!
```

Recall earlier in the “[exporting API Members](#)” section that we talked about how the `bar` and `baz` bindings are bound to the actual identifiers inside the `"world"` module. That means if the module changes those values, `hello.bar` and `hello.baz` now reference the updated values.

But the immutable/read-only nature of your local imported bindings enforces that you cannot change them from the imported bindings, hence the `TypeError`s. That’s pretty important, because without those protections, your changes would end up affecting all other consumers of the module (remember: singleton), which could create some very surprising side effects!

Moreover, though a module *can* change its API members from the inside, you should be very cautious of intentionally designing your modules in that fashion. ES6 modules are *intended* to be static, so deviations from that principle should be rare and should be carefully and verbosely documented.

Warning: There are module design philosophies where you actually intend to let a consumer change the value of a property on your API, or module APIs are designed to be “extended” by having other “plug-ins” add to the API namespace. As we just asserted, ES6 module APIs should be thought of and designed as static and unchangeable, which strongly restricts and discourages these alternative module design patterns. You can get around these limitations by exporting a plain object, which of course can then be changed at will. But be careful and think twice before going down that road.

Declarations that occur as a result of an `import` are “hoisted” (see the [Scope & Closures](#) title of this series). Consider:

```
foo();

import { foo } from "foo";
```

`foo()` can run because not only did the static resolution of the `import ..` statement figure out what `foo` is during compilation, but it also “hoisted” the

declaration to the top of the module's scope, thus making it available throughout the module.

Finally, the most basic form of the `import` looks like this:

```
import "foo";
```

This form does not actually import any of the module's bindings into your scope. It loads (if not already loaded), compiles (if not already compiled), and evaluates (if not already run) the `"foo"` module.

In general, that sort of import is probably not going to be terribly useful. There may be niche cases where a module's definition has side effects (such as assigning things to the `window`/global object). You could also envision using `import "foo"` as a sort of preload for a module that may be needed later.

Circular Module Dependency

A imports B. B imports A. How does this actually work?

I'll state off the bat that designing systems with intentional circular dependency is generally something I try to avoid. That having been said, I recognize there are reasons people do this and it can solve some sticky design situations.

Let's consider how ES6 handles this. First, module "A":

```
import bar from "B";

export default function foo(x) {
  if (x > 10) return bar( x - 1 );
  return x * 2;
}
```

Now, module "B":

```
import foo from "A";

export default function bar(y) {
  if (y > 5) return foo( y / 2 );
  return y * 3;
}
```

These two functions, `foo(..)` and `bar(..)`, would work as standard function declarations if they were in the same scope, because the declarations are "hoisted" to the whole scope and thus available to each other regardless of authoring order.

With modules, you have declarations in entirely different scopes, so ES6 has to do extra work to help make these circular references work.

In a rough conceptual sense, this is how circular `import` dependencies are validated and resolved:

- If the "A" module is loaded first, the first step is to scan the file and analyze all the exports, so it can register all those bindings available for import. Then it processes the `import .. from "B"`, which signals that it needs to go fetch "B".
- Once the engine loads "B", it does the same analysis of its export bindings. When it sees the `import .. from "A"`, it knows the API of "A" already, so it can verify the `import` is valid. Now that it knows the "B" API, it can also validate the `import .. from "B"` in the waiting "A" module.

In essence, the mutual imports, along with the static verification that's done to validate both `import` statements, virtually composes the two separate module scopes (via the bindings), such that `foo(..)` can call `bar(..)` and vice versa. This is symmetric to if they had originally been declared in the same scope.

Now let's try using the two modules together. First, we'll try `foo(..)`:

```
import foo from "foo";
foo( 25 );           // 11
```

Or we can try `bar(..)`:

```
import bar from "bar";
bar( 25 );           // 11.5
```

By the time either the `foo(25)` or `bar(25)` calls are executed, all the analysis/compilation of all modules has completed. That means `foo(..)` internally knows directly about `bar(..)` and `bar(..)` internally knows directly about `foo(..)`.

If all we need is to interact with `foo(..)`, then we only need to import the "foo" module. Likewise with `bar(..)` and the "bar" module.

Of course, we *can* import and use both of them if we want to:

```
import foo from "foo";
import bar from "bar";

foo( 25 );           // 11
bar( 25 );           // 11.5
```

The static loading semantics of the `import` statement mean that a "foo" and "bar" that mutually depend on each other via `import` will ensure that both are loaded, parsed, and compiled before either of them runs. So their circular dependency is statically resolved and this works as you'd expect.

Module Loading

We asserted at the beginning of this "Modules" section that the `import` statement uses a separate mechanism, provided by the hosting environment (browser, Node.js, etc.), to actually resolve the module specifier string into some useful

instruction for finding and loading the desired module. That mechanism is the system *Module Loader*.

The default module loader provided by the environment will interpret a module specifier as a URL if in the browser, and (generally) as a local filesystem path if on a server such as Node.js. The default behavior is to assume the loaded file is authored in the ES6 standard module format.

Moreover, you will be able to load a module into the browser via an HTML tag, similar to how current script programs are loaded. At the time of this writing, it's not fully clear if this tag will be `<script type="module">` or `<module>`. ES6 doesn't control that decision, but discussions in the appropriate standards bodies are already well along in parallel of ES6.

Whatever the tag looks like, you can be sure that under the covers it will use the default loader (or a customized one you've pre-specified, as we'll discuss in the next section).

Just like the tag you'll use in markup, the module loader itself is not specified by ES6. It is a separate, parallel standard (<http://whatwg.github.io/loader/>) controlled currently by the WHATWG browser standards group.

At the time of this writing, the following discussions reflect an early pass at the API design, and things are likely to change.

Loading Modules Outside of Modules One use for interacting directly with the module loader is if a non-module needs to load a module. Consider:

```
// normal script loaded in browser via `<script>`,
// `import` is illegal here

Reflect.Loader.import( "foo" ) // returns a promise for `"foo"`
  .then( function(foo){
    foo.bar();
  } );
```

The `Reflect.Loader.import(..)` utility imports the entire module onto the named parameter (as a namespace), just like the `import * as foo ..` namespace import we discussed earlier.

Note: The `Reflect.Loader.import(..)` utility returns a promise that is fulfilled once the module is ready. To import multiple modules, you can compose promises from multiple `Reflect.Loader.import(..)` calls using `Promise.all([..])`. For more information about Promises, see “Promises” in Chapter 4.

You can also use `Reflect.Loader.import(..)` in a real module to dynamically/conditionally load a module, where `import` itself would not work. You might, for instance, choose to load a module containing a polyfill for some ES7+ feature if a feature test reveals it's not defined by the current engine.

For performance reasons, you'll want to avoid dynamic loading whenever possible, as it hampers the ability of the JS engine to fire off early fetches from its static analysis.

Customized Loading Another use for directly interacting with the module loader is if you want to customize its behavior through configuration or even redefinition.

At the time of this writing, there's a polyfill for the module loader API being developed (<https://github.com/ModuleLoader/es6-module-loader>). While details are scarce and highly subject to change, we can explore what possibilities may eventually land.

The `Reflect.Loader.import(..)` call may support a second argument for specifying various options to customize the import/load task. For example:

```
Reflect.Loader.import( "foo", { address: "/path/to/foo.js" } )
  .then( function(foo){
    // ...
  } )
```

It's also expected that a customization will be provided (through some means) for hooking into the process of loading a module, where a translation/transpilation could occur after load but before the engine compiles the module.

For example, you could load something that's not already an ES6-compliant module format (e.g., CoffeeScript, TypeScript, CommonJS, AMD). Your translation step could then convert it to an ES6-compliant module for the engine to then process.

Classes

From nearly the beginning of JavaScript, syntax and development patterns have all strived (read: struggled) to put on a facade of supporting class-oriented development. With things like `new` and `instanceof` and a `.constructor` property, who couldn't help but be teased that JS had classes hidden somewhere inside its prototype system?

Of course, JS "classes" aren't nearly the same as classical classes. The differences are well documented, so I won't belabor that point any further here.

Note: To learn more about the patterns used in JS to fake "classes," and an alternative view of prototypes called "delegation," see the second half of the *this & Object Prototypes* title of this series.

class

Although JS's prototype mechanism doesn't work like traditional classes, that doesn't stop the strong tide of demand on the language to extend the syntactic sugar so that expressing "classes" looks more like real classes. Enter the ES6 `class` keyword and its associated mechanism.

This feature is the result of a highly contentious and drawn-out debate, and represents a smaller subset compromise from several strongly opposed views on how to approach JS classes. Most developers who want full classes in JS will find parts of the new syntax quite inviting, but will find important bits still missing. Don't worry, though. TC39 is already working on additional features to augment classes in the post-ES6 timeframe.

At the heart of the new ES6 class mechanism is the `class` keyword, which identifies a *block* where the contents define the members of a function's prototype. Consider:

```
class Foo {
  constructor(a,b) {
    this.x = a;
    this.y = b;
  }

  gimmeXY() {
    return this.x * this.y;
  }
}
```

Some things to note:

- `class Foo` implies creating a (special) function of the name `Foo`, much like you did pre-ES6.
- `constructor(..)` identifies the signature of that `Foo(..)` function, as well as its body contents.
- Class methods use the same "concise method" syntax available to object literals, as discussed in Chapter 2. This also includes the concise generator form as discussed earlier in this chapter, as well as the ES5 getter/setter syntax. However, class methods are non-enumerable whereas object methods are by default enumerable.
- Unlike object literals, there are no commas separating members in a `class` body! In fact, they're not even allowed.

The `class` syntax definition in the previous snippet can be roughly thought of as this pre-ES6 equivalent, which probably will look fairly familiar to those who've done prototype-style coding before:

```
function Foo(a,b) {
  this.x = a;
```

```

    this.y = b;
}

Foo.prototype.gimmeXY = function() {
    return this.x * this.y;
}

```

In either the pre-ES6 form or the new ES6 `class` form, this “class” can now be instantiated and used just as you’d expect:

```

var f = new Foo( 5, 15 );

f.x;                      // 5
f.y;                      // 15
f.gimmeXY();              // 75

```

Caution! Though `class Foo` seems much like `function Foo()`, there are important differences:

- A `Foo(..)` call of `class Foo` *must* be made with `new`, as the pre-ES6 option of `Foo.call(obj)` will *not* work.
- While `function Foo` is “hoisted” (see the *Scope & Closures* title of this series), `class Foo` is not; the `extends ..` clause specifies an expression that cannot be “hoisted.” So, you must declare a `class` before you can instantiate it.
- `class Foo` in the top global scope creates a lexical `Foo` identifier in that scope, but unlike `function Foo` does not create a global object property of that name.

The established `instanceof` operator still works with ES6 classes, because `class` just creates a constructor function of the same name. However, ES6 introduces a way to customize how `instanceof` works, using `Symbol.hasInstance` (see “Well-Known Symbols” in Chapter 7).

Another way of thinking about `class`, which I find more convenient, is as a *macro* that is used to automatically populate a `prototype` object. Optionally, it also wires up the `[[Prototype]]` relationship if using `extends` (see the next section).

An ES6 `class` isn’t really an entity itself, but a meta concept that wraps around other concrete entities, such as functions and properties, and ties them together.

Tip: In addition to the declaration form, a `class` can also be an expression, as in: `var x = class Y { .. }`. This is primarily useful for passing a class definition (technically, the constructor itself) as a function argument or assigning it to an object property.

extends and super

ES6 classes also have syntactic sugar for establishing the `[[Prototype]]` delegation link between two function prototypes – commonly mislabeled “inheritance” or confusingly labeled “prototype inheritance” – using the class-oriented familiar terminology `extends`:

```
class Bar extends Foo {
  constructor(a,b,c) {
    super( a, b );
    this.z = c;
  }

  gimmeXYZ() {
    return super.gimmeXY() * this.z;
  }
}

var b = new Bar( 5, 15, 25 );

b.x;                      // 5
b.y;                      // 15
b.z;                      // 25
b.gimmeXYZ();             // 1875
```

A significant new addition is `super`, which is actually something not directly possible pre-ES6 (without some unfortunate hack trade-offs). In the constructor, `super` automatically refers to the “parent constructor,” which in the previous example is `Foo(..)`. In a method, it refers to the “parent object,” such that you can then make a property/method access off it, such as `super.gimmeXY()`.

`Bar extends Foo` of course means to link the `[[Prototype]]` of `Bar.prototype` to `Foo.prototype`. So, `super` in a method like `gimmeXYZ()` specifically means `Foo.prototype`, whereas `super` means `Foo` when used in the `Bar` constructor.

Note: `super` is not limited to `class` declarations. It also works in object literals, in much the same way we’re discussing here. See “Object `super`” in Chapter 2 for more information.

There Be `super` Dragons It is not insignificant to note that `super` behaves differently depending on where it appears. In fairness, most of the time, that won’t be a problem. But surprises await if you deviate from a narrow norm.

There may be cases where in the constructor you would want to reference the `Foo.prototype`, such as to directly access one of its properties/methods. However, `super` in the constructor cannot be used in that way; `super.prototype` will not work. `super(..)` means roughly to call `new Foo(..)`, but isn’t actually a usable reference to `Foo` itself.

Symmetrically, you may want to reference the `Foo(..)` function from inside a non-constructor method. `super.constructor` will point at `Foo(..)` the function, but beware that this function can *only* be invoked with `new`. `new super.constructor(..)` would be valid, but it wouldn't be terribly useful in most cases, because you can't make that call use or reference the current `this` object context, which is likely what you'd want.

Also, `super` looks like it might be driven by a function's context just like `this` – that is, that they'd both be dynamically bound. However, `super` is not dynamic like `this` is. When a constructor or method makes a `super` reference inside it at declaration time (in the `class` body), that `super` is statically bound to that specific class hierarchy, and cannot be overridden (at least in ES6).

What does that mean? It means that if you're in the habit of taking a method from one "class" and "borrowing" it for another class by overriding its `this`, say with `call(..)` or `apply(..)`, that may very well create surprises if the method you're borrowing has a `super` in it. Consider this class hierarchy:

```
class ParentA {
  constructor() { this.id = "a"; }
  foo() { console.log( "ParentA:", this.id ); }
}

class ParentB {
  constructor() { this.id = "b"; }
  foo() { console.log( "ParentB:", this.id ); }
}

class ChildA extends ParentA {
  foo() {
    super.foo();
    console.log( "ChildA:", this.id );
  }
}

class ChildB extends ParentB {
  foo() {
    super.foo();
    console.log( "ChildB:", this.id );
  }
}

var a = new ChildA();           // ParentA: a
a.foo();                      // ChildA: a
var b = new ChildB();           // ParentB: b
b.foo();                      // ChildB: b
```

All seems fairly natural and expected in this previous snippet. However, if you try to borrow `b.foo()` and use it in the context of `a` – by virtue of dynamic `this` binding, such borrowing is quite common and used in many different ways, including mixins most notably – you may find this result an ugly surprise:

```
// borrow `b.foo()` to use in `a` context
b.foo.call( a );           // ParentB: a
                           // ChildB: a
```

As you can see, the `this.id` reference was dynamically rebound so that : `a` is reported in both cases instead of : `b`. But `b.foo()`'s `super.foo()` reference wasn't dynamically rebound, so it still reported `ParentB` instead of the expected `ParentA`.

`super` references `super`, it is statically bound to the `ChildB/ParentB` hierarchy and cannot be used against the `ChildA/ParentA` hierarchy. There is no ES6 solution to this limitation.

`super` seems to work intuitively if you have a static class hierarchy with no cross-pollination. But in all fairness, one of the main benefits of doing `this`-aware coding is exactly that sort of flexibility. Simply, `class + super` requires you to avoid such techniques.

The choice boils down to narrowing your object design to these static hierarchies – `class`, `extends`, and `super` will be quite nice – or dropping all attempts to “fake” classes and instead embrace dynamic and flexible, classless objects and `[[Prototype]]` delegation (see the *this & Object Prototypes* title of this series).

Subclass Constructor Constructors are not required for classes or subclasses; a default constructor is substituted in both cases if omitted. However, the default substituted constructor is different for a direct class versus an extended class.

Specifically, the default subclass constructor automatically calls the parent constructor, and passes along any arguments. In other words, you could think of the default subclass constructor sort of like this:

```
constructor(...args) {
  super(...args);
}
```

This is an important detail to note. Not all class languages have the subclass constructor automatically call the parent constructor. C++ does, but Java does not. But more importantly, in pre-ES6 classes, such automatic “parent constructor” calling does not happen. Be careful when converting to ES6 `class` if you've been relying on such calls *not* happening.

Another perhaps surprising deviation/limitation of ES6 subclass constructors: in a constructor of a subclass, you cannot access `this` until `super(..)` has been called. The reason is nuanced and complicated, but it boils down to the

fact that the parent constructor is actually the one creating/initializing your instance's `this`. Pre-ES6, it works oppositely; the `this` object is created by the “subclass constructor,” and then you call a “parent constructor” with the context of the “subclass” `this`.

Let's illustrate. This works pre-ES6:

```
function Foo() {
  this.a = 1;
}

function Bar() {
  this.b = 2;
  Foo.call( this );
}

// `Bar` "extends" `Foo`
Bar.prototype = Object.create( Foo.prototype );
```

But this ES6 equivalent is not allowed:

```
class Foo {
  constructor() { this.a = 1; }

}

class Bar extends Foo {
  constructor() {
    this.b = 2;           // not allowed before `super()`
    super();             // to fix swap these two statements
  }
}
```

In this case, the fix is simple. Just swap the two statements in the subclass `Bar` constructor. However, if you've been relying pre-ES6 on being able to skip calling the “parent constructor,” beware because that won't be allowed anymore.

extending Natives One of the most heralded benefits to the new `class` and `extend` design is the ability to (finally!) subclass the built-in natives, like `Array`. Consider:

```
class MyCoolArray extends Array {
  first() { return this[0]; }
  last() { return this[this.length - 1]; }
}

var a = new MyCoolArray( 1, 2, 3 );

a.length;           // 3
```

```

a;           // [1,2,3]

a.first();    // 1
a.last();    // 3

```

Prior to ES6, a fake “subclass” of `Array` using manual object creation and linking to `Array.prototype` only partially worked. It missed out on the special behaviors of a real array, such as the automatically updating `length` property. ES6 subclasses should fully work with “inherited” and augmented behaviors as expected!

Another common pre-ES6 “subclass” limitation is with the `Error` object, in creating custom error “subclasses.” When genuine `Error` objects are created, they automatically capture special `stack` information, including the line number and file where the error is created. Pre-ES6 custom error “subclasses” have no such special behavior, which severely limits their usefulness.

ES6 to the rescue:

```

class Oops extends Error {
  constructor(reason) {
    super(reason);
    this.oops = reason;
  }
}

// later:
var ouch = new Oops( "I messed up!" );
throw ouch;

```

The `ouch` custom error object in this previous snippet will behave like any other genuine error object, including capturing `stack`. That’s a big improvement!

`new.target`

ES6 introduces a new concept called a *meta property* (see Chapter 7), in the form of `new.target`.

If that looks strange, it is; pairing a keyword with a `.` and a property name is definitely an out-of-the-ordinary pattern for JS.

`new.target` is a new “magical” value available in all functions, though in normal functions it will always be `undefined`. In any constructor, `new.target` always points at the constructor that `new` actually directly invoked, even if the constructor is in a parent class and was delegated to by a `super(..)` call from a child constructor. Consider:

```

class Foo {
  constructor() {
    console.log( "Foo: ", new.target.name );
  }
}

```

```

        }
    }

    class Bar extends Foo {
        constructor() {
            super();
            console.log( "Bar: ", new.target.name );
        }
        baz() {
            console.log( "baz: ", new.target );
        }
    }

    var a = new Foo();
    // Foo: Foo

    var b = new Bar();
    // Foo: Bar    <-- respects the `new` call-site
    // Bar: Bar

    b.baz();
    // baz: undefined

```

The `new.target` meta property doesn't have much purpose in class constructors, except accessing a static property/method (see the next section).

If `new.target` is `undefined`, you know the function was not called with `new`. You can then force a `new` invocation if that's necessary.

static

When a subclass `Bar` extends a parent class `Foo`, we already observed that `Bar.prototype` is `[[Prototype]]`-linked to `Foo.prototype`. But additionally, `Bar()` is `[[Prototype]]`-linked to `Foo()`. That part may not have such an obvious reasoning.

However, it's quite useful in the case where you declare `static` methods (not just properties) for a class, as these are added directly to that class's function object, not to the function object's `prototype` object. Consider:

```

class Foo {
    static cool() { console.log( "cool" ); }
    wow() { console.log( "wow" ); }
}

class Bar extends Foo {
    static awesome() {
        super.cool();

```

```

        console.log( "awesome" );
    }
    neat() {
        super.wow();
        console.log( "neat" );
    }
}

Foo.cool();           // "cool"
Bar.cool();           // "cool"
Bar.awesome();        // "cool"
                     // "awesome"

var b = new Bar();
b.neat();            // "wow"
                     // "neat"

b.awesome;           // undefined
b.cool;              // undefined

```

Be careful not to get confused that `static` members are on the class's prototype chain. They're actually on the dual/parallel chain between the function constructors.

Symbol.species Constructor Getter One place where `static` can be useful is in setting the `Symbol.species` getter (known internally in the specification as `@@species`) for a derived (child) class. This capability allows a child class to signal to a parent class what constructor should be used – when not intending the child class's constructor itself – if any parent class method needs to vend a new instance.

For example, many methods on `Array` create and return a new `Array` instance. If you define a derived class from `Array`, but you want those methods to continue to vend actual `Array` instances instead of from your derived class, this works:

```

class MyCoolArray extends Array {
    // force `species` to be parent constructor
    static get [Symbol.species]() { return Array; }
}

var a = new MyCoolArray( 1, 2, 3 ),
    b = a.map( function(v){ return v * 2; } );

b instanceof MyCoolArray; // false
b instanceof Array;      // true

```

To illustrate how a parent class method can use a child's species declaration

somewhat like `Array#map(..)` is doing, consider:

```
class Foo {
  // defer `species` to derived constructor
  static get [Symbol.species]() { return this; }
  spawn() {
    return new this.constructor[Symbol.species]();
  }
}

class Bar extends Foo {
  // force `species` to be parent constructor
  static get [Symbol.species]() { return Foo; }
}

var a = new Foo();
var b = a.spawn();
b instanceof Foo;           // true

var x = new Bar();
var y = x.spawn();
y instanceof Bar;           // false
y instanceof Foo;           // true
```

The parent class `Symbol.species` does `return this` to defer to any derived class, as you'd normally expect. `Bar` then overrides to manually declare `Foo` to be used for such instance creation. Of course, a derived class can still vend instances of itself using `new this.constructor(..)`.

Review

ES6 introduces several new features that aid in code organization:

- Iterators provide sequential access to data or operations. They can be consumed by new language features like `for..of` and `....`.
- Generators are locally pause/resume capable functions controlled by an iterator. They can be used to programmatically (and interactively, through `yield/next(..)` message passing) *generate* values to be consumed via iteration.
- Modules allow private encapsulation of implementation details with a publicly exported API. Module definitions are file-based, singleton instances, and statically resolved at compile time.
- Classes provide cleaner syntax around prototype-based coding. The addition of `super` also solves tricky issues with relative references in the `[[Prototype]]` chain.

These new tools should be your first stop when trying to improve the architecture of your JS projects by embracing ES6.

You Don't Know JS: ES6 & Beyond

Chapter 4: Async Flow Control

It's no secret if you've written any significant amount of JavaScript that asynchronous programming is a required skill. The primary mechanism for managing asynchrony has been the function callback.

However, ES6 adds a new feature that helps address significant shortcomings in the callbacks-only approach to async: *Promises*. In addition, we can revisit generators (from the previous chapter) and see a pattern for combining the two that's a major step forward in async flow control programming in JavaScript.

Promises

Let's clear up some misconceptions: Promises are not about replacing callbacks. Promises provide a trustable intermediary – that is, between your calling code and the async code that will perform the task – to manage callbacks.

Another way of thinking about a Promise is as an event listener, on which you can register to listen for an event that lets you know when a task has completed. It's an event that will only ever fire once, but it can be thought of as an event nonetheless.

Promises can be chained together, which can sequence a series of asynchronously completing steps. Together with higher-level abstractions like the `all(..)` method (in classic terms, a “gate”) and the `race(..)` method (in classic terms, a “latch”), promise chains provide a mechanism for async flow control.

Yet another way of conceptualizing a Promise is that it's a *future value*, a time-independent container wrapped around a value. This container can be reasoned about identically whether the underlying value is final or not. Observing the resolution of a Promise extracts this value once available. In other words, a Promise is said to be the async version of a sync function's return value.

A Promise can only have one of two possible resolution outcomes: fulfilled or rejected, with an optional single value. If a Promise is fulfilled, the final value is called a fulfillment. If it's rejected, the final value is called a reason (as in, a “reason for rejection”). Promises can only be resolved (fulfillment or rejection) *once*. Any further attempts to fulfill or reject are simply ignored. Thus, once a Promise is resolved, it's an immutable value that cannot be changed.

Clearly, there are several different ways to think about what a Promise is. No single perspective is fully sufficient, but each provides a separate aspect of the whole. The big takeaway is that they offer a significant improvement over callbacks-only async, namely that they provide order, predictability, and trustability.

Making and Using Promises

To construct a promise instance, use the `Promise(..)` constructor:

```
var p = new Promise( function pr(resolve,reject){  
    // ..  
} );
```

The `Promise(..)` constructor takes a single function (`pr(..)`), which is called immediately and receives two control functions as arguments, usually named `resolve(..)` and `reject(..)`. They are used as:

- If you call `reject(..)`, the promise is rejected, and if any value is passed to `reject(..)`, it is set as the reason for rejection.
- If you call `resolve(..)` with no value, or any non-promise value, the promise is fulfilled.
- If you call `resolve(..)` and pass another promise, this promise simply adopts the state – whether immediate or eventual – of the passed promise (either fulfillment or rejection).

Here's how you'd typically use a promise to refactor a callback-reliant function call. If you start out with an `ajax(..)` utility that expects to be able to call an error-first style callback:

```
function ajax(url,cb) {  
    // make request, eventually call `cb(..)`  
}  
  
// ..  
  
ajax( "http://some.url.1", function handler(err,contents){  
    if (err) {  
        // handle ajax error  
    }  
    else {  
        // handle `contents` success  
    }  
} );
```

You can convert it to:

```
function ajax(url) {  
    return new Promise( function pr(resolve,reject){  
        // make request, eventually call  
        // either `resolve(..)` or `reject(..)`  
    } );  
}  
  
// ..
```

```

ajax( "http://some.url.1" )
.then(
  function fulfilled(contents){
    // handle `contents` success
  },
  function rejected(reason){
    // handle ajax error reason
  }
);

```

Promises have a `then(..)` method that accepts one or two callback functions. The first function (if present) is treated as the handler to call if the promise is fulfilled successfully. The second function (if present) is treated as the handler to call if the promise is rejected explicitly, or if any error/exception is caught during resolution.

If one of the arguments is omitted or otherwise not a valid function – typically you'll use `null` instead – a default placeholder equivalent is used. The default success callback passes its fulfillment value along and the default error callback propagates its rejection reason along.

The shorthand for calling `then(null, handleRejection)` is `catch(handleRejection)`.

Both `then(..)` and `catch(..)` automatically construct and return another promise instance, which is wired to receive the resolution from whatever the return value is from the original promise's fulfillment or rejection handler (whichever is actually called). Consider:

```

ajax( "http://some.url.1" )
.then(
  function fulfilled(contents){
    return contents.toUpperCase();
  },
  function rejected(reason){
    return "DEFAULT VALUE";
  }
)
.then( function fulfilled(data){
  // handle data from original promise's
  // handlers
} );

```

In this snippet, we're returning an immediate value from either `fulfilled(..)` or `rejected(..)`, which then is received on the next event turn in the second `then(..)`'s `fulfilled(..)`. If we instead return a new promise, that new promise is subsumed and adopted as the resolution:

```
ajax( "http://some.url.1" )
```

```

.then(
  function fulfilled(contents){
    return ajax(
      "http://some.url.2?v=" + contents
    );
  },
  function rejected(reason){
    return ajax(
      "http://backup.url.3?err=" + reason
    );
  }
)
.then( function fulfilled(contents){
  // `contents` comes from the subsequent
  // `ajax(..)` call, whichever it was
} );

```

It's important to note that an exception (or rejected promise) in the first `fulfilled(..)` will *not* result in the first `rejected(..)` being called, as that handler only responds to the resolution of the first original promise. Instead, the second promise, which the second `then(..)` is called against, receives that rejection.

In this previous snippet, we are not listening for that rejection, which means it will be silently held onto for future observation. If you never observe it by calling a `then(..)` or `catch(..)`, then it will go unhandled. Some browser developer consoles may detect these unhandled rejections and report them, but this is not reliably guaranteed; you should always observe promise rejections.

Note: This was just a brief overview of Promise theory and behavior. For a much more in-depth exploration, see Chapter 3 of the *Async & Performance* title of this series.

Thenables

Promises are genuine instances of the `Promise(..)` constructor. However, there are promise-like objects called *thenables* that generally can interoperate with the Promise mechanisms.

Any object (or function) with a `then(..)` function on it is assumed to be a thenable. Any place where the Promise mechanisms can accept and adopt the state of a genuine promise, they can also handle a thenable.

Thenables are basically a general label for any promise-like value that may have been created by some other system than the actual `Promise(..)` constructor. In that perspective, a thenable is generally less trustable than a genuine Promise. Consider this misbehaving thenable, for example:

```
var th = {
```

```

    then: function thener( fulfilled ) {
      // call `fulfilled(..)` once every 100ms forever
      setInterval( fulfilled, 100 );
    }
  };

```

If you received that thenable and chained it with `th.then(..)`, you'd likely be surprised that your fulfillment handler is called repeatedly, when normal Promises are supposed to only ever be resolved once.

Generally, if you're receiving what purports to be a promise or thenable back from some other system, you shouldn't just trust it blindly. In the next section, we'll see a utility included with ES6 Promises that helps address this trust concern.

But to further understand the perils of this issue, consider that *any* object in *any* piece of code that's ever been defined to have a method on it called `then(..)` can be potentially confused as a thenable – if used with Promises, of course – regardless of if that thing was ever intended to even remotely be related to Promise-style async coding.

Prior to ES6, there was never any special reservation made on methods called `then(..)`, and as you can imagine there's been at least a few cases where that method name has been chosen prior to Promises ever showing up on the radar screen. The most likely case of mistaken thenable will be async libraries that use `then(..)` but which are not strictly Promises-compliant – there are several out in the wild.

The onus will be on you to guard against directly using values with the Promise mechanism that would be incorrectly assumed to be a thenable.

Promise API

The `Promise` API also provides some static methods for working with Promises.

`Promise.resolve(..)` creates a promise resolved to the value passed in. Let's compare how it works to the more manual approach:

```

var p1 = Promise.resolve( 42 );

var p2 = new Promise( function pr(resolve){
  resolve( 42 );
} );

```

`p1` and `p2` will have essentially identical behavior. The same goes for resolving with a promise:

```

var theP = ajax( ... );

var p1 = Promise.resolve( theP );

```

```
var p2 = new Promise( function pr(resolve){
    resolve( theP );
} );
```

Tip: `Promise.resolve(..)` is the solution to the thenable trust issue raised in the previous section. Any value that you are not already certain is a trustable promise – even if it could be an immediate value – can be normalized by passing it to `Promise.resolve(..)`. If the value is already a recognizable promise or thenable, its state/resolution will simply be adopted, insulating you from misbehavior. If it's instead an immediate value, it will be “wrapped” in a genuine promise, thereby normalizing its behavior to be `async`.

`Promise.reject(..)` creates an immediately rejected promise, the same as its `Promise(..)` constructor counterpart:

```
var p1 = Promise.reject( "Oops" );

var p2 = new Promise( function pr(resolve,reject){
    reject( "Oops" );
} );
```

While `resolve(..)` and `Promise.resolve(..)` can accept a promise and adopt its state/resolution, `reject(..)` and `Promise.reject(..)` do not differentiate what value they receive. So, if you reject with a promise or thenable, the promise/thenable itself will be set as the rejection reason, not its underlying value.

`Promise.all([..])` accepts an array of one or more values (e.g., immediate values, promises, thenables). It returns a promise back that will be fulfilled if all the values fulfill, or reject immediately once the first of any of them rejects.

Starting with these values/promises:

```
var p1 = Promise.resolve( 42 );
var p2 = new Promise( function pr(resolve){
    setTimeout( function(){
        resolve( 43 );
    }, 100 );
} );
var v3 = 44;
var p4 = new Promise( function pr(resolve,reject){
    setTimeout( function(){
        reject( "Oops" );
    }, 10 );
} );
```

Let's consider how `Promise.all([..])` works with combinations of those values:

```

Promise.all( [p1,p2,v3] )
.then( function fulfilled(vals){
    console.log( vals );           // [42,43,44]
} );

Promise.all( [p1,p2,v3,p4] )
.then(
    function fulfilled(vals){
        // never gets here
    },
    function rejected(reason){
        console.log( reason );      // Oops
    }
);

```

While `Promise.all([...])` waits for all fulfillments (or the first rejection), `Promise.race([...])` waits only for either the first fulfillment or rejection. Consider:

```

// NOTE: re-setup all test values to
// avoid timing issues misleading you!

Promise.race( [p2,p1,v3] )
.then( function fulfilled(val){
    console.log( val );           // 42
} );

Promise.race( [p2,p4] )
.then(
    function fulfilled(val){
        // never gets here
    },
    function rejected(reason){
        console.log( reason );      // Oops
    }
);

```

Warning: While `Promise.all([])` will fulfill right away (with no values), `Promise.race([])` will hang forever. This is a strange inconsistency, and speaks to the suggestion that you should never use these methods with empty arrays.

Generators + Promises

It *is* possible to express a series of promises in a chain to represent the async flow control of your program. Consider:

```

step1()
.then(
  step2,
  step1Failed
)
.then(
  function step3(msg) {
    return Promise.all( [
      step3a( msg ),
      step3b( msg ),
      step3c( msg )
    ] )
  }
)
.then(step4);

```

However, there's a much better option for expressing async flow control, and it will probably be much more preferable in terms of coding style than long promise chains. We can use what we learned in Chapter 3 about generators to express our async flow control.

The important pattern to recognize: a generator can yield a promise, and that promise can then be wired to resume the generator with its fulfillment value.

Consider the previous snippet's async flow control expressed with a generator:

```

function *main() {

  try {
    var ret = yield step1();
  }
  catch (err) {
    ret = yield step1Failed( err );
  }

  ret = yield step2( ret );

  // step 3
  ret = yield Promise.all( [
    step3a( ret ),
    step3b( ret ),
    step3c( ret )
  ] );

  yield step4( ret );
}

```

On the surface, this snippet may seem more verbose than the promise chain

equivalent in the earlier snippet. However, it offers a much more attractive – and more importantly, a more understandable and reason-able – synchronous-looking coding style (with = assignment of “return” values, etc.) That’s especially true in that `try..catch` error handling can be used across those hidden `async` boundaries.

Why are we using Promises with the generator? It’s certainly possible to do `async` generator coding without Promises.

Promises are a trustable system that uninverts the inversion of control of normal callbacks or thunks (see the *Async & Performance* title of this series). So, combining the trustability of Promises and the synchronicity of code in generators effectively addresses all the major deficiencies of callbacks. Also, utilities like `Promise.all([..])` are a nice, clean way to express concurrency at a generator’s single `yield` step.

So how does this magic work? We’re going to need a *runner* that can run our generator, receive a `yielded` promise, and wire it up to resume the generator with either the `fulfillment` success value, or throw an error into the generator with the rejection reason.

Many `async`-capable utilities/libraries have such a “runner”; for example, `Q.spawn(..)` and my `asynquence`’s `runner(..)` plug-in. But here’s a stand-alone runner to illustrate how the process works:

```
function run(gen) {
  var args = [].slice.call( arguments, 1 ), it;

  it = gen.apply( this, args );

  return Promise.resolve()
    .then( function handleNext(value){
      var next = it.next( value );

      return (function handleResult(next){
        if (next.done) {
          return next.value;
        }
        else {
          return Promise.resolve( next.value )
            .then(
              handleNext,
              function handleErr(err) {
                return Promise.resolve(
                  it.throw( err )
                )
                .then( handleResult );
              }
            );
        }
      });
    });
}
```

```

        );
    }
})( next );
} );
}

```

Note: For a more prolifically commented version of this utility, see the *Async & Performance* title of this series. Also, the run utilities provided with various async libraries are often more powerful/capable than what we've shown here. For example, asynquence's `runner(..)` can handle yielded promises, sequences, thunks, and immediate (non-promise) values, giving you ultimate flexibility.

So now running `*main()` as listed in the earlier snippet is as easy as:

```

run( main )
.then(
  function fulfilled(){
    // `*main()` completed successfully
  },
  function rejected(reason){
    // Oops, something went wrong
  }
);

```

Essentially, anywhere that you have more than two asynchronous steps of flow control logic in your program, you can *and should* use a promise-yielding generator driven by a run utility to express the flow control in a synchronous fashion. This will make for much easier to understand and maintain code.

This yield-a-promise-resume-the-generator pattern is going to be so common and so powerful, the next version of JavaScript after ES6 is almost certainly going to introduce a new function type that will do it automatically without needing the run utility. We'll cover `async functions` (as they're expected to be called) in Chapter 8.

Review

As JavaScript continues to mature and grow in its widespread adoption, asynchronous programming is more and more of a central concern. Callbacks are not fully sufficient for these tasks, and totally fall down the more sophisticated the need.

Thankfully, ES6 adds Promises to address one of the major shortcomings of callbacks: lack of trust in predictable behavior. Promises represent the future completion value from a potentially async task, normalizing behavior across sync and async boundaries.

But it's the combination of Promises with generators that fully realizes the

benefits of rearranging our async flow control code to de-emphasize and abstract away that ugly callback soup (aka “hell”).

Right now, we can manage these interactions with the aide of various async libraries’ runners, but JavaScript is eventually going to support this interaction pattern with dedicated syntax alone!

You Don't Know JS: ES6 & Beyond

Chapter 5: Collections

Structured collection and access to data is a critical component of just about any JS program. From the beginning of the language up to this point, the array and the object have been our primary mechanism for creating data structures. Of course, many higher-level data structures have been built on top of these, as user-land libraries.

As of ES6, some of the most useful (and performance-optimizing!) data structure abstractions have been added as native components of the language.

We'll start this chapter first by looking at *TypedArrays*, technically contemporary to ES5 efforts several years ago, but only standardized as companions to WebGL and not JavaScript itself. As of ES6, these have been adopted directly by the language specification, which gives them first-class status.

Maps are like objects (key/value pairs), but instead of just a string for the key, you can use any value – even another object or map! Sets are similar to arrays (lists of values), but the values are unique; if you add a duplicate, it's ignored. There are also weak (in relation to memory/garbage collection) counterparts: WeakMap and WeakSet.

TypedArrays

As we cover in the *Types & Grammar* title of this series, JS does have a set of built-in types, like `number` and `string`. It'd be tempting to look at a feature named “typed array” and assume it means an array of a specific type of values, like an array of only strings.

However, typed arrays are really more about providing structured access to binary data using array-like semantics (indexed access, etc.). The “type” in the name refers to a “view” layered on type of the bucket of bits, which is essentially a mapping of whether the bits should be viewed as an array of 8-bit signed integers, 16-bit signed integers, and so on.

How do you construct such a bit-bucket? It's called a “buffer,” and you construct it most directly with the `ArrayBuffer(..)` constructor:

```
var buf = new ArrayBuffer( 32 );
buf.byteLength; // 32
```

`buf` is now a binary buffer that is 32-bytes long (256-bits), that's pre-initialized to all 0s. A buffer by itself doesn't really allow you any interaction except for checking its `byteLength` property.

Tip: Several web platform features use or return array buffers, such as `FileReader#readAsArrayBuffer(..)`, `XMLHttpRequest#send(..)`, and `ImageData` (canvas data).

But on top of this array buffer, you can then layer a “view,” which comes in the form of a typed array. Consider:

```
var arr = new Uint16Array( buf );
arr.length; // 16
```

`arr` is a typed array of 16-bit unsigned integers mapped over the 256-bit `buf` buffer, meaning you get 16 elements.

Endianness

It's very important to understand that the `arr` is mapped using the endian-setting (big-endian or little-endian) of the platform the JS is running on. This can be an issue if the binary data is created with one endianness but interpreted on a platform with the opposite endianness.

Endian means if the low-order byte (collection of 8-bits) of a multi-byte number – such as the 16-bit unsigned ints we created in the earlier snippet – is on the right or the left of the number's bytes.

For example, let's imagine the base-10 number 3085, which takes 16-bits to represent. If you have just one 16-bit number container, it'd be represented in binary as 0000110000001101 (hexadecimal 0c0d) regardless of endianness.

But if 3085 was represented with two 8-bit numbers, the endianness would significantly affect its storage in memory:

- 0000110000001101 / 0c0d (big endian)
- 0000110100001100 / 0d0c (little endian)

If you received the bits of 3085 as 0000110100001100 from a little-endian system, but you layered a view on top of it in a big-endian system, you'd instead see value 3340 (base-10) and 0d0c (base-16).

Little endian is the most common representation on the web these days, but there are definitely browsers where that's not true. It's important that you understand the endianness of both the producer and consumer of a chunk of binary data.

From MDN, here's a quick way to test the endianness of your JavaScript:

```
var littleEndian = (function() {
  var buffer = new ArrayBuffer( 2 );
  new DataView( buffer ).setInt16( 0, 256, true );
  return new Int16Array( buffer )[0] === 256;
})();
```

`littleEndian` will be `true` or `false`; for most browsers, it should return `true`. This test uses `DataView(..)`, which allows more low-level, fine-grained control over accessing (setting/getting) the bits from the view you layer over the buffer.

The third parameter of the `setInt16(..)` method in the previous snippet is for telling the `DataView` what endianness you’re wanting it to use for that operation.

Warning: Do not confuse endianness of underlying binary storage in array buffers with how a given number is represented when exposed in a JS program. For example, `(3085).toString(2)` returns `"110000001101"`, which with an assumed leading four "0"s appears to be the big-endian representation. In fact, this representation is based on a single 16-bit view, not a view of two 8-bit bytes. The `DataView` test above is the best way to determine endianness for your JS environment.

Multiple Views

A single buffer can have multiple views attached to it, such as:

```
var buf = new ArrayBuffer( 2 );

var view8 = new Uint8Array( buf );
var view16 = new Uint16Array( buf );

view16[0] = 3085;
view8[0];                      // 13
view8[1];                      // 12

view8[0].toString( 16 );        // "d"
view8[1].toString( 16 );        // "c"

// swap (as if endian!)
var tmp = view8[0];
view8[0] = view8[1];
view8[1] = tmp;

view16[0];                      // 3340
```

The typed array constructors have multiple signature variations. We’ve shown so far only passing them an existing buffer. However, that form also takes two extra parameters: `byteOffset` and `length`. In other words, you can start the typed array view at a location other than 0 and you can make it span less than the full length of the buffer.

If the buffer of binary data includes data in non-uniform size/location, this technique can be quite useful.

For example, consider a binary buffer that has a 2-byte number (aka “word”) at the beginning, followed by two 1-byte numbers, followed by a 32-bit floating point number. Here’s how you can access that data with multiple views on the same buffer, offsets, and lengths:

```
var first = new Uint16Array( buf, 0, 2 )[0],
```

```

second = new Uint8Array( buf, 2, 1 )[0],
third = new Uint8Array( buf, 3, 1 )[0],
fourth = new Float32Array( buf, 4, 4 )[0];

```

TypedArray Constructors

In addition to the `(buffer, [offset, [length]])` form examined in the previous section, typed array constructors also support these forms:

- `[constructor](length)`: Creates a new view over a new buffer of `length` bytes
- `[constructor](typedArr)`: Creates a new view and buffer, and copies the contents from the `typedArr` view
- `[constructor](obj)`: Creates a new view and buffer, and iterates over the array-like or object `obj` to copy its contents

The following typed array constructors are available as of ES6:

- `Int8Array` (8-bit signed integers), `Uint8Array` (8-bit unsigned integers)
 - `Uint8ClampedArray` (8-bit unsigned integers, each value clamped on setting to the 0-255 range)
- `Int16Array` (16-bit signed integers), `Uint16Array` (16-bit unsigned integers)
- `Int32Array` (32-bit signed integers), `Uint32Array` (32-bit unsigned integers)
- `Float32Array` (32-bit floating point, IEEE-754)
- `Float64Array` (64-bit floating point, IEEE-754)

Instances of typed array constructors are almost the same as regular native arrays. Some differences include having a fixed length and the values all being of the same “type.”

However, they share most of the same `prototype` methods. As such, you likely will be able to use them as regular arrays without needing to convert.

For example:

```

var a = new Int32Array( 3 );
a[0] = 10;
a[1] = 20;
a[2] = 30;

a.map( function(v){
    console.log( v );
} );
// 10 20 30

a.join( "-" );
// "10-20-30"

```

Warning: You can't use certain `Array.prototype` methods with TypedArrays that don't make sense, such as the mutators (`splice(..)`, `push(..)`, etc.) and `concat(..)`.

Be aware that the elements in TypedArrays really are constrained to the declared bit sizes. If you have a `Uint8Array` and try to assign something larger than an 8-bit value into one of its elements, the value wraps around so as to stay within the bit length.

This could cause problems if you were trying to, for instance, square all the values in a TypedArray. Consider:

```
var a = new Uint8Array( 3 );
a[0] = 10;
a[1] = 20;
a[2] = 30;

var b = a.map( function(v){
    return v * v;
} );

b;           // [100, 144, 132]
```

The 20 and 30 values, when squared, resulted in bit overflow. To get around such a limitation, you can use the `TypedArray#from(..)` function:

```
var a = new Uint8Array( 3 );
a[0] = 10;
a[1] = 20;
a[2] = 30;

var b = Uint16Array.from( a, function(v){
    return v * v;
} );

b;           // [100, 400, 900]
```

See the “`Array.from(..)` Static Function” section in Chapter 6 for more information about the `Array.from(..)` that is shared with TypedArrays. Specifically, the “Mapping” section explains the mapping function accepted as its second argument.

One interesting behavior to consider is that TypedArrays have a `sort(..)` method much like regular arrays, but this one defaults to numeric sort comparisons instead of coercing values to strings for lexicographic comparison. For example:

```
var a = [ 10, 1, 2, ];
a.sort();           // [1,10,2]
```

```
var b = new Uint8Array( [ 10, 1, 2 ] );
b.sort();                                // [1,2,10]
```

The `TypedArray#sort(..)` takes an optional compare function argument just like `Array#sort(..)`, which works in exactly the same way.

Maps

If you have a lot of JS experience, you know that objects are the primary mechanism for creating unordered key/value-pair data structures, otherwise known as maps. However, the major drawback with objects-as-maps is the inability to use a non-string value as the key.

For example, consider:

```
var m = {};

var x = { id: 1 },
y = { id: 2 };

m[x] = "foo";
m[y] = "bar";

m[x];                                // "bar"
m[y];                                // "bar"
```

What's going on here? The two objects `x` and `y` both stringify to "[object Object]", so only that one key is being set in `m`.

Some have implemented fake maps by maintaining a parallel array of non-string keys alongside an array of the values, such as:

```
var keys = [], vals = [];

var x = { id: 1 },
y = { id: 2 };

keys.push( x );
vals.push( "foo" );

keys.push( y );
vals.push( "bar" );

keys[0] === x;                         // true
vals[0];                                // "foo"

keys[1] === y;                         // true
vals[1];                                // "bar"
```

Of course, you wouldn't want to manage those parallel arrays yourself, so you could define a data structure with methods that automatically do the management under the covers. Besides having to do that work yourself, the main drawback is that access is no longer $O(1)$ time-complexity, but instead is $O(n)$.

But as of ES6, there's no longer any need to do this! Just use `Map(..)`:

```
var m = new Map();

var x = { id: 1 },
y = { id: 2 };

m.set( x, "foo" );
m.set( y, "bar" );

m.get( x );           // "foo"
m.get( y );           // "bar"
```

The only drawback is that you can't use the `[]` bracket access syntax for setting and retrieving values. But `get(..)` and `set(..)` work perfectly suitably instead.

To delete an element from a map, don't use the `delete` operator, but instead use the `delete(..)` method:

```
m.set( x, "foo" );
m.set( y, "bar" );

m.delete( y );
```

You can clear the entire map's contents with `clear()`. To get the length of a map (i.e., the number of keys), use the `size` property (not `length`):

```
m.set( x, "foo" );
m.set( y, "bar" );
m.size;           // 2

m.clear();
m.size;           // 0
```

The `Map(..)` constructor can also receive an iterable (see "Iterators" in Chapter 3), which must produce a list of arrays, where the first item in each array is the key and the second item is the value. This format for iteration is identical to that produced by the `entries()` method, explained in the next section. That makes it easy to make a copy of a map:

```
var m2 = new Map( m.entries() );

// same as:
var m2 = new Map( m );
```

Because a map instance is an iterable, and its default iterator is the same as `entries()`, the second shorter form is more preferable.

Of course, you can just manually specify an `entries` list (array of key/value arrays) in the `Map(..)` constructor form:

```
var x = { id: 1 },
      y = { id: 2 };

var m = new Map( [
  [ x, "foo" ],
  [ y, "bar" ]
] );

m.get( x );           // "foo"
m.get( y );           // "bar"
```

Map Values

To get the list of values from a map, use `values(..)`, which returns an iterator. In Chapters 2 and 3, we covered various ways to process an iterator sequentially (like an array), such as the `...` spread operator and the `for..of` loop. Also, “Arrays” in Chapter 6 covers the `Array.from(..)` method in detail. Consider:

```
var m = new Map();

var x = { id: 1 },
      y = { id: 2 };

m.set( x, "foo" );
m.set( y, "bar" );

var vals = [ ...m.values() ];

vals;           // ["foo", "bar"]
Array.from( m.values() ); // ["foo", "bar"]
```

As discussed in the previous section, you can iterate over a map’s entries using `entries()` (or the default map iterator). Consider:

```
var m = new Map();

var x = { id: 1 },
      y = { id: 2 };

m.set( x, "foo" );
m.set( y, "bar" );
```

```

var vals = [ ...m.entries() ];

vals[0][0] === x;           // true
vals[0][1];                // "foo"

vals[1][0] === y;           // true
vals[1][1];                // "bar"

```

Map Keys

To get the list of keys, use `keys()`, which returns an iterator over the keys in the map:

```

var m = new Map();

var x = { id: 1 },
y = { id: 2 };

m.set( x, "foo" );
m.set( y, "bar" );

var keys = [ ...m.keys() ];

keys[0] === x;           // true
keys[1] === y;           // true

```

To determine if a map has a given key, use `has(..)`:

```

var m = new Map();

var x = { id: 1 },
y = { id: 2 };

m.set( x, "foo" );

m.has( x );           // true
m.has( y );           // false

```

Maps essentially let you associate some extra piece of information (the value) with an object (the key) without actually putting that information on the object itself.

While you can use any kind of value as a key for a map, you typically will use objects, as strings and other primitives are already eligible as keys of normal objects. In other words, you'll probably want to continue to use normal objects for maps unless some or all of the keys need to be objects, in which case map is more appropriate.

Warning: If you use an object as a map key and that object is later discarded

(all references unset) in attempt to have garbage collection (GC) reclaim its memory, the map itself will still retain its entry. You will need to remove the entry from the map for it to be GC-eligible. In the next section, we'll see WeakMaps as a better option for object keys and GC.

WeakMaps

WeakMaps are a variation on maps, which has most of the same external behavior but differs underneath in how the memory allocation (specifically its GC) works.

WeakMaps take (only) objects as keys. Those objects are held *weakly*, which means if the object itself is GC'd, the entry in the WeakMap is also removed. This isn't observable behavior, though, as the only way an object can be GC'd is if there's no more references to it – once there are no more references to it, you have no object reference to check if it exists in the WeakMap.

Otherwise, the API for WeakMap is similar, though more limited:

```
var m = new WeakMap();

var x = { id: 1 },
y = { id: 2 };

m.set( x, "foo" );

m.has( x );           // true
m.has( y );           // false
```

WeakMaps do not have a `size` property or `clear()` method, nor do they expose any iterators over their keys, values, or entries. So even if you unset the `x` reference, which will remove its entry from `m` upon GC, there is no way to tell. You'll just have to take JavaScript's word for it!

Just like Maps, WeakMaps let you soft-associate information with an object. But they are particularly useful if the object is not one you completely control, such as a DOM element. If the object you're using as a map key can be deleted and should be GC-eligible when it is, then a WeakMap is a more appropriate option.

It's important to note that a WeakMap only holds its *keys* weakly, not its values. Consider:

```
var m = new WeakMap();

var x = { id: 1 },
y = { id: 2 },
z = { id: 3 },
w = { id: 4 };
```

```

m.set( x, y );

x = null;           // { id: 1 } is GC-eligible
y = null;           // { id: 2 } is GC-eligible
                    // only because { id: 1 } is

m.set( z, w );

w = null;           // { id: 4 } is not GC-eligible

```

For this reason, WeakMaps are in my opinion better named “WeakKeyMaps.”

Sets

A set is a collection of unique values (duplicates are ignored).

The API for a set is similar to map. The `add(..)` method takes the place of the `set(..)` method (somewhat ironically), and there is no `get(..)` method.

Consider:

```

var s = new Set();

var x = { id: 1 },
y = { id: 2 };

s.add( x );
s.add( y );
s.add( x );

s.size;           // 2

s.delete( y );
s.size;           // 1

s.clear();
s.size;           // 0

```

The `Set(..)` constructor form is similar to `Map(..)`, in that it can receive an iterable, like another set or simply an array of values. However, unlike how `Map(..)` expects *entries* list (array of key/value arrays), `Set(..)` expects a *values* list (array of values):

```

var x = { id: 1 },
y = { id: 2 };

var s = new Set( [x,y] );

```

A set doesn't need a `get(..)` because you don't retrieve a value from a set, but rather test if it is present or not, using `has(..)`:

```
var s = new Set();

var x = { id: 1 },
y = { id: 2 };

s.add( x );

s.has( x );           // true
s.has( y );           // false
```

Note: The comparison algorithm in `has(..)` is almost identical to `Object.is(..)` (see Chapter 6), except that `-0` and `0` are treated as the same rather than distinct.

Set Iterators

Sets have the same iterator methods as maps. Their behavior is different for sets, but symmetric with the behavior of map iterators. Consider:

```
var s = new Set();

var x = { id: 1 },
y = { id: 2 };

s.add( x ).add( y );

var keys = [ ...s.keys() ],
vals = [ ...s.values() ],
entries = [ ...s.entries() ];

keys[0] === x;
keys[1] === y;

vals[0] === x;
vals[1] === y;

entries[0][0] === x;
entries[0][1] === x;
entries[1][0] === y;
entries[1][1] === y;
```

The `keys()` and `values()` iterators both yield a list of the unique values in the set. The `entries()` iterator yields a list of entry arrays, where both items of the array are the unique set value. The default iterator for a set is its `values()` iterator.

The inherent uniqueness of a set is its most useful trait. For example:

```
var s = new Set( [1,2,3,4,"1",2,4,"5"] ),
    uniques = [ ...s ];

uniques;                                // [1,2,3,4,"1", "5"]
```

Set uniqueness does not allow coercion, so 1 and "1" are considered distinct values.

WeakSets

Whereas a WeakMap holds its keys weakly (but its values strongly), a WeakSet holds its values weakly (there aren't really keys).

```
var s = new WeakSet();

var x = { id: 1 },
    y = { id: 2 };

s.add( x );
s.add( y );

x = null;                                // `x` is GC-eligible
y = null;                                // `y` is GC-eligible
```

Warning: WeakSet values must be objects, not primitive values as is allowed with sets.

Review

ES6 defines a number of useful collections that make working with data in structured ways more efficient and effective.

TypedArrays provide “view”s of binary data buffers that align with various integer types, like 8-bit unsigned integers and 32-bit floats. The array access to binary data makes operations much easier to express and maintain, which enables you to more easily work with complex data like video, audio, canvas data, and so on.

Maps are key-value pairs where the key can be an object instead of just a string/primitive. Sets are unique lists of values (of any type).

WeakMaps are maps where the key (object) is weakly held, so that GC is free to collect the entry if it's the last reference to an object. WeakSets are sets where the value is weakly held, again so that GC can remove the entry if it's the last reference to that object.

You Don't Know JS: ES6 & Beyond

Chapter 6: API Additions

From conversions of values to mathematic calculations, ES6 adds many static properties and methods to various built-in natives and objects to help with common tasks. In addition, instances of some of the natives have new capabilities via various new prototype methods.

Note: Most of these features can be faithfully polyfilled. We will not dive into such details here, but check out “ES6 Shim” (<https://github.com/paulmillr/es6-shim/>) for standards-compliant shims/polyfills.

Array

One of the most commonly extended features in JS by various user libraries is the `Array` type. It should be no surprise that ES6 adds a number of helpers to `Array`, both static and prototype (instance).

`Array.of(...)` Static Function

There's a well known gotcha with the `Array(...)` constructor, which is that if there's only one argument passed, and that argument is a number, instead of making an array of one element with that number value in it, it constructs an empty array with a `length` property equal to the number. This action produces the unfortunate and quirky “empty slots” behavior that's reviled about JS arrays.

`Array.of(...)` replaces `Array(...)` as the preferred function-form constructor for arrays, because `Array.of(...)` does not have that special single-number-argument case. Consider:

```
var a = Array( 3 );
a.length;           // 3
a[0];              // undefined

var b = Array.of( 3 );
b.length;           // 1
b[0];              // 3

var c = Array.of( 1, 2, 3 );
c.length;           // 3
c;                  // [1,2,3]
```

Under what circumstances would you want to use `Array.of(...)` instead of just creating an array with literal syntax, like `c = [1,2,3]`? There's two possible cases.

If you have a callback that's supposed to wrap argument(s) passed to it in an array, `Array.of(..)` fits the bill perfectly. That's probably not terribly common, but it may scratch an itch for you.

The other scenario is if you subclass `Array` (see “Classes” in Chapter 3) and want to be able to create and initialize elements in an instance of your subclass, such as:

```
class MyCoolArray extends Array {
  sum() {
    return this.reduce( function reducer(acc,curr){
      return acc + curr;
    }, 0 );
  }
}

var x = new MyCoolArray( 3 );
x.length;                                // 3 -- oops!
x.sum();                                   // 0 -- oops!

var y = [3];                                // Array, not MyCoolArray
y.length;                                  // 1
y.sum();                                    // `sum` is not a function

var z = MyCoolArray.of( 3 );
z.length;                                  // 1
z.sum();                                   // 3
```

You can't just (easily) create a constructor for `MyCoolArray` that overrides the behavior of the `Array` parent constructor, because that constructor is necessary to actually create a well-behaving array value (initializing the `this`). The “inherited” static `of(..)` method on the `MyCoolArray` subclass provides a nice solution.

`Array.from(..)` Static Function

An “array-like object” in JavaScript is an object that has a `length` property on it, specifically with an integer value of zero or higher.

These values have been notoriously frustrating to work with in JS; it's been quite common to need to transform them into an actual array, so that the various `Array.prototype` methods (`map(..)`, `indexOf(..)` etc.) are available to use with it. That process usually looks like:

```
// array-like object
var arrLike = {
  length: 3,
  0: "foo",
```

```

    1: "bar"
};

var arr = Array.prototype.slice.call( arrLike );

```

Another common task where `slice(..)` is often used is in duplicating a real array:

```
var arr2 = arr.slice();
```

In both cases, the new ES6 `Array.from(..)` method can be a more understandable and graceful – if also less verbose – approach:

```
var arr = Array.from( arrLike );
```

```
var arrCopy = Array.from( arr );
```

`Array.from(..)` looks to see if the first argument is an iterable (see “Iterators” in Chapter 3), and if so, it uses the iterator to produce values to “copy” into the returned array. Because real arrays have an iterator for those values, that iterator is automatically used.

But if you pass an array-like object as the first argument to `Array.from(..)`, it behaves basically the same as `slice()` (no arguments!) or `apply(..)` does, which is that it simply loops over the value, accessing numerically named properties from 0 up to whatever the value of `length` is.

Consider:

```

var arrLike = {
  length: 4,
  2: "foo"
};

Array.from( arrLike );
// [ undefined, undefined, "foo", undefined ]

```

Because positions 0, 1, and 3 didn’t exist on `arrLike`, the result was the `undefined` value for each of those slots.

You could produce a similar outcome like this:

```

var emptySlotsArr = [];
emptySlotsArr.length = 4;
emptySlotsArr[2] = "foo";

Array.from( emptySlotsArr );
// [ undefined, undefined, "foo", undefined ]

```

Avoiding Empty Slots There’s a subtle but important difference in the previous snippet between the `emptySlotsArr` and the result of the `Array.from(..)`

call. `Array.from(..)` never produces empty slots.

Prior to ES6, if you wanted to produce an array initialized to a certain length with actual `undefined` values in each slot (no empty slots!), you had to do extra work:

```
var a = Array( 4 );                                // four empty slots!

var b = Array.apply( null, { length: 4 } );        // four `undefined` values

But Array.from(..) now makes this easier:

var c = Array.from( { length: 4 } );                // four `undefined` values
```

Warning: Using an empty slot array like `a` in the previous snippets would work with some array functions, but others ignore empty slots (like `map(..)`, etc.). You should never intentionally work with empty slots, as it will almost certainly lead to strange/unpredictable behavior in your programs.

Mapping The `Array.from(..)` utility has another helpful trick up its sleeve. The second argument, if provided, is a mapping callback (almost the same as the regular `Array#map(..)` expects) which is called to map/transform each value from the source to the returned target. Consider:

```
var arrLike = {
  length: 4,
  2: "foo"
};

Array.from( arrLike, function mapper(val, idx){
  if (typeof val == "string") {
    return val.toUpperCase();
  }
  else {
    return idx;
  }
} );
// [ 0, 1, "FOO", 3 ]
```

Note: As with other array methods that take callbacks, `Array.from(..)` takes an optional third argument that if set will specify the `this` binding for the callback passed as the second argument. Otherwise, `this` will be `undefined`.

See “[TypedArrays](#)” in Chapter 5 for an example of using `Array.from(..)` in translating values from an array of 8-bit values to an array of 16-bit values.

Creating Arrays and Subtypes

In the last couple of sections, we’ve discussed `Array.of(..)` and `Array.from(..)`, both of which create a new array in a similar way to a

constructor. But what do they do in subclasses? Do they create instances of the base `Array` or the derived subclass?

```
class MyCoolArray extends Array {  
  ..  
}  
  
MyCoolArray.from( [1, 2] ) instanceof MyCoolArray; // true  
  
Array.from(  
  MyCoolArray.from( [1, 2] )  
) instanceof MyCoolArray; // false
```

Both `of(..)` and `from(..)` use the constructor that they're accessed from to construct the array. So if you use the base `Array.of(..)` you'll get an `Array` instance, but if you use `MyCoolArray.of(..)`, you'll get a `MyCoolArray` instance.

In “Classes” in Chapter 3, we covered the `@@species` setting which all the built-in classes (like `Array`) have defined, which is used by any prototype methods if they create a new instance. `slice(..)` is a great example:

```
var x = new MyCoolArray( 1, 2, 3 );  
  
x.slice( 1 ) instanceof MyCoolArray; // true
```

Generally, that default behavior will probably be desired, but as we discussed in Chapter 3, you *can* override if you want:

```
class MyCoolArray extends Array {  
  // force `species` to be parent constructor  
  static get [Symbol.species]() { return Array; }  
}  
  
var x = new MyCoolArray( 1, 2, 3 );  
  
x.slice( 1 ) instanceof MyCoolArray; // false  
x.slice( 1 ) instanceof Array; // true
```

It's important to note that the `@@species` setting is only used for the prototype methods, like `slice(..)`. It's not used by `of(..)` and `from(..)`; they both just use the `this` binding (whatever constructor is used to make the reference). Consider:

```
class MyCoolArray extends Array {  
  // force `species` to be parent constructor  
  static get [Symbol.species]() { return Array; }  
}  
  
var x = new MyCoolArray( 1, 2, 3 );
```

```
MyCoolArray.from( x ) instanceof MyCoolArray;      // true
MyCoolArray.of( [2, 3] ) instanceof MyCoolArray;    // true
```

copyWithin(..) Prototype Method

`Array#copyWithin(..)` is a new mutator method available to all arrays (including Typed Arrays; see Chapter 5). `copyWithin(..)` copies a portion of an array to another location in the same array, overwriting whatever was there before.

The arguments are *target* (the index to copy to), *start* (the inclusive index to start the copying from), and optionally *end* (the exclusive index to stop copying). If any of the arguments are negative, they're taken to be relative from the end of the array.

Consider:

```
[1,2,3,4,5].copyWithin( 3, 0 );           // [1,2,3,1,2]
[1,2,3,4,5].copyWithin( 3, 0, 1 );        // [1,2,3,1,5]
[1,2,3,4,5].copyWithin( 0, -2 );          // [4,5,3,4,5]
[1,2,3,4,5].copyWithin( 0, -2, -1 );      // [4,2,3,4,5]
```

The `copyWithin(..)` method does not extend the array's length, as the first example in the previous snippet shows. Copying simply stops when the end of the array is reached.

Contrary to what you might think, the copying doesn't always go in left-to-right (ascending index) order. It's possible this would result in repeatedly copying an already copied value if the from and target ranges overlap, which is presumably not desired behavior.

So internally, the algorithm avoids this case by copying in reverse order to avoid that gotcha. Consider:

```
[1,2,3,4,5].copyWithin( 2, 1 );      // ???
```

If the algorithm was strictly moving left to right, then the 2 should be copied to overwrite the 3, then *that* copied 2 should be copied to overwrite 4, then *that* copied 2 should be copied to overwrite 5, and you'd end up with `[1,2,2,2,2]`.

Instead, the copying algorithm reverses direction and copies 4 to overwrite 5, then copies 3 to overwrite 4, then copies 2 to overwrite 3, and the final result is `[1,2,2,3,4]`. That's probably more "correct" in terms of expectation, but it can be confusing if you're only thinking about the copying algorithm in a naive left-to-right fashion.

fill(..) Prototype Method

Filling an existing array entirely (or partially) with a specified value is natively supported as of ES6 with the `Array#fill(..)` method:

```
var a = Array( 4 ).fill( undefined );
a;
// [undefined,undefined,undefined,undefined]
```

`fill(..)` optionally takes *start* and *end* parameters, which indicate a subset portion of the array to fill, such as:

```
var a = [ null, null, null, null ].fill( 42, 1, 3 );
a;
// [null,42,42,null]
```

find(..) Prototype Method

The most common way to search for a value in an array has generally been the `indexOf(..)` method, which returns the index the value is found at or `-1` if not found:

```
var a = [1,2,3,4,5];

(a.indexOf( 3 ) != -1);           // true
(a.indexOf( 7 ) != -1);           // false

(a.indexOf( "2" ) != -1);         // false
```

The `indexOf(..)` comparison requires a strict `==` match, so a search for `"2"` fails to find a value of `2`, and vice versa. There's no way to override the matching algorithm for `indexOf(..)`. It's also unfortunate/ungraceful to have to make the manual comparison to the `-1` value.

Tip: See the *Types & Grammar* title of this series for an interesting (and controversially confusing) technique to work around the `-1` ugliness with the `~` operator.

Since ES5, the most common workaround to have control over the matching logic has been the `some(..)` method. It works by calling a function callback for each element, until one of those calls returns a `true`/truthy value, and then it stops. Because you get to define the callback function, you have full control over how a match is made:

```
var a = [1,2,3,4,5];

a.some( function matcher(v){
    return v == "2";
} );                         // true
```

```
a.some( function matcher(v){
  return v == 7;
} ); // false
```

But the downside to this approach is that you only get the `true/false` indicating if a suitably matched value was found, but not what the actual matched value was.

ES6's `find(..)` addresses this. It works basically the same as `some(..)`, except that once the callback returns a `true/truthy` value, the actual array value is returned:

```
var a = [1,2,3,4,5];

a.find( function matcher(v){
  return v == "2";
} ); // 2

a.find( function matcher(v){
  return v == 7; // undefined
} );
```

Using a custom `matcher(..)` function also lets you match against complex values like objects:

```
var points = [
  { x: 10, y: 20 },
  { x: 20, y: 30 },
  { x: 30, y: 40 },
  { x: 40, y: 50 },
  { x: 50, y: 60 }
];

points.find( function matcher(point) {
  return (
    point.x % 3 == 0 &&
    point.y % 4 == 0
  );
} ); // { x: 30, y: 40 }
```

Note: As with other array methods that take callbacks, `find(..)` takes an optional second argument that if set will specify the `this` binding for the callback passed as the first argument. Otherwise, `this` will be `undefined`.

`findIndex(..)` Prototype Method

While the previous section illustrates how `some(..)` yields a boolean result for a search of an array, and `find(..)` yields the matched value itself from the

array search, there's also a need for finding the positional index of the matched value.

`indexOf(..)` does that, but there's no control over its matching logic; it always uses `==` strict equality. So ES6's `findIndex(..)` is the answer:

```
var points = [
  { x: 10, y: 20 },
  { x: 20, y: 30 },
  { x: 30, y: 40 },
  { x: 40, y: 50 },
  { x: 50, y: 60 }
];

points.findIndex( function matcher(point) {
  return (
    point.x % 3 == 0 &&
    point.y % 4 == 0
  );
} ); // 2

points.findIndex( function matcher(point) {
  return (
    point.x % 6 == 0 &&
    point.y % 7 == 0
  );
} ); // -1
```

Don't use `findIndex(..) != -1` (the way it's always been done with `indexOf(..)`) to get a boolean from the search, because `some(..)` already yields the `true/false` you want. And don't do `a[a.findIndex(..)]` to get the matched value, because that's what `find(..)` accomplishes. And finally, use `indexOf(..)` if you need the index of a strict match, or `findIndex(..)` if you need the index of a more customized match.

Note: As with other array methods that take callbacks, `findIndex(..)` takes an optional second argument that if set will specify the `this` binding for the callback passed as the first argument. Otherwise, `this` will be `undefined`.

`entries()`, `values()`, `keys()` Prototype Methods

In Chapter 3, we illustrated how data structures can provide a patterned item-by-item enumeration of their values, via an iterator. We then expounded on this approach in Chapter 5, as we explored how the new ES6 collections (Map, Set, etc.) provide several methods for producing different kinds of iterations.

Because it's not new to ES6, `Array` might not be thought of traditionally as a "collection," but it is one in the sense that it provides these same iterator

methods: `entries()`, `values()`, and `keys()`. Consider:

```
var a = [1,2,3];

[...a.values()];           // [1,2,3]
[...a.keys()];            // [0,1,2]
[...a.entries()];          // [ [0,1], [1,2], [2,3] ]

[...a[Symbol.iterator]()] ; // [1,2,3]
```

Just like with `Set`, the default `Array` iterator is the same as what `values()` returns.

In “Avoiding Empty Slots” earlier in this chapter, we illustrated how `Array.from(..)` treats empty slots in an array as just being present slots with `undefined` in them. That’s actually because under the covers, the array iterators behave that way:

```
var a = [];
a.length = 3;
a[1] = 2;

[...a.values()];           // [undefined,2,undefined]
[...a.keys()];            // [0,1,2]
[...a.entries()];          // [ [0,undefined], [1,2], [2,undefined] ]
```

Object

A few additional static helpers have been added to `Object`. Traditionally, functions of this sort have been seen as focused on the behaviors/capabilities of object values.

However, starting with ES6, `Object` static functions will also be for general-purpose global APIs of any sort that don’t already belong more naturally in some other location (i.e., `Array.from(..)`).

`Object.is(..)` Static Function

The `Object.is(..)` static function makes value comparisons in an even more strict fashion than the `==` comparison.

`Object.is(..)` invokes the underlying `SameValue` algorithm (ES6 spec, section 7.2.9). The `SameValue` algorithm is basically the same as the `==` Strict Equality Comparison Algorithm (ES6 spec, section 7.2.13), with two important exceptions.

Consider:

```
var x = NaN, y = 0, z = -0;
```

```

x === x;                      // false
y === z;                      // true

Object.is( x, x );            // true
Object.is( y, z );            // false

```

You should continue to use `==` for strict equality comparisons; `Object.is(..)` shouldn't be thought of as a replacement for the operator. However, in cases where you're trying to strictly identify a `NaN` or `-0` value, `Object.is(..)` is now the preferred option.

Note: ES6 also adds a `Number.isNaN(..)` utility (discussed later in this chapter) which may be a slightly more convenient test; you may prefer `Number.isNaN(x)` over `Object.is(x,NaN)`. You *can* accurately test for `-0` with a clumsy `x == 0 && 1 / x === -Infinity`, but in this case `Object.is(x,-0)` is much better.

`Object.getOwnPropertySymbols(..)` Static Function

The “Symbols” section in Chapter 2 discusses the new `Symbol` primitive value type in ES6.

Symbols are likely going to be mostly used as special (meta) properties on objects. So the `Object.getOwnPropertySymbols(..)` utility was introduced, which retrieves only the symbol properties directly on an object:

```

var o = {
  foo: 42,
  [ Symbol( "bar" ) ]: "hello world",
  baz: true
};

Object.getOwnPropertySymbols( o ); // [ Symbol(bar) ]

```

`Object.setPrototypeOf(..)` Static Function

Also in Chapter 2, we mentioned the `Object.setPrototypeOf(..)` utility, which (unsurprisingly) sets the `[[Prototype]]` of an object for the purposes of *behavior delegation* (see the *this & Object Prototypes* title of this series). Consider:

```

var o1 = {
  foo() { console.log( "foo" ); }
};

var o2 = {
  // .. o2's definition ..
};

Object.setPrototypeOf( o2, o1 );

```

```
// delegates to `o1.foo()`  
o2.foo(); // foo
```

Alternatively:

```
var o1 = {  
  foo() { console.log( "foo" ); }  
};  
  
var o2 = Object.setPrototypeOf( {  
  // .. o2's definition ..  
}, o1 );  
  
// delegates to `o1.foo()`  
o2.foo(); // foo
```

In both previous snippets, the relationship between `o2` and `o1` appears at the end of the `o2` definition. More commonly, the relationship between an `o2` and `o1` is specified at the top of the `o2` definition, as it is with classes, and also with `__proto__` in object literals (see “Setting `[[Prototype]]`” in Chapter 2).

Warning: Setting a `[[Prototype]]` right after object creation is reasonable, as shown. But changing it much later is generally not a good idea and will usually lead to more confusion than clarity.

`Object.assign(..)` Static Function

Many JavaScript libraries/frameworks provide utilities for copying/mixing one object’s properties into another (e.g., jQuery’s `extend(..)`). There are various nuanced differences between these different utilities, such as whether a property with value `undefined` is ignored or not.

ES6 adds `Object.assign(..)`, which is a simplified version of these algorithms. The first argument is the *target*, and any other arguments passed are the *sources*, which will be processed in listed order. For each source, its enumerable and own (e.g., not “inherited”) keys, including symbols, are copied as if by plain `=` assignment. `Object.assign(..)` returns the target object.

Consider this object setup:

```
var target = {},  
  o1 = { a: 1 }, o2 = { b: 2 },  
  o3 = { c: 3 }, o4 = { d: 4 };  
  
// setup read-only property  
Object.defineProperty( o3, "e", {  
  value: 5,  
  enumerable: true,
```

```

        writable: false,
        configurable: false
    } );

    // setup non-enumerable property
Object.defineProperty( o3, "f", {
    value: 6,
    enumerable: false
} );

o3[ Symbol( "g" ) ] = 7;

    // setup non-enumerable symbol
Object.defineProperty( o3, Symbol( "h" ), {
    value: 8,
    enumerable: false
} );

Object.setPrototypeOf( o3, o4 );

```

Only the properties `a`, `b`, `c`, `e`, and `Symbol("g")` will be copied to `target`:

```

Object.assign( target, o1, o2, o3 );

target.a;                      // 1
target.b;                      // 2
target.c;                      // 3

Object.getOwnPropertyDescriptor( target, "e" );
// { value: 5, writable: true, enumerable: true,
//   configurable: true }

Object.getOwnPropertySymbols( target );
// [Symbol("g")]

```

The `d`, `f`, and `Symbol("h")` properties are omitted from copying; non-enumerable properties and non-owned properties are all excluded from the assignment. Also, `e` is copied as a normal property assignment, not duplicated as a read-only property.

In an earlier section, we showed using `setPrototypeOf(..)` to set up a `[[Prototype]]` relationship between an `o2` and `o1` object. There's another form that leverages `Object.assign(..)`:

```

var o1 = {
    foo() { console.log( "foo" ); }
};

```

```

var o2 = Object.assign(
  Object.create( o1 ),
  {
    // ... o2's definition ...
  }
);

// delegates to `o1.foo()`
o2.foo(); // foo

```

Note: `Object.create(..)` is the ES5 standard utility that creates an empty object that is `[[Prototype]]`-linked. See the *this & Object Prototypes* title of this series for more information.

Math

ES6 adds several new mathematic utilities that fill in holes or aid with common operations. All of these can be manually calculated, but most of them are now defined natively so that in some cases the JS engine can either more optimally perform the calculations, or perform them with better decimal precision than their manual counterparts.

It's likely that asm.js/transpiled JS code (see the *Async & Performance* title of this series) is the more likely consumer of many of these utilities rather than direct developers.

Trigonometry:

- `cosh(..)` - Hyperbolic cosine
- `acosh(..)` - Hyperbolic arccosine
- `sinh(..)` - Hyperbolic sine
- `asinh(..)` - Hyperbolic arcsine
- `tanh(..)` - Hyperbolic tangent
- `atanh(..)` - Hyperbolic arctangent
- `hypot(..)` - The squareroot of the sum of the squares (i.e., the generalized Pythagorean theorem)

Arithmetic:

- `cbrt(..)` - Cube root
- `clz32(..)` - Count leading zeros in 32-bit binary representation
- `expm1(..)` - The same as `exp(x) - 1`
- `log2(..)` - Binary logarithm (log base 2)
- `log10(..)` - Log base 10
- `log1p(..)` - The same as `log(x + 1)`
- `imul(..)` - 32-bit integer multiplication of two numbers

Meta:

- `sign(..)` - Returns the sign of the number

- `trunc(..)` - Returns only the integer part of a number
- `fround(..)` - Rounds to nearest 32-bit (single precision) floating-point value

Number

Importantly, for your program to properly work, it must accurately handle numbers. ES6 adds some additional properties and functions to assist with common numeric operations.

Two additions to `Number` are just references to the preexisting globals: `Number.parseInt(..)` and `Number.parseFloat(..)`.

Static Properties

ES6 adds some helpful numeric constants as static properties:

- `Number.EPSILON` - The minimum value between any two numbers: 2^{-52} (see Chapter 2 of the *Types & Grammar* title of this series regarding using this value as a tolerance for imprecision in floating-point arithmetic)
- `Number.MAX_SAFE_INTEGER` - The highest integer that can “safely” be represented unambiguously in a JS number value: $2^{53} - 1$
- `Number.MIN_SAFE_INTEGER` - The lowest integer that can “safely” be represented unambiguously in a JS number value: $-(2^{53} - 1)$ or $(-2)^{53} + 1$.

Note: See Chapter 2 of the *Types & Grammar* title of this series for more information about “safe” integers.

Number.isNaN(..) Static Function

The standard global `isNaN(..)` utility has been broken since its inception, in that it returns `true` for things that are not numbers, not just for the actual `NaN` value, because it coerces the argument to a number type (which can falsely result in a `NaN`). ES6 adds a fixed utility `Number.isNaN(..)` that works as it should:

```
var a = NaN, b = "NaN", c = 42;

isNaN( a );                      // true
isNaN( b );                      // true -- oops!
isNaN( c );                      // false

Number.isNaN( a );                // true
Number.isNaN( b );                // false -- fixed!
Number.isNaN( c );                // false
```

Number.isFinite(..) Static Function

There's a temptation to look at a function name like `isFinite(..)` and assume it's simply "not infinite". That's not quite correct, though. There's more nuance to this new ES6 utility. Consider:

```
var a = NaN, b = Infinity, c = 42;

Number.isFinite( a );           // false
Number.isFinite( b );           // false

Number.isFinite( c );           // true
```

The standard global `isFinite(..)` coerces its argument, but `Number.isFinite(..)` omits the coercive behavior:

```
var a = "42";

isFinite( a );                 // true
Number.isFinite( a );           // false
```

You may still prefer the coercion, in which case using the global `isFinite(..)` is a valid choice. Alternatively, and perhaps more sensibly, you can use `Number.isFinite(+x)`, which explicitly coerces `x` to a number before passing it in (see Chapter 4 of the *Types & Grammar* title of this series).

Integer-Related Static Functions

JavaScript number values are always floating point (IEEE-754). So the notion of determining if a number is an "integer" is not about checking its type, because JS makes no such distinction.

Instead, you need to check if there's any non-zero decimal portion of the value. The easiest way to do that has commonly been:

```
x === Math.floor( x );
```

ES6 adds a `Number.isInteger(..)` helper utility that potentially can determine this quality slightly more efficiently:

```
Number.isInteger( 4 );           // true
Number.isInteger( 4.2 );          // false
```

Note: In JavaScript, there's no difference between 4, 4., 4.0, or 4.0000. All of these would be considered an "integer", and would thus yield `true` from `Number.isInteger(..)`.

In addition, `Number.isInteger(..)` filters out some clearly not-integer values that `x === Math.floor(x)` could potentially mix up:

```
Number.isInteger( NaN );          // false
Number.isInteger( Infinity );      // false
```

Working with “integers” is sometimes an important bit of information, as it can simplify certain kinds of algorithms. JS code by itself will not run faster just from filtering for only integers, but there are optimization techniques the engine can take (e.g., `asm.js`) when only integers are being used.

Because of `Number.isInteger(..)`’s handling of `NaN` and `Infinity` values, defining a `isFloat(..)` utility would not be just as simple as `!Number.isInteger(..)`. You’d need to do something like:

```
function isFloat(x) {
  return Number.isFinite( x ) && !Number.isInteger( x );
}

isFloat( 4.2 );           // true
isFloat( 4 );            // false

isFloat( NaN );          // false
isFloat( Infinity );    // false
```

Note: It may seem strange, but `Infinity` should neither be considered an integer nor a float.

ES6 also defines a `Number.isSafeInteger(..)` utility, which checks to make sure the value is both an integer and within the range of `Number.MIN_SAFE_INTEGER`-`Number.MAX_SAFE_INTEGER` (inclusive).

```
var x = Math.pow( 2, 53 ),
    y = Math.pow( -2, 53 );

Number.isSafeInteger( x - 1 );    // true
Number.isSafeInteger( y + 1 );    // true

Number.isSafeInteger( x );        // false
Number.isSafeInteger( y );        // false
```

String

Strings already have quite a few helpers prior to ES6, but even more have been added to the mix.

Unicode Functions

“Unicode-Aware String Operations” in Chapter 2 discusses `String.fromCodePoint(..)`, `String#codePointAt(..)`, and `String#normalize(..)` in detail. They have been added to improve Unicode support in JS string values.

```
String.fromCodePoint( 0x1d49e );           // " "
"ab d".codePointAt( 2 ).toString( 16 );    // "1d49e"
```

The `normalize(..)` string prototype method is used to perform Unicode normalizations that either combine characters with adjacent “combining marks” or decompose combined characters.

Generally, the normalization won’t create a visible effect on the contents of the string, but will change the contents of the string, which can affect how things like the `length` property are reported, as well as how character access by position behave:

```
var s1 = "e\u0301";
s1.length;                                // 2

var s2 = s1.normalize();
s2.length;                                // 1
s2 === "\xE9";                           // true
```

`normalize(..)` takes an optional argument that specifies the normalization form to use. This argument must be one of the following four values: "NFC" (default), "NFD", "NFKC", or "NFKD".

Note: Normalization forms and their effects on strings is well beyond the scope of what we’ll discuss here. See “Unicode Normalization Forms” (<http://www.unicode.org/reports/tr15/>) for more information.

`String.raw(..)` Static Function

The `String.raw(..)` utility is provided as a built-in tag function to use with template string literals (see Chapter 2) for obtaining the raw string value without any processing of escape sequences.

This function will almost never be called manually, but will be used with tagged template literals:

```
var str = "bc";

String.raw`\ta${str}d\xE9`;
// "\abcd\xE9", not " abcdé"
```

In the resultant string, `\` and `t` are separate raw characters, not the one escape sequence character `\t`. The same is true with the Unicode escape sequence.

`repeat(..)` Prototype Function

In languages like Python and Ruby, you can repeat a string as:

```
"foo" * 3;                                // "foofoofoo"
```

That doesn’t work in JS, because `*` multiplication is only defined for numbers, and thus `"foo"` coerces to the `NaN` number.

However, ES6 defines a string prototype method `repeat(..)` to accomplish the task:

```
"foo".repeat( 3 );           // "foofoofoo"
```

String Inspection Functions

In addition to `String#indexOf(..)` and `String#lastIndexOf(..)` from prior to ES6, three new methods for searching/inspection have been added: `startsWith(..)`, `endsWith(..)`, and `includes(..)`.

```
var palindrome = "step on no pets";  
  
palindrome.startsWith( "step on" ); // true  
palindrome.startsWith( "on", 5 ); // true  
  
palindrome.endsWith( "no pets" ); // true  
palindrome.endsWith( "no", 10 ); // true  
  
palindrome.includes( "on" ); // true  
palindrome.includes( "on", 6 ); // false
```

For all the string search/inspection methods, if you look for an empty string "", it will either be found at the beginning or the end of the string.

Warning: These methods will not by default accept a regular expression for the search string. See “Regular Expression Symbols” in Chapter 7 for information about disabling the `isRegExp` check that is performed on this first argument.

Review

ES6 adds many extra API helpers on the various built-in native objects:

- `Array` adds `of(..)` and `from(..)` static functions, as well as prototype functions like `copyWithin(..)` and `fill(..)`.
- `Object` adds static functions like `is(..)` and `assign(..)`.
- `Math` adds static functions like `acosh(..)` and `clz32(..)`.
- `Number` adds static properties like `Number.EPSILON`, as well as static functions like `Number.isFinite(..)`.
- `String` adds static functions like `String.fromCodePoint(..)` and `String.raw(..)`, as well as prototype functions like `repeat(..)` and `includes(..)`.

Most of these additions can be polyfilled (see ES6 Shim), and were inspired by utilities in common JS libraries/frameworks.

You Don't Know JS: ES6 & Beyond

Chapter 7: Meta Programming

Meta programming is programming where the operation targets the behavior of the program itself. In other words, it's programming the programming of your program. Yeah, a mouthful, huh?

For example, if you probe the relationship between one object `a` and another `b` – are they `[[Prototype]]` linked? – using `a.isPrototypeOf(b)`, this is commonly referred to as introspection, a form of meta programming. Macros (which don't exist in JS, yet) – where the code modifies itself at compile time – are another obvious example of meta programming. Enumerating the keys of an object with a `for..in` loop, or checking if an object is an *instance of* a “class constructor”, are other common meta programming tasks.

Meta programming focuses on one or more of the following: code inspecting itself, code modifying itself, or code modifying default language behavior so other code is affected.

The goal of meta programming is to leverage the language's own intrinsic capabilities to make the rest of your code more descriptive, expressive, and/or flexible. Because of the *meta* nature of meta programming, it's somewhat difficult to put a more precise definition on it than that. The best way to understand meta programming is to see it through examples.

ES6 adds several new forms/features for meta programming on top of what JS already had.

Function Names

There are cases where your code may want to introspect on itself and ask what the name of some function is. If you ask what a function's name is, the answer is surprisingly somewhat ambiguous. Consider:

```
function daz() {
  // ...
}

var obj = {
  foo: function() {
    // ...
  },
  bar: function baz() {
    // ...
  },
  bam: daz,
  zim() {
```

```
// ...
}
};
```

In this previous snippet, “what is the name of `obj.foo()`” is slightly nuanced. Is it “`foo`”, “`undefined`”, or `undefined`? And what about `obj.bar()` – is it named “`bar`” or “`baz`”? Is `obj.bam()` named “`bam`” or “`daz`”? What about `obj.zim()`?

Moreover, what about functions which are passed as callbacks, like:

```
function foo(cb) {
  // what is the name of `cb()` here?
}

foo( function(){
  // I'm anonymous!
} );
```

There are quite a few ways that functions can be expressed in programs, and it’s not always clear and unambiguous what the “name” of that function should be.

More importantly, we need to distinguish whether the “name” of a function refers to its `name` property – yes, functions have a property called `name` – or whether it refers to the lexical binding name, such as `bar` in `function bar() { ... }`.

The lexical binding name is what you use for things like recursion:

```
function foo(i) {
  if (i < 10) return foo( i * 2 );
  return i;
}
```

The `name` property is what you’d use for meta programming purposes, so that’s what we’ll focus on in this discussion.

The confusion comes because by default, the lexical name a function has (if any) is also set as its `name` property. Actually there was no official requirement for that behavior by the ES5 (and prior) specifications. The setting of the `name` property was nonstandard but still fairly reliable. As of ES6, it has been standardized.

Tip: If a function has a `name` value assigned, that’s typically the name used in stack traces in developer tools.

Inferences

But what happens to the `name` property if a function has no lexical name?

As of ES6, there are now inference rules which can determine a sensible `name` property value to assign a function even if that function doesn't have a lexical name to use.

Consider:

```
var abc = function() {  
  // ..  
};  
  
abc.name; // "abc"
```

Had we given the function a lexical name like `abc = function def() { .. }`, the `name` property would of course be "`def`". But in the absence of the lexical name, intuitively the "`abc`" name seems appropriate.

Here are other forms that will infer a name (or not) in ES6:

```
(function(){ .. }) ; // name:  
(function*(){ .. }) ; // name:  
window.foo = function(){ .. } ; // name:  
  
class Awesome {  
  constructor() { .. } // name: Awesome  
  funny() { .. } // name: funny  
}  
  
var c = class Awesome { .. } ; // name: Awesome  
  
var o = {  
  foo() { .. } , // name: foo  
  *bar() { .. } , // name: bar  
  baz: () => { .. } , // name: baz  
  bam: function(){ .. } , // name: bam  
  get qux() { .. } , // name: get qux  
  set fuz() { .. } , // name: set fuz  
  ["b" + "iz"] :  
    function(){ .. } , // name: biz  
  [Symbol( "buz" )] :  
    function(){ .. } // name: [buz]  
};  
  
var x = o.foo.bind( o ) ; // name: bound foo  
(function(){ .. }).bind( o ) ; // name: bound  
  
export default function() { .. } // name: default  
  
var y = new Function() ; // name: anonymous
```

```

var GeneratorFunction =
  function*(){}.__proto__.constructor;
var z = new GeneratorFunction(); // name: anonymous

```

The `name` property is not writable by default, but it is configurable, meaning you can use `Object.defineProperty(..)` to manually change it if so desired.

Meta Properties

In the “`new.target`” section of Chapter 3, we introduced a concept new to JS in ES6: the meta property. As the name suggests, meta properties are intended to provide special meta information in the form of a property access that would otherwise not have been possible.

In the case of `new.target`, the keyword `new` serves as the context for a property access. Clearly `new` is itself not an object, which makes this capability special. However, when `new.target` is used inside a constructor call (a function/method invoked with `new`), `new` becomes a virtual context, so that `new.target` can refer to the target constructor that `new` invoked.

This is a clear example of a meta programming operation, as the intent is to determine from inside a constructor call what the original `new` target was, generally for the purposes of introspection (examining typing/structure) or static property access.

For example, you may want to have different behavior in a constructor depending on if it’s directly invoked or invoked via a child class:

```

class Parent {
  constructor() {
    if (new.target === Parent) {
      console.log( "Parent instantiated" );
    }
    else {
      console.log( "A child instantiated" );
    }
  }
}

class Child extends Parent {}

var a = new Parent();
// Parent instantiated

var b = new Child();
// A child instantiated

```

There’s a slight nuance here, which is that the `constructor()` inside the `Parent` class definition is actually given the lexical name of the class (`Parent`), even

though the syntax implies that the class is a separate entity from the constructor.

Warning: As with all meta programming techniques, be careful of creating code that's too clever for your future self or others maintaining your code to understand. Use these tricks with caution.

Well Known Symbols

In the “Symbols” section of Chapter 2, we covered the new ES6 primitive type `symbol`. In addition to symbols you can define in your own program, JS predefines a number of built-in symbols, referred to as *Well Known Symbols* (WKS).

These symbol values are defined primarily to expose special meta properties that are being exposed to your JS programs to give you more control over JS's behavior.

We'll briefly introduce each and discuss their purpose.

`Symbol.iterator`

In Chapters 2 and 3, we introduced and used the `@@iterator` symbol, automatically used by `...` spreads and `for..of` loops. We also saw `@@iterator` as defined on the new ES6 collections as defined in Chapter 5.

`Symbol.iterator` represents the special location (property) on any object where the language mechanisms automatically look to find a method that will construct an iterator instance for consuming that object's values. Many objects come with a default one defined.

However, we can define our own iterator logic for any object value by setting the `Symbol.iterator` property, even if that's overriding the default iterator. The meta programming aspect is that we are defining behavior which other parts of JS (namely, operators and looping constructs) use when processing an object value we define.

Consider:

```
var arr = [4,5,6,7,8,9];

for (var v of arr) {
  console.log( v );
}

// 4 5 6 7 8 9

// define iterator that only produces values
// from odd indexes
arr[Symbol.iterator] = function*() {
  var idx = 1;
  do {
```

```

        yield this[idx];
    } while ((idx += 2) < this.length);
};

for (var v of arr) {
    console.log( v );
}
// 5 7 9

```

Symbol.toStringTag and Symbol.hasInstance

One of the most common meta programming tasks is to introspect on a value to find out what *kind* it is, usually to decide what operations are appropriate to perform on it. With objects, the two most common inspection techniques are `toString()` and `instanceof`.

Consider:

```

function Foo() {}

var a = new Foo();

a.toString();           // [object Object]
a instanceof Foo;      // true

```

As of ES6, you can control the behavior of these operations:

```

function Foo(greeting) {
    this.greeting = greeting;
}

Foo.prototype[Symbol.toStringTag] = "Foo";

Object.defineProperty( Foo, Symbol.hasInstance, {
    value: function(inst) {
        return inst.greeting == "hello";
    }
} );

var a = new Foo( "hello" ),
    b = new Foo( "world" );

b[Symbol.toStringTag] = "cool";

a.toString();           // [object Foo]
String( b );           // [object cool]

a instanceof Foo;      // true

```

```
b instanceof Foo; // false
```

The `@@toStringTag` symbol on the prototype (or instance itself) specifies a string value to use in the `[object ___]` stringification.

The `@@hasInstance` symbol is a method on the constructor function which receives the instance object value and lets you decide by returning `true` or `false` if the value should be considered an instance or not.

Note: To set `@@hasInstance` on a function, you must use `Object.defineProperty(..)`, as the default one on `Function.prototype` is `writable: false`. See the *this & Object Prototypes* title of this series for more information.

Symbol.species

In “Classes” in Chapter 3, we introduced the `@@species` symbol, which controls which constructor is used by built-in methods of a class that needs to spawn new instances.

The most common example is when subclassing `Array` and wanting to define which constructor (`Array(..)` or your subclass) inherited methods like `slice(..)` should use. By default, `slice(..)` called on an instance of a subclass of `Array` would produce a new instance of that subclass, which is frankly what you’ll likely often want.

However, you can meta program by overriding a class’s default `@@species` definition:

```
class Cool {
  // defer `@@species` to derived constructor
  static get [Symbol.species]() { return this; }

  again() {
    return new this.constructor[Symbol.species]();
  }
}

class Fun extends Cool {}

class Awesome extends Cool {
  // force `@@species` to be parent constructor
  static get [Symbol.species]() { return Cool; }
}

var a = new Fun(),
  b = new Awesome(),
  c = a.again(),
  d = b.again();
```

```

c instanceof Fun;           // true
d instanceof Awesome;      // false
d instanceof Cool;         // true

```

The `Symbol.species` setting defaults on the built-in native constructors to the `return this` behavior as illustrated in the previous snippet in the `Cool` definition. It has no default on user classes, but as shown that behavior is easy to emulate.

If you need to define methods that generate new instances, use the meta programming of the `new this.constructor[Symbol.species](..)` pattern instead of the hard-wiring of `new this.constructor(..)` or `new XYZ(..)`. Derived classes will then be able to customize `Symbol.species` to control which constructor vends those instances.

`Symbol.toPrimitive`

In the *Types & Grammar* title of this series, we discussed the `ToPrimitive` abstract coercion operation, which is used when an object must be coerced to a primitive value for some operation (such as `==` comparison or `+` addition). Prior to ES6, there was no way to control this behavior.

As of ES6, the `@@toPrimitive` symbol as a property on any object value can customize that `ToPrimitive` coercion by specifying a method.

Consider:

```

var arr = [1,2,3,4,5];

arr + 10;           // 1,2,3,4,510

arr[Symbol.toPrimitive] = function(hint) {
  if (hint == "default" || hint == "number") {
    // sum all numbers
    return this.reduce( function(acc,curr){
      return acc + curr;
    }, 0 );
  }
};

arr + 10;           // 25

```

The `Symbol.toPrimitive` method will be provided with a *hint* of `"string"`, `"number"`, or `"default"` (which should be interpreted as `"number"`), depending on what type the operation invoking `ToPrimitive` is expecting. In the previous snippet, the additive `+` operation has no hint (`"default"` is passed). A multiplicative `*` operation would hint `"number"` and a `String(arr)` would hint `"string"`.

Warning: The `==` operator will invoke the `ToPrimitive` operation with no hint – the `@@toPrimitive` method, if any is called with hint "default" – on an object if the other value being compared is not an object. However, if both comparison values are objects, the behavior of `==` is identical to `==`, which is that the references themselves are directly compared. In this case, `@@toPrimitive` is not invoked at all. See the *Types & Grammar* title of this series for more information about coercion and the abstract operations.

Regular Expression Symbols

There are four well known symbols that can be overridden for regular expression objects, which control how those regular expressions are used by the four corresponding `String.prototype` functions of the same name:

- `@@match`: The `Symbol.match` value of a regular expression is the method used to match all or part of a string value with the given regular expression. It's used by `String.prototype.match(..)` if you pass it a regular expression for the pattern matching.

The default algorithm for matching is laid out in section 21.2.5.6 of the ES6 specification (<http://www.ecma-international.org/ecma-262/6.0/#sec-regexp.prototype-@@match>). You could override this default algorithm and provide extra regex features, such as look-behind assertions.

`Symbol.match` is also used by the `isRegExp` abstract operation (see the note in "String Inspection Functions" in Chapter 6) to determine if an object is intended to be used as a regular expression. To force this check to fail for an object so it's not treated as a regular expression, set the `Symbol.match` value to `false` (or something falsy).

- `@@replace`: The `Symbol.replace` value of a regular expression is the method used by `String.prototype.replace(..)` to replace within a string one or all occurrences of character sequences that match the given regular expression pattern.

The default algorithm for replacing is laid out in section 21.2.5.8 of the ES6 specification (<http://www.ecma-international.org/ecma-262/6.0/#sec-regexp.prototype-@@replace>).

One cool use for overriding the default algorithm is to provide additional `replacer` argument options, such as supporting `"abaca".replace(/a/g, [1,2,3])` producing `"1b2c3"` by consuming the iterable for successive replacement values.

- `@@search`: The `Symbol.search` value of a regular expression is the method used by `String.prototype.search(..)` to search for a sub-string within another string as matched by the given regular expression.

The default algorithm for searching is laid out in section 21.2.5.9 of the ES6 specification (<http://www.ecma-international.org/ecma-262/6.0/#sec-regexp.prototype-@@search>).

- **@@split**: The `Symbol.split` value of a regular expression is the method used by `String.prototype.split(..)` to split a string into sub-strings at the location(s) of the delimiter as matched by the given regular expression.

The default algorithm for splitting is laid out in section 21.2.5.11 of the ES6 specification (<http://www.ecma-international.org/ecma-262/6.0/#sec-regexp.prototype-@@split>).

Overriding the built-in regular expression algorithms is not for the faint of heart! JS ships with a highly optimized regular expression engine, so your own user code will likely be a lot slower. This kind of meta programming is neat and powerful, but it should only be used in cases where it's really necessary or beneficial.

`Symbol.isConcatSpreadable`

The `@@isConcatSpreadable` symbol can be defined as a boolean property (`Symbol.isConcatSpreadable`) on any object (like an array or other iterable) to indicate if it should be *spread out* if passed to an array `concat(..)`.

Consider:

```
var a = [1,2,3],  
    b = [4,5,6];  
  
b[Symbol.isConcatSpreadable] = false;  
  
[] .concat( a, b );      // [1,2,3,[4,5,6]]
```

`Symbol.unscopables`

The `@@unscopables` symbol can be defined as an object property (`Symbol.unscopables`) on any object to indicate which properties can and cannot be exposed as lexical variables in a `with` statement.

Consider:

```
var o = { a:1, b:2, c:3 },  
       a = 10, b = 20, c = 30;  
  
o[Symbol.unscopables] = {  
    a: false,  
    b: true,  
    c: false  
};
```

```

with (o) {
  console.log( a, b, c );      // 1 20 3
}

```

A `true` in the `@@unscopables` object indicates the property should be *unscopable*, and thus filtered out from the lexical scope variables. `false` means it's OK to be included in the lexical scope variables.

Warning: The `with` statement is disallowed entirely in `strict` mode, and as such should be considered deprecated from the language. Don't use it. See the *Scope & Closures* title of this series for more information. Because `with` should be avoided, the `@@unscopables` symbol is also moot.

Proxies

One of the most obviously meta programming features added to ES6 is the `Proxy` feature.

A proxy is a special kind of object you create that “wraps” – or sits in front of – another normal object. You can register special handlers (aka *traps*) on the proxy object which are called when various operations are performed against the proxy. These handlers have the opportunity to perform extra logic in addition to *forwarding* the operations on to the original target/wrapped object.

One example of the kind of *trap* handler you can define on a proxy is `get` that intercepts the `[[Get]]` operation – performed when you try to access a property on an object. Consider:

```

var obj = { a: 1 },
  handlers = {
    get(target, key, context) {
      // note: target === obj,
      // context === pobj
      console.log( "accessing: ", key );
      return Reflect.get(
        target, key, context
      );
    }
  },
  pobj = new Proxy( obj, handlers );

obj.a;
// 1

pobj.a;
// accessing: a
// 1

```

We declare a `get(..)` handler as a named method on the `handler` object (second

argument to `Proxy(..)`), which receives a reference to the *target* object (`obj`), the *key* property name ("a"), and the `self/receiver/proxy` (`pobj`).

After the `console.log(..)` tracing statement, we “forward” the operation onto `obj` via `Reflect.get(..)`. We will cover the `Reflect` API in the next section, but note that each available proxy trap has a corresponding `Reflect` function of the same name.

These mappings are symmetric on purpose. The proxy handlers each intercept when a respective meta programming task is performed, and the `Reflect` utilities each perform the respective meta programming task on an object. Each proxy handler has a default definition that automatically calls the corresponding `Reflect` utility. You will almost certainly use both `Proxy` and `Reflect` in tandem.

Here’s a list of handlers you can define on a proxy for a *target* object/function, and how/when they are triggered:

- `get(..)`: via `[[Get]]`, a property is accessed on the proxy (`Reflect.get(..)`,
. property operator, or `[..]` property operator)
- `set(..)`: via `[[Set]]`, a property value is set on the proxy
(`Reflect.set(..)`, the `=` assignment operator, or destructuring assignment if it targets an object property)
- `deleteProperty(..)`: via `[[Delete]]`, a property is deleted from the proxy (`Reflect.deleteProperty(..)` or `delete`)
- `apply(..)` (if *target* is a function): via `[[Call]]`, the proxy is invoked as a normal function/method (`Reflect.apply(..)`, `call(..)`, `apply(..)`, or the `(..)` call operator)
- `construct(..)` (if *target* is a constructor function): via `[[Construct]]`, the proxy is invoked as a constructor function (`Reflect.construct(..)` or `new`)
- `getOwnPropertyDescriptor(..)`: via `[[GetOwnProperty]]`, a property descriptor is retrieved from the proxy (`Object.getOwnPropertyDescriptor(..)` or `Reflect.getOwnPropertyDescriptor(..)`)
- `defineProperty(..)`: via `[[DefineOwnProperty]]`, a property descriptor is set on the proxy (`Object.defineProperty(..)` or `Reflect.defineProperty(..)`)
- `getPrototypeOf(..)`: via `[[GetPrototypeOf]]`, the `[[Prototype]]` of the proxy is retrieved (`Object.getPrototypeOf(..)`, `Reflect.getPrototypeOf(..)`, `__proto__`, `Object#isPrototypeOf(..)`, or `instanceof`)
- `setPrototypeOf(..)`: via `[[SetPrototypeOf]]`, the `[[Prototype]]` of the proxy is set (`Object.setPrototypeOf(..)`, `Reflect.setPrototypeOf(..)`, or `__proto__`)
- `preventExtensions(..)`: via `[[PreventExtensions]]`, the proxy is made non-extensible (`Object.preventExtensions(..)` or `Reflect.preventExtensions(..)`)
- `isExtensible(..)`: via `[[IsExtensible]]`, the extensibility of the proxy is probed (`Object.isExtensible(..)` or `Reflect.isExtensible(..)`)
- `ownKeys(..)`: via `[[OwnPropertyKeys]]`, the set of owned prop-

- `getOwnPropertyNames(..)`: via `Object.getOwnPropertyNames(..)`, `Object.keys(..)`, `Object.getOwnPropertySymbols(..)`, `Reflect.ownKeys(..)`, or `JSON.stringify(..)`
- `enumerate(..)`: via `[[Enumerate]]`, an iterator is requested for the proxy's enumerable owned and “inherited” properties (`Reflect.enumerate(..)` or `for..in`)
- `has(..)`: via `[[HasProperty]]`, the proxy is probed to see if it has an owned or “inherited” property (`Reflect.has(..)`, `Object#hasOwnProperty(..)`, or “`prop`” in `obj`)

Tip: For more information about each of these meta programming tasks, see the “[Reflect API](#)” section later in this chapter.

In addition to the notations in the preceding list about actions that will trigger the various traps, some traps are triggered indirectly by the default actions of another trap. For example:

```
var handlers = {
  getOwnPropertyDescriptor(target,prop) {
    console.log(
      "getOwnPropertyDescriptor"
    );
    return Object.getOwnPropertyDescriptor(
      target, prop
    );
  },
  defineProperty(target,prop,desc){
    console.log( "defineProperty" );
    return Object.defineProperty(
      target, prop, desc
    );
  }
},
proxy = new Proxy( {}, handlers );

proxy.a = 2;
// getOwnPropertyDescriptor
// defineProperty
```

The `getOwnPropertyDescriptor(..)` and `defineProperty(..)` handlers are triggered by the default `set(..)` handler's steps when setting a property value (whether newly adding or updating). If you also define your own `set(..)` handler, you may or may not make the corresponding calls against `context` (not `target!`) which would trigger these proxy traps.

Proxy Limitations

These meta programming handlers trap a wide array of fundamental operations you can perform against an object. However, there are some operations which are not (yet, at least) available to intercept.

For example, none of these operations are trapped and forwarded from `pobj` proxy to `obj` target:

```
var obj = { a:1, b:2 },
  handlers = { .. },
  pobj = new Proxy( obj, handlers );

typeof obj;
String( obj );
obj + "";
obj == pobj;
obj === pobj
```

Perhaps in the future, more of these underlying fundamental operations in the language will be interceptable, giving us even more power to extend JavaScript from within itself.

Warning: There are certain *invariants* – behaviors which cannot be overridden – that apply to the use of proxy handlers. For example, the result from the `isExtensible(..)` handler is always coerced to a `boolean`. These invariants restrict some of your ability to customize behaviors with proxies, but they do so only to prevent you from creating strange and unusual (or inconsistent) behavior. The conditions for these invariants are complicated so we won't fully go into them here, but this post (<http://www.2ality.com/2014/12/es6-proxies.html#invariants>) does a great job of covering them.

Revocable Proxies

A regular proxy always traps for the target object, and cannot be modified after creation – as long as a reference is kept to the proxy, proxying remains possible. However, there may be cases where you want to create a proxy that can be disabled when you want to stop allowing it to proxy. The solution is to create a *revocable proxy*:

```
var obj = { a: 1 },
  handlers = {
    get(target,key,context) {
      // note: target === obj,
      // context === pobj
      console.log( "accessing: ", key );
      return target[key];
    }
  },
```

```

{ proxy: pobj, revoke: prevoke } =
  Proxy.revocable( obj, handlers );

pobj.a;
// accessing: a
// 1

// later:
prevoke();

pobj.a;
// TypeError

```

A revocable proxy is created with `Proxy.revocable(..)`, which is a regular function, not a constructor like `Proxy(..)`. Otherwise, it takes the same two arguments: *target* and *handlers*.

The return value of `Proxy.revocable(..)` is not the proxy itself as with `new Proxy(..)`. Instead, it's an object with two properties: *proxy* and *revoke* – we used object destructuring (see “Destructuring” in Chapter 2) to assign these properties to `pobj` and `prevoke()` variables, respectively.

Once a revocable proxy is revoked, any attempts to access it (trigger any of its traps) will throw a `TypeError`.

An example of using a revocable proxy might be handing out a proxy to another party in your application that manages data in your model, instead of giving them a reference to the real model object itself. If your model object changes or is replaced, you want to invalidate the proxy you handed out so the other party knows (via the errors!) to request an updated reference to the model.

Using Proxies

The meta programming benefits of these Proxy handlers should be obvious. We can almost fully intercept (and thus override) the behavior of objects, meaning we can extend object behavior beyond core JS in some very powerful ways. We'll look at a few example patterns to explore the possibilities.

Proxy First, Proxy Last As we mentioned earlier, you typically think of a proxy as “wrapping” the target object. In that sense, the proxy becomes the primary object that the code interfaces with, and the actual target object remains hidden/protected.

You might do this because you want to pass the object somewhere that can't be fully “trusted,” and so you need to enforce special rules around its access rather than passing the object itself.

Consider:

```

var messages = [],
  handlers = {
    get(target, key) {
      // string value?
      if (typeof target[key] == "string") {
        // filter out punctuation
        return target[key]
          .replace( /[^\w]/g, "" );
      }

      // pass everything else through
      return target[key];
    },
    set(target, key, val) {
      // only set unique strings, lowercased
      if (typeof val == "string") {
        val = val.toLowerCase();
        if (target.indexOf( val ) == -1) {
          target.push(val);
        }
      }
      return true;
    }
  },
  messages_proxy =
    new Proxy( messages, handlers );

// elsewhere:
messages_proxy.push(
  "heLlO...", 42, "wOrld!!", "WoRld!!"
);

messages_proxy.forEach( function(val){
  console.log(val);
} );
// hello world

messages.forEach( function(val){
  console.log(val);
} );
// hello... world!!

```

I call this *proxy first* design, as we interact first (primarily, entirely) with the proxy.

We enforce some special rules on interacting with `messages_proxy` that aren't enforced for `messages` itself. We only add elements if the value is a string

and is also unique; we also lowercase the value. When retrieving values from `messages_proxy`, we filter out any punctuation in the strings.

Alternatively, we can completely invert this pattern, where the target interacts with the proxy instead of the proxy interacting with the target. Thus, code really only interacts with the main object. The easiest way to accomplish this fallback is to have the proxy object in the `[[Prototype]]` chain of the main object.

Consider:

```
var handlers = {
  get(target, key, context) {
    return function() {
      context.speak(key + "!");
    };
  }
},
catchall = new Proxy( {}, handlers ),
greeter = {
  speak(who = "someone") {
    console.log( "hello", who );
  }
};

// setup `greeter` to fall back to `catchall`
Object.setPrototypeOf( greeter, catchall );

greeter.speak();           // hello someone
greeter.speak( "world" );  // hello world

greeter.everyone();        // hello everyone!
```

We interact directly with `greeter` instead of `catchall`. When we call `speak(..)`, it's found on `greeter` and used directly. But when we try to access a method like `everyone()`, that function doesn't exist on `greeter`.

The default object property behavior is to check up the `[[Prototype]]` chain (see the *this & Object Prototypes* title of this series), so `catchall` is consulted for an `everyone` property. The proxy `get()` handler then kicks in and returns a function that calls `speak(..)` with the name of the property being accessed ("`everyone`").

I call this pattern *proxy last*, as the proxy is used only as a last resort.

“No Such Property/Method” A common complaint about JS is that objects aren't by default very defensive in the situation where you try to access or set a property that doesn't already exist. You may wish to predefine all the

properties/methods for an object, and have an error thrown if a nonexistent property name is subsequently used.

We can accomplish this with a proxy, either in *proxy first* or *proxy last* design. Let's consider both.

```
var obj = {
  a: 1,
  foo() {
    console.log( "a:", this.a );
  }
},
handlers = {
  get(target, key, context) {
    if (Reflect.has( target, key )) {
      return Reflect.get(
        target, key, context
      );
    }
    else {
      throw "No such property/method!";
    }
  },
  set(target, key, val, context) {
    if (Reflect.has( target, key )) {
      return Reflect.set(
        target, key, val, context
      );
    }
    else {
      throw "No such property/method!";
    }
  }
},
pobj = new Proxy( obj, handlers );

pobj.a = 3;
pobj.foo();           // a: 3

pobj.b = 4;           // Error: No such property/method!
pobj.bar();           // Error: No such property/method!
```

For both `get(..)` and `set(..)`, we only forward the operation if the target object's property already exists; error thrown otherwise. The proxy object (`pobj`) is the main object code should interact with, as it intercepts these actions to provide the protections.

Now, let's consider inverting with *proxy last* design:

```

var handlers = {
  get() {
    throw "No such property/method!";
  },
  set() {
    throw "No such property/method!";
  }
},
pobj = new Proxy( {}, handlers ),
obj = {
  a: 1,
  foo() {
    console.log( "a:", this.a );
  }
};

// setup `obj` to fall back to `pobj`
Object.setPrototypeOf( obj, pobj );

obj.a = 3;
obj.foo();           // a: 3

obj.b = 4;           // Error: No such property/method!
obj.bar();           // Error: No such property/method!

```

The *proxy last* design here is a fair bit simpler with respect to how the handlers are defined. Instead of needing to intercept the `[[Get]]` and `[[Set]]` operations and only forward them if the target property exists, we instead rely on the fact that if either `[[Get]]` or `[[Set]]` get to our `pobj` fallback, the action has already traversed the whole `[[Prototype]]` chain and not found a matching property. We are free at that point to unconditionally throw the error. Cool, huh?

Proxy Hacking the `[[Prototype]]` Chain The `[[Get]]` operation is the primary channel by which the `[[Prototype]]` mechanism is invoked. When a property is not found on the immediate object, `[[Get]]` automatically hands off the operation to the `[[Prototype]]` object.

That means you can use the `get(..)` trap of a proxy to emulate or extend the notion of this `[[Prototype]]` mechanism.

The first hack we'll consider is creating two objects which are circularly linked via `[[Prototype]]` (or, at least it appears that way!). You cannot actually create a real circular `[[Prototype]]` chain, as the engine will throw an error. But a proxy can fake it!

Consider:

```
var handlers = {
```

```

get(target, key, context) {
  if (Reflect.has( target, key )) {
    return Reflect.get(
      target, key, context
    );
  }
  // fake circular `[[Prototype]]` link
  else {
    return Reflect.get(
      target[
        Symbol.for( "[[Prototype]]" )
      ],
      key,
      context
    );
  }
},
obj1 = new Proxy(
{
  name: "obj-1",
  foo() {
    console.log( "foo:", this.name );
  }
},
{
  handlers
}),
obj2 = Object.assign(
  Object.create( obj1 ),
{
  name: "obj-2",
  bar() {
    console.log( "bar:", this.name );
    this.foo();
  }
});
// fake circular `[[Prototype]]` link
obj1[ Symbol.for( "[[Prototype]]" ) ] = obj2;

obj1.bar();
// bar: obj-1 <-- through proxy faking [[Prototype]]
// foo: obj-1 <-- `this` context still preserved

obj2.foo();

```

```
// foo: obj-2 <-- through [[Prototype]]
```

Note: We didn't need to proxy/forward `[[Set]]` in this example, so we kept things simpler. To be fully `[[Prototype]]` emulation compliant, you'd want to implement a `set(..)` handler that searches the `[[Prototype]]` chain for a matching property and respects its descriptor behavior (e.g., set, writable). See the *this & Object Prototypes* title of this series.

In the previous snippet, `obj2` is `[[Prototype]]` linked to `obj1` by virtue of the `Object.create(..)` statement. But to create the reverse (circular) linkage, we create property on `obj1` at the symbol location `Symbol.for("[[Prototype]]")` (see "Symbols" in Chapter 2). This symbol may look sort of special/magical, but it isn't. It just allows me a conveniently named hook that semantically appears related to the task I'm performing.

Then, the proxy's `get(..)` handler looks first to see if a requested `key` is on the proxy. If not, the operation is manually handed off to the object reference stored in the `Symbol.for("[[Prototype]]")` location of `target`.

One important advantage of this pattern is that the definitions of `obj1` and `obj2` are mostly not intruded by the setting up of this circular relationship between them. Although the previous snippet has all the steps intertwined for brevity's sake, if you look closely, the proxy handler logic is entirely generic (doesn't know about `obj1` or `obj2` specifically). So, that logic could be pulled out into a simple helper that wires them up, like a `setCircularPrototypeOf(..)` for example. We'll leave that as an exercise for the reader.

Now that we've seen how we can use `get(..)` to emulate a `[[Prototype]]` link, let's push the hackery a bit further. Instead of a circular `[[Prototype]]`, what about multiple `[[Prototype]]` linkages (aka "multiple inheritance")? This turns out to be fairly straightforward:

```
var obj1 = {
  name: "obj-1",
  foo() {
    console.log( "obj1.foo:", this.name );
  },
},
obj2 = {
  name: "obj-2",
  foo() {
    console.log( "obj2.foo:", this.name );
  },
  bar() {
    console.log( "obj2.bar:", this.name );
  }
},
handlers = {
  get(target, key, context) {
```

```

        if (Reflect.has( target, key )) {
            return Reflect.get(
                target, key, context
            );
        }
        // fake multiple `[[Prototype]]` links
        else {
            for (var P of target[
                Symbol.for( "[[Prototype]]" )
            ]) {
                if (Reflect.has( P, key )) {
                    return Reflect.get(
                        P, key, context
                    );
                }
            }
        }
    },
    obj3 = new Proxy(
    {
        name: "obj-3",
        baz() {
            this.foo();
            this.bar();
        }
    },
    handlers
);

// fake multiple `[[Prototype]]` links
obj3[ Symbol.for( "[[Prototype]]" ) ] = [
    obj1, obj2
];

obj3.baz();
// obj1.foo: obj-3
// obj2.bar: obj-3

```

Note: As mentioned in the note after the earlier circular `[[Prototype]]` example, we didn't implement the `set(...)` handler, but it would be necessary for a complete solution that emulates `[[Set]]` actions as normal `[[Prototype]]`s behave.

`obj3` is set up to multiple-delegate to both `obj1` and `obj2`. In `obj3.baz()`, the `this.foo()` call ends up pulling `foo()` from `obj1` (first-come, first-served, even though there's also a `foo()` on `obj2`). If we reordered the linkage as `obj2`,

`obj1`, the `obj2.foo()` would have been found and used.

But as is, the `this.bar()` call doesn't find a `bar()` on `obj1`, so it falls over to check `obj2`, where it finds a match.

`obj1` and `obj2` represent two parallel `[[Prototype]]` chains of `obj3`. `obj1` and/or `obj2` could themselves have normal `[[Prototype]]` delegation to other objects, or either could themselves be a proxy (like `obj3` is) that can multiple-delegate.

Just as with the circular `[[Prototype]]` example earlier, the definitions of `obj1`, `obj2`, and `obj3` are almost entirely separate from the generic proxy logic that handles the multiple-delegation. It would be trivial to define a utility like `setPrototypesOf(..)` (notice the "s"!) that takes a main object and a list of objects to fake the multiple `[[Prototype]]` linkage to. Again, we'll leave that as an exercise for the reader.

Hopefully the power of proxies is now becoming clearer after these various examples. There are many other powerful meta programming tasks that proxies enable.

Reflect API

The `Reflect` object is a plain object (like `Math`), not a function/constructor like the other built-in natives.

It holds static functions which correspond to various meta programming tasks that you can control. These functions correspond one-to-one with the handler methods (*traps*) that Proxies can define.

Some of the functions will look familiar as functions of the same names on `Object`:

- `Reflect.getOwnPropertyDescriptor(..)`
- `Reflect.defineProperty(..)`
- `Reflect.getPrototypeOf(..)`
- `Reflect.setPrototypeOf(..)`
- `Reflect.preventExtensions(..)`
- `Reflect.isExtensible(..)`

These utilities in general behave the same as their `Object.*` counterparts. However, one difference is that the `Object.*` counterparts attempt to coerce their first argument (the target object) to an object if it's not already one. The `Reflect.*` methods simply throw an error in that case.

An object's keys can be accessed/inspected using these utilities:

- `Reflect.ownKeys(..)`: Returns the list of all owned keys (not "inherited"), as returned by both `Object.getOwnPropertyNames(..)` and `Object.getOwnPropertySymbols(..)`. See the "Property Enumeration Order" section for information about the order of keys.

- `Reflect.enumerate(..)`: Returns an iterator that produces the set of all non-symbol keys (owned and “inherited”) that are *enumerable* (see the *this & Object Prototypes* title of this series). Essentially, this set of keys is the same as those processed by a `for..in` loop. See the “Property Enumeration Order” section for information about the order of keys.
- `Reflect.has(..)`: Essentially the same as the `in` operator for checking if a property is on an object or its `[[Prototype]]` chain. For example, `Reflect.has(o, "foo")` essentially performs `"foo" in o`.

Function calls and constructor invocations can be performed manually, separate of the normal syntax (e.g., `(..)` and `new`) using these utilities:

- `Reflect.apply(..)`: For example, `Reflect.apply(foo, thisObj, [42, "bar"])` calls the `foo(..)` function with `thisObj` as its `this`, and passes in the `42` and `"bar"` arguments.
- `Reflect.construct(..)`: For example, `Reflect.construct(foo, [42, "bar"])` essentially calls `new foo(42, "bar")`.

Object property access, setting, and deletion can be performed manually using these utilities:

- `Reflect.get(..)`: For example, `Reflect.get(o, "foo")` retrieves `o.foo`.
- `Reflect.set(..)`: For example, `Reflect.set(o, "foo", 42)` essentially performs `o.foo = 42`.
- `Reflect.deleteProperty(..)`: For example, `Reflect.deleteProperty(o, "foo")` essentially performs `delete o.foo`.

The meta programming capabilities of `Reflect` give you programmatic equivalents to emulate various syntactic features, exposing previously hidden-only abstract operations. For example, you can use these capabilities to extend features and APIs for *domain specific languages* (DSLs).

Property Ordering

Prior to ES6, the order used to list an object’s keys/properties was implementation dependent and undefined by the specification. Generally, most engines have enumerated them in creation order, though developers have been strongly encouraged not to ever rely on this ordering.

As of ES6, the order for listing owned properties is now defined (ES6 specification, section 9.1.12) by the `[[OwnPropertyKeys]]` algorithm, which produces all owned properties (strings or symbols), regardless of enumerability. This ordering is only guaranteed for `Reflect.ownKeys(..)` (and by extension, `Object.getOwnPropertyNames(..)` and `Object.getOwnPropertySymbols(..)`).

The ordering is:

1. First, enumerate any owned properties that are integer indexes, in ascending numeric order.

2. Next, enumerate the rest of the owned string property names in creation order.
3. Finally, enumerate owned symbol properties in creation order.

Consider:

```
var o = {};  
  
o[Symbol("c")] = "yay";  
o[2] = true;  
o[1] = true;  
o.b = "awesome";  
o.a = "cool";  
  
Reflect.ownKeys( o );           // [1,2,"b","a",Symbol(c)]  
Object.getOwnPropertyNames( o ); // [1,2,"b","a"]  
Object.getOwnPropertySymbols( o ); // [Symbol(c)]
```

On the other hand, the `[[Enumerate]]` algorithm (ES6 specification, section 9.1.11) produces only enumerable properties, from the target object as well as its `[[Prototype]]` chain. It is used by both `Reflect.enumerate(..)` and `for..in`. The observable ordering is implementation dependent and not controlled by the specification.

By contrast, `Object.keys(..)` invokes the `[[OwnPropertyKeys]]` algorithm to get a list of all owned keys. However, it filters out non-enumerable properties and then reorders the list to match legacy implementation-dependent behavior, specifically with `JSON.stringify(..)` and `for..in`. So, by extension the ordering *also* matches that of `Reflect.enumerate(..)`.

In other words, all four mechanisms (`Reflect.enumerate(..)`, `Object.keys(..)`, `for..in`, and `JSON.stringify(..)`) will match with the same implementation-dependent ordering, though they technically get there in different ways.

Implementations are allowed to match these four to the ordering of `[[OwnPropertyKeys]]`, but are not required to. Nevertheless, you will likely observe the following ordering behavior from them:

```
var o = { a: 1, b: 2 };  
var p = Object.create( o );  
p.c = 3;  
p.d = 4;  
  
for (var prop of Reflect.enumerate( p )) {  
  console.log( prop );  
}  
// c d a b  
  
for (var prop in p) {
```

```

        console.log( prop );
    }
    // c d a b

JSON.stringify( p );
// {"c":3,"d":4}

Object.keys( p );
// ["c", "d"]

```

Boiling this all down: as of ES6, `Reflect.ownKeys(..)`, `Object.getOwnPropertyNames(..)`, and `Object.getOwnPropertySymbols(..)` all have predictable and reliable ordering guaranteed by the specification. So it's safe to build code that relies on this ordering.

`Reflect.enumerate(..)`, `Object.keys(..)`, and `for..in` (as well as `JSON.stringify(..)` by extension) continue to share an observable ordering with each other, as they always have. But that ordering will not necessarily be the same as that of `Reflect.ownKeys(..)`. Care should still be taken in relying on their implementation-dependent ordering.

Feature Testing

What is a feature test? It's a test that you run to determine if a feature is available or not. Sometimes, the test is not just for existence, but for conformance to specified behavior – features can exist but be buggy.

This is a meta programming technique, to test the environment your program runs in to then determine how your program should behave.

The most common use of feature tests in JS is checking for the existence of an API and if it's not present, defining a polyfill (see Chapter 1). For example:

```

if (!Number.isNaN) {
    Number.isNaN = function(x) {
        return x !== x;
    };
}

```

The `if` statement in this snippet is meta programming: we're probing our program and its runtime environment to determine if and how we should proceed.

But what about testing for features that involve new syntax?

You might try something like:

```

try {
    a = () => {};
    ARROW_FUNCS_ENABLED = true;
}

```

```

catch (err) {
  ARROW_FUNCS_ENABLED = false;
}

```

Unfortunately, this doesn't work, because our JS programs are compiled. Thus, the engine will choke on the `() => {}` syntax if it is not already supporting ES6 arrow functions. Having a syntax error in your program prevents it from running, which prevents your program from subsequently responding differently if the feature is supported or not.

To meta program with feature tests around syntax-related features, we need a way to insulate the test from the initial compile step our program runs through. For instance, if we could store the code for the test in a string, then the JS engine wouldn't by default try to compile the contents of that string, until we asked it to.

Did your mind just jump to using `eval(..)`?

Not so fast. See the *Scope & Closures* title of this series for why `eval(..)` is a bad idea. But there's another option with less downsides: the `Function(..)` constructor.

Consider:

```

try {
  new Function( "( () => {} )" );
  ARROW_FUNCS_ENABLED = true;
}
catch (err) {
  ARROW_FUNCS_ENABLED = false;
}

```

OK, so now we're meta programming by determining if a feature like arrow functions *can* compile in the current engine or not. You might then wonder, what would we do with this information?

With existence checks for APIs, and defining fallback API polyfills, there's a clear path for what to do with either test success or failure. But what can we do with the information that we get from `ARROW_FUNCS_ENABLED` being `true` or `false`?

Because the syntax can't appear in a file if the engine doesn't support that feature, you can't just have different functions defined in the file with and without the syntax in question.

What you can do is use the test to determine which of a set of JS files you should load. For example, if you had a set of these feature tests in a bootstrapper for your JS application, it could then test the environment to determine if your ES6 code can be loaded and run directly, or if you need to load a transpiled version of your code (see Chapter 1).

This technique is called *split delivery*.

It recognizes the reality that your ES6 authored JS programs will sometimes be able to entirely run “natively” in ES6+ browsers, but other times need transpilation to run in pre-ES6 browsers. If you always load and use the transpiled code, even in the new ES6-compliant environments, you’re running suboptimal code at least some of the time. This is not ideal.

Split delivery is more complicated and sophisticated, but it represents a more mature and robust approach to bridging the gap between the code you write and the feature support in browsers your programs must run in.

FeatureTests.io

Defining feature tests for all of the ES6+ syntax, as well as the semantic behaviors, is a daunting task you probably don’t want to tackle yourself. Because these tests require dynamic compilation (`new Function(..)`), there’s some unfortunate performance cost.

Moreover, running these tests every single time your app runs is probably wasteful, as on average a user’s browser only updates once in a several week period at most, and even then, new features aren’t necessarily showing up with every update.

Finally, managing the list of feature tests that apply to your specific code base – rarely will your programs use the entirety of ES6 – is unruly and error-prone.

The “<https://featuretests.io>” feature-tests-as-a-service offers solutions to these frustrations.

You can load the service’s library into your page, and it loads the latest test definitions and runs all the feature tests. It does so using background processing with Web Workers, if possible, to reduce the performance overhead. It also uses LocalStorage persistence to cache the results in a way that can be shared across all sites you visit which use the service, which drastically reduces how often the tests need to run on each browser instance.

You get runtime feature tests in each of your users’ browsers, and you can use those tests results dynamically to serve users the most appropriate code (no more, no less) for their environments.

Moreover, the service provides tools and APIs to scan your files to determine what features you need, so you can fully automate your split delivery build processes.

FeatureTests.io makes it practical to use feature tests for all parts of ES6 and beyond to make sure that only the best code is ever loaded and run for any given environment.

Tail Call Optimization (TCO)

Normally, when a function call is made from inside another function, a second *stack frame* is allocated to separately manage the variables/state of that other function invocation. Not only does this allocation cost some processing time, but it also takes up some extra memory.

A call stack chain typically has at most 10-15 jumps from one function to another and another. In those scenarios, the memory usage is not likely any kind of practical problem.

However, when you consider recursive programming (a function calling itself repeatedly) – or mutual recursion with two or more functions calling each other – the call stack could easily be hundreds, thousands, or more levels deep. You can probably see the problems that could cause, if memory usage grows unbounded.

JavaScript engines have to set an arbitrary limit to prevent such programming techniques from crashing by running the browser and device out of memory. That's why we get the frustrating “`RangeError: Maximum call stack size exceeded`” thrown if the limit is hit.

Warning: The limit of call stack depth is not controlled by the specification. It's implementation dependent, and will vary between browsers and devices. You should never code with strong assumptions of exact observable limits, as they may very well change from release to release.

Certain patterns of function calls, called *tail calls*, can be optimized in a way to avoid the extra allocation of stack frames. If the extra allocation can be avoided, there's no reason to arbitrarily limit the call stack depth, so the engines can let them run unbounded.

A tail call is a `return` statement with a function call, where nothing has to happen after the call except returning its value.

This optimization can only be applied in `strict` mode. Yet another reason to always be writing all your code as `strict!`

Here's a function call that is *not* in tail position:

```
"use strict";

function foo(x) {
    return x * 2;
}

function bar(x) {
    // not a tail call
    return 1 + foo( x );
}
```

```
bar( 10 );           // 21
```

`1 + ..` has to be performed after the `foo(x)` call completes, so the state of that `bar(..)` invocation needs to be preserved.

But the following snippet demonstrates calls to `foo(..)` and `bar(..)` where both *are* in tail position, as they're the last thing to happen in their code path (other than the `return`):

```
"use strict";\n\nfunction foo(x) {\n    return x * 2;\n}\n\nfunction bar(x) {\n    x = x + 1;\n    if (x > 10) {\n        return foo( x );\n    }\n    else {\n        return bar( x + 1 );\n    }\n}\n\nbar( 5 );           // 24\nbar( 15 );          // 32
```

In this program, `bar(..)` is clearly recursive, but `foo(..)` is just a regular function call. In both cases, the function calls are in *proper tail position*. The `x + 1` is evaluated before the `bar(..)` call, and whenever that call finishes, all that happens is the `return`.

Proper Tail Calls (PTC) of these forms can be optimized – called tail call optimization (TCO) – so that the extra stack frame allocation is unnecessary. Instead of creating a new stack frame for the next function call, the engine just reuses the existing stack frame. That works because a function doesn't need to preserve any of the current state, as nothing happens with that state after the PTC.

TCO means there's practically no limit to how deep the call stack can be. That trick slightly improves regular function calls in normal programs, but more importantly opens the door to using recursion for program expression even if the call stack could be tens of thousands of calls deep.

We're no longer relegated to simply theorizing about recursion for problem solving, but can actually use it in real JavaScript programs!

As of ES6, all PTC should be optimizable in this way, recursion or not.

Tail Call Rewrite

The hitch, however, is that only PTC can be optimized; non-PTC will still work of course, but will cause stack frame allocation as they always did. You'll have to be careful about structuring your functions with PTC if you expect the optimizations to kick in.

If you have a function that's not written with PTC, you may find the need to manually rearrange your code to be eligible for TCO.

Consider:

```
"use strict";\n\nfunction foo(x) {\n    if (x <= 1) return 1;\n    return (x / 2) + foo( x - 1 );\n}\n\nfoo( 123456 );           // RangeError
```

The call to `foo(x-1)` isn't a PTC because its result has to be added to `(x / 2)` before returning.

However, to make this code eligible for TCO in an ES6 engine, we can rewrite it as follows:

```
"use strict";\n\nvar foo = (function(){\n    function _foo(acc,x) {\n        if (x <= 1) return acc;\n        return _foo( (x / 2) + acc, x - 1 );\n    }\n\n    return function(x) {\n        return _foo( 1, x );\n    };\n})();\n\nfoo( 123456 );           // 3810376848.5
```

If you run the previous snippet in an ES6 engine that implements TCO, you'll get the `3810376848.5` answer as shown. However, it'll still fail with a `RangeError` in non-TCO engines.

Non-TCO Optimizations

There are other techniques to rewrite the code so that the call stack isn't growing with each call.

One such technique is called *trampolining*, which amounts to having each partial result represented as a function that either returns another partial result function or the final result. Then you can simply loop until you stop getting a function, and you'll have the result. Consider:

```
"use strict";

function trampoline( res ) {
    while (typeof res == "function") {
        res = res();
    }
    return res;
}

var foo = (function(){
    function _foo(acc,x) {
        if (x <= 1) return acc;
        return function partial(){
            return _foo( (x / 2) + acc, x - 1 );
        };
    }

    return function(x) {
        return trampoline( _foo( 1, x ) );
    };
})();

foo( 123456 );           // 3810376848.5
```

This reworking required minimal changes to factor out the recursion into the loop in `trampoline(..)`:

1. First, we wrapped the `return _foo ..` line in the `return partial() { .. }` function expression.
2. Then we wrapped the `_foo(1,x)` call in the `trampoline(..)` call.

The reason this technique doesn't suffer the call stack limitation is that each of those inner `partial(..)` functions is just returned back to the `while` loop in `trampoline(..)`, which runs it and then loop iterates again. In other words, `partial(..)` doesn't recursively call itself, it just returns another function. The stack depth remains constant, so it can run as long as it needs to.

Trampolining expressed in this way uses the closure that the inner `partial()` function has over the `x` and `acc` variables to keep the state from iteration to iteration. The advantage is that the looping logic is pulled out into a reusable `trampoline(..)` utility function, which many libraries provide versions of. You can reuse `trampoline(..)` multiple times in your program with different trampolined algorithms.

Of course, if you really wanted to deeply optimize (and the reusability wasn't a concern), you could discard the closure state and inline the state tracking of `acc` into just one function's scope along with a loop. This technique is generally called *recursion unrolling*:

```
"use strict";\n\nfunction foo(x) {\n    var acc = 1;\n    while (x > 1) {\n        acc = (x / 2) + acc;\n        x = x - 1;\n    }\n    return acc;\n}\n\nfoo( 123456 );           // 3810376848.5
```

This expression of the algorithm is simpler to read, and will likely perform the best (strictly speaking) of the various forms we've explored. That may seem like a clear winner, and you may wonder why you would ever try the other approaches.

There are some reasons why you might not want to always manually unroll your recursions:

- Instead of factoring out the trampolining (loop) logic for reusability, we've inlined it. This works great when there's only one example to consider, but as soon as you have a half dozen or more of these in your program, there's a good chance you'll want some reusability to keep things shorter and more manageable.
- The example here is deliberately simple enough to illustrate the different forms. In practice, there are many more complications in recursion algorithms, such as mutual recursion (more than just one function calling itself).

The farther you go down this rabbit hole, the more manual and intricate the *unrolling* optimizations are. You'll quickly lose all the perceived value of readability. The primary advantage of recursion, even in the PTC form, is that it preserves the algorithm readability, and offloads the performance optimization to the engine.

If you write your algorithms with PTC, the ES6 engine will apply TCO to let your code run in constant stack depth (by reusing stack frames). You get the readability of recursion with most of the performance benefits and no limitations of run length.

Meta?

What does TCO have to do with meta programming?

As we covered in the “Feature Testing” section earlier, you can determine at run-time what features an engine supports. This includes TCO, though determining it is quite brute force. Consider:

```
"use strict";  
  
try {  
  (function foo(x){  
    if (x < 5E5) return foo( x + 1 );  
  })( 1 );  
  
  TCO_ENABLED = true;  
}  
catch (err) {  
  TCO_ENABLED = false;  
}
```

In a non-TCO engine, the recursive loop will fail out eventually, throwing an exception caught by the `try..catch`. Otherwise, the loop completes easily thanks to TCO.

Yuck, right?

But how could meta programming around the TCO feature (or rather, the lack thereof) benefit our code? The simple answer is that you could use such a feature test to decide to load a version of your application’s code that uses recursion, or an alternative one that’s been converted/transpiled to not need recursion.

Self-Adjusting Code But here’s another way of looking at the problem:

```
"use strict";  
  
function foo(x) {  
  function _foo() {  
    if (x > 1) {  
      acc = acc + (x / 2);  
      x = x - 1;  
      return _foo();  
    }  
  }  
  
  var acc = 1;  
  
  while (x > 1) {  
    try {
```

```

        _foo();
    }
    catch (err) { }
}

return acc;
}

foo( 123456 );           // 3810376848.5

```

This algorithm works by attempting to do as much of the work with recursion as possible, but keeping track of the progress via scoped variables `x` and `acc`. If the entire problem can be solved with recursion without an error, great. If the engine kills the recursion at some point, we simply catch that with the `try...catch` and then try again, picking up where we left off.

I consider this a form of meta programming in that you are probing during runtime the ability of the engine to fully (recursively) finish the task, and working around any (non-TCO) engine limitations that may restrict you.

At first (or even second!) glance, my bet is this code seems much uglier to you compared to some of the earlier versions. It also runs a fair bit slower (on larger runs in a non-TCO environment).

The primary advantage, other than it being able to complete any size task even in non-TCO engines, is that this “solution” to the recursion stack limitation is much more flexible than the trampolining or manual unrolling techniques shown previously.

Essentially, `_foo()` in this case is a sort of stand-in for practically any recursive task, even mutual recursion. The rest is the boilerplate that should work for just about any algorithm.

The only “catch” is that to be able to resume in the event of a recursion limit being hit, the state of the recursion must be in scoped variables that exist outside the recursive function(s). We did that by leaving `x` and `acc` outside of the `_foo()` function, instead of passing them as arguments to `_foo()` as earlier.

Almost any recursive algorithm can be adapted to work this way. That means it's the most widely applicable way of leveraging TCO with recursion in your programs, with minimal rewriting.

This approach still uses a PTC, meaning that this code will *progressively enhance* from running using the loop many times (recursion batches) in an older browser to fully leveraging TCO'd recursion in an ES6+ environment. I think that's pretty cool!

Review

Meta programming is when you turn the logic of your program to focus on itself (or its runtime environment), either to inspect its own structure or to modify it. The primary value of meta programming is to extend the normal mechanisms of the language to provide additional capabilities.

Prior to ES6, JavaScript already had quite a bit of meta programming capability, but ES6 significantly ramps that up with several new features.

From function name inferences for anonymous functions to meta properties that give you information about things like how a constructor was invoked, you can inspect the program structure while it runs more than ever before. Well Known Symbols let you override intrinsic behaviors, such as coercion of an object to a primitive value. Proxies can intercept and customize various low-level operations on objects, and `Reflect` provides utilities to emulate them.

Feature testing, even for subtle semantic behaviors like Tail Call Optimization, shifts the meta programming focus from your program to the JS engine capabilities itself. By knowing more about what the environment can do, your programs can adjust themselves to the best fit as they run.

Should you meta program? My advice is: first focus on learning how the core mechanics of the language really work. But once you fully know what JS itself can do, it's time to start leveraging these powerful meta programming capabilities to push the language further!

You Don't Know JS: ES6 & Beyond

Chapter 8: Beyond ES6

At the time of this writing, the final draft of ES6 (*ECMAScript 2015*) is shortly headed toward its final official vote of approval by ECMA. But even as ES6 is being finalized, the TC39 committee is already hard at work on features for ES7/2016 and beyond.

As we discussed in Chapter 1, it's expected that the cadence of progress for JS is going to accelerate from updating once every several years to having an official version update once per year (hence the year-based naming). That alone is going to radically change how JS developers learn about and keep up with the language.

But even more importantly, the committee is actually going to work feature by feature. As soon as a feature is spec-complete and has its kinks worked out through implementation experiments in a few browsers, that feature will be considered stable enough to start using. We're all strongly encouraged to adopt features once they're ready instead of waiting for some official standards vote. If you haven't already learned ES6, the time is *past due* to get on board!

As the time of this writing, a list of future proposals and their status can be seen here (<https://github.com/tc39/ecma262#current-proposals>).

Transpilers and polyfills are how we'll bridge to these new features even before all browsers we support have implemented them. Babel, Traceur, and several other major transpilers already have support for some of the post-ES6 features that are most likely to stabilize.

With that in mind, it's already time for us to look at some of them. Let's jump in!

Warning: These features are all in various stages of development. While they're likely to land, and probably will look similar, take the contents of this chapter with more than a few grains of salt. This chapter will evolve in future editions of this title as these (and other!) features finalize.

async functions

In “Generators + Promises” in Chapter 4, we mentioned that there's a proposal for direct syntactic support for the pattern of generators `yielding` promises to a runner-like utility that will resume it on promise completion. Let's take a brief look at that proposed feature, called `async function`.

Recall this generator example from Chapter 4:

```
run( function *main() {  
  var ret = yield step1();  
  ...  
  return stepN();  
})
```

```

try {
    ret = yield step2( ret );
}
catch (err) {
    ret = yield step2Failed( err );
}

ret = yield Promise.all([
    step3a( ret ),
    step3b( ret ),
    step3c( ret )
]);

yield step4( ret );
} )
.then(
    function fulfilled(){
        // `main()` completed successfully
    },
    function rejected(reason){
        // Ops, something went wrong
    }
);

```

The proposed `async` function syntax can express this same flow control logic without needing the `run(..)` utility, because JS will automatically know how to look for promises to wait and resume. Consider:

```

async function main() {
    var ret = await step1();

    try {
        ret = await step2( ret );
    }
    catch (err) {
        ret = await step2Failed( err );
    }

    ret = await Promise.all( [
        step3a( ret ),
        step3b( ret ),
        step3c( ret )
    ] );

    await step4( ret );
}

```

```

main()
.then(
  function fulfilled(){
    // `main()` completed successfully
  },
  function rejected(reason){
    // Oops, something went wrong
  }
);

```

Instead of the `function *main() { .. }` declaration, we declare with the `async function main() { .. }` form. And instead of `yielding` a promise, we `await` the promise. The call to run the function `main()` actually returns a promise that we can directly observe. That's the equivalent to the promise that we get back from a `run(main)` call.

Do you see the symmetry? `async function` is essentially syntactic sugar for the generators + promises + `run(..)` pattern; under the covers, it operates the same!

If you're a C# developer and this `async/await` looks familiar, it's because this feature is directly inspired by C#'s feature. It's nice to see language precedence informing convergence!

Babel, Traceur and other transpilers already have early support for the current status of `async functions`, so you can start using them already. However, in the next section "Caveats", we'll see why you perhaps shouldn't jump on that ship quite yet.

Note: There's also a proposal for `async function*`, which would be called an "async generator." You can both `yield` and `await` in the same code, and even combine those operations in the same statement: `x = await yield y`. The "async generator" proposal seems to be more in flux – namely, its return value is not fully worked out yet. Some feel it should be an *observable*, which is kind of like the combination of an iterator and a promise. For now, we won't go further into that topic, but stay tuned as it evolves.

Caveats

One unresolved point of contention with `async function` is that because it only returns a promise, there's no way from the outside to *cancel* an `async function` instance that's currently running. This can be a problem if the `async` operation is resource intensive, and you want to free up the resources as soon as you're sure the result won't be needed.

For example:

```
async function request(url) {
```

```

var resp = await (
  new Promise( function(resolve,reject){
    var xhr = new XMLHttpRequest();
    xhr.open( "GET", url );
    xhr.onreadystatechange = function(){
      if (xhr.readyState == 4) {
        if (xhr.status == 200) {
          resolve( xhr );
        }
        else {
          reject( xhr.statusText );
        }
      }
    };
    xhr.send();
  } )
);

return resp.responseText;
}

var pr = request( "http://some.url.1" );

pr.then(
  function fulfilled(responseText){
    // ajax success
  },
  function rejected(reason){
    // Oops, something went wrong
  }
);

```

This `request(..)` that I've conceived is somewhat like the `fetch(..)` utility that's recently been proposed for inclusion into the web platform. So the concern is, what happens if you want to use the `pr` value to somehow indicate that you want to cancel a long-running Ajax request, for example?

Promises are not cancelable (at the time of writing, anyway). In my opinion, as well as many others, they never should be (see the *Async & Performance* title of this series). And even if a promise did have a `cancel()` method on it, does that necessarily mean that calling `pr.cancel()` should actually propagate a cancellation signal all the way back up the promise chain to the `async function`?

Several possible resolutions to this debate have surfaced:

- `async functions` won't be cancelable at all (status quo)
- A "cancel token" can be passed to an `async` function at call time

- Return value changes to a cancelable-promise type that's added
- Return value changes to something else non-promise (e.g., observable, or control token with promise and cancel capabilities)

At the time of this writing, `async` functions return regular promises, so it's less likely that the return value will entirely change. But it's too early to tell where things will land. Keep an eye on this discussion.

`Object.observe(..)`

One of the holy grails of front-end web development is data binding – listening for updates to a data object and syncing the DOM representation of that data. Most JS frameworks provide some mechanism for these sorts of operations.

It appears likely that post ES6, we'll see support added directly to the language, via a utility called `Object.observe(..)`. Essentially, the idea is that you can set up a listener to observe an object's changes, and have a callback called any time a change occurs. You can then update the DOM accordingly, for instance.

There are six types of changes that you can observe:

- add
- update
- delete
- reconfigure
- setPrototypeOf
- preventExtensions

By default, you'll be notified of all these change types, but you can filter down to only the ones you care about.

Consider:

```
var obj = { a: 1, b: 2 };

Object.observe(
  obj,
  function(changes){
    for (var change of changes) {
      console.log( change );
    }
  },
  [ "add", "update", "delete" ]
);

obj.c = 3;
// { name: "c", object: obj, type: "add" }

obj.a = 42;
```

```
// { name: "a", object: obj, type: "update", oldValue: 1 }

delete obj.b;
// { name: "b", object: obj, type: "delete", oldValue: 2 }
```

In addition to the main "add", "update", and "delete" change types:

- The "reconfigure" change event is fired if one of the object's properties is reconfigured with `Object.defineProperty(..)`, such as changing its `writable` attribute. See the *this & Object Prototypes* title of this series for more information.
- The "preventExtensions" change event is fired if the object is made non-extensible via `Object.preventExtensions(..)`.

Because both `Object.seal(..)` and `Object.freeze(..)` also imply `Object.preventExtensions(..)`, they'll also fire its corresponding change event. In addition, "reconfigure" change events will also be fired for each property on the object.

- The "setPrototype" change event is fired if the `[[Prototype]]` of an object is changed, either by setting it with the `__proto__` setter, or using `Object.setPrototypeOf(..)`.

Notice that these change events are notified immediately after said change. Don't confuse this with proxies (see Chapter 7) where you can intercept the actions before they occur. Object observation lets you respond after a change (or set of changes) occurs.

Custom Change Events

In addition to the six built-in change event types, you can also listen for and fire custom change events.

Consider:

```
function observer(changes){
  for (var change of changes) {
    if (change.type == "recalc") {
      change.object.c =
        change.object.oldValue +
        change.object.a +
        change.object.b;
    }
  }
}

function changeObj(a,b) {
  var notifier = Object.getNotifier( obj );
```

```

obj.a = a * 2;
obj.b = b * 3;

// queue up change events into a set
notifier.notify( {
  type: "recalc",
  name: "c",
  oldValue: obj.c
} );
}

var obj = { a: 1, b: 2, c: 3 };

Object.observe(
  obj,
  observer,
  ["recalc"]
);

changeObj( 3, 11 );

obj.a;           // 12
obj.b;           // 30
obj.c;           // 3

```

The change set ("recalc" custom event) has been queued for delivery to the observer, but not delivered yet, which is why `obj.c` is still 3.

The changes are by default delivered at the end of the current event loop (see the *Async & Performance* title of this series). If you want to deliver them immediately, use `Object.deliverChangeRecords(observer)`. Once the change events are delivered, you can observe `obj.c` updated as expected:

```
obj.c;           // 42
```

In the previous example, we called `notifier.notify(..)` with the complete change event record. An alternative form for queuing change records is to use `performChange(..)`, which separates specifying the type of the event from the rest of event record's properties (via a function callback). Consider:

```

notifier.performChange( "recalc", function(){
  return {
    name: "c",
    // `this` is the object under observation
    oldValue: this.c
  };
} );

```

In certain circumstances, this separation of concerns may map more cleanly to

your usage pattern.

Ending Observation

Just like with normal event listeners, you may wish to stop observing an object's change events. For that, you use `Object.unobserve(..)`.

For example:

```
var obj = { a: 1, b: 2 };

Object.observe( obj, function observer(changes) {
  for (var change of changes) {
    if (change.type == "setPrototypeOf") {
      Object.unobserve(
        change.object, observer
      );
      break;
    }
  }
} );
```

In this trivial example, we listen for change events until we see the "setPrototypeOf" event come through, at which time we stop observing any more change events.

Exponentiation Operator

An operator has been proposed for JavaScript to perform exponentiation in the same way that `Math.pow(..)` does. Consider:

```
var a = 2;

a ** 4;           // Math.pow( a, 4 ) == 16

a **= 3;          // a = Math.pow( a, 3 )
a;                // 8
```

Note: `**` is essentially the same as it appears in Python, Ruby, Perl, and others.

Objects Properties and ...

As we saw in the "Too Many, Too Few, Just Enough" section of Chapter 2, the `...` operator is pretty obvious in how it relates to spreading or gathering arrays. But what about objects?

Such a feature was considered for ES6, but was deferred to be considered after ES6 (aka "ES7" or "ES2016" or ...). Here's how it might work in that "beyond ES6" timeframe:

```

var o1 = { a: 1, b: 2 },
o2 = { c: 3 },
o3 = { ...o1, ...o2, d: 4 };

console.log( o3.a, o3.b, o3.c, o3.d );
// 1 2 3 4

```

The ... operator might also be used to gather an object's destructured properties back into an object:

```

var o1 = { b: 2, c: 3, d: 4 };
var { b, ...o2 } = o1;

console.log( b, o2.c, o2.d );
// 2 3 4

```

Here, the ...o2 re-gathers the destructured c and d properties back into an o2 object (o2 does not have a b property like o1 does).

Again, these are just proposals under consideration beyond ES6. But it'll be cool if they do land.

Array#includes(..)

One extremely common task JS developers need to perform is searching for a value inside an array of values. The way this has always been done is:

```

var vals = [ "foo", "bar", 42, "baz" ];

if (vals.indexOf( 42 ) >= 0) {
  // found it!
}

```

The reason for the `>= 0` check is because `indexOf(..)` returns a numeric value of 0 or greater if found, or -1 if not found. In other words, we're using an index-returning function in a boolean context. But because -1 is truthy instead of falsy, we have to be more manual with our checks.

In the *Types & Grammar* title of this series, I explored another pattern that I slightly prefer:

```

var vals = [ "foo", "bar", 42, "baz" ];

if (~vals.indexOf( 42 )) {
  // found it!
}

```

The ~ operator here conforms the return value of `indexOf(..)` to a value range that is suitably boolean coercible. That is, -1 produces 0 (falsy), and anything else produces a non-zero (truthy) value, which is what we're for deciding if we found the value or not.

While I think that's an improvement, others strongly disagree. However, no one can argue that `indexOf(..)`'s searching logic is perfect. It fails to find `NaN` values in the array, for example.

So a proposal has surfaced and gained a lot of support for adding a real boolean-returning array search method, called `includes(..)`:

```
var vals = [ "foo", "bar", 42, "baz" ];

if (vals.includes( 42 )) {
  // found it!
}
```

Note: `Array#includes(..)` uses matching logic that will find `NaN` values, but will not distinguish between `-0` and `0` (see the *Types & Grammar* title of this series). If you don't care about `-0` values in your programs, this will likely be exactly what you're hoping for. If you *do* care about `-0`, you'll need to do your own searching logic, likely using the `Object.is(..)` utility (see Chapter 6).

SIMD

We cover Single Instruction, Multiple Data (SIMD) in more detail in the *Async & Performance* title of this series, but it bears a brief mention here, as it's one of the next likely features to land in a future JS.

The SIMD API exposes various low-level (CPU) instructions that can operate on more than a single number value at a time. For example, you'll be able to specify two *vectors* of 4 or 8 numbers each, and multiply the respective elements all at once (data parallelism!).

Consider:

```
var v1 = SIMD.float32x4( 3.14159, 21.0, 32.3, 55.55 );
var v2 = SIMD.float32x4( 2.1, 3.2, 4.3, 5.4 );

SIMD.float32x4.mul( v1, v2 );
// [ 6.597339, 67.2, 138.89, 299.97 ]
```

SIMD will include several other operations besides `mul(..)` (multiplication), such as `sub()`, `div()`, `abs()`, `neg()`, `sqrt()`, and many more.

Parallel math operations are critical for the next generations of high performance JS applications.

WebAssembly (WASM)

Brendan Eich made a late breaking announcement near the completion of the first edition of this title that has the potential to significantly impact the future path of JavaScript: WebAssembly (WASM). We will not be able to cover WASM

in detail here, as it's extremely early at the time of this writing. But this title would be incomplete without at least a brief mention of it.

One of the strongest pressures on the recent (and near future) design changes of the JS language has been the desire that it become a more suitable target for transpilation/cross-compilation from other languages (like C/C++, ClojureScript, etc.). Obviously, performance of code running as JavaScript has been a primary concern.

As discussed in the *Async & Performance* title of this series, a few years ago a group of developers at Mozilla introduced an idea to JavaScript called ASM.js. ASM.js is a subset of valid JS that most significantly restricts certain actions that make code hard for the JS engine to optimize. The result is that ASM.js compatible code running in an ASM-aware engine can run remarkably faster, nearly on par with native optimized C equivalents. Many viewed ASM.js as the most likely backbone on which performance-hungry applications would ride in JavaScript.

In other words, all roads to running code in the browser *lead through JavaScript*.

That is, until the WASM announcement. WASM provides an alternate path for other languages to target the browser's runtime environment without having to first pass through JavaScript. Essentially, if WASM takes off, JS engines will grow an extra capability to execute a binary format of code that can be seen as somewhat similar to a bytecode (like that which runs on the JVM).

WASM proposes a format for a binary representation of a highly compressed AST (syntax tree) of code, which can then give instructions directly to the JS engine and its underpinnings, without having to be parsed by JS, or even behave by the rules of JS. Languages like C or C++ can be compiled directly to the WASM format instead of ASM.js, and gain an extra speed advantage by skipping the JS parsing.

The near term for WASM is to have parity with ASM.js and indeed JS. But eventually, it's expected that WASM would grow new capabilities that surpass anything JS could do. For example, the pressure for JS to evolve radical features like threads – a change that would certainly send major shockwaves through the JS ecosystem – has a more hopeful future as a future WASM extension, relieving the pressure to change JS.

In fact, this new roadmap opens up many new roads for many languages to target the web runtime. That's an exciting new future path for the web platform!

What does it mean for JS? Will JS become irrelevant or “die”? Absolutely not. ASM.js will likely not see much of a future beyond the next couple of years, but the majority of JS is quite safely anchored in the web platform story.

Proponents of WASM suggest its success will mean that the design of JS will be protected from pressures that would have eventually stretched it beyond assumed breaking points of reasonability. It is projected that WASM will become

the preferred target for high-performance parts of applications, as authored in any of a myriad of different languages.

Interestingly, JavaScript is one of the lesser likely languages to target WASM in the future. There may be future changes that carve out subsets of JS that might be tenable for such targeting, but that path doesn't seem high on the priority list.

While JS likely won't be much of a WASM funnel, JS code and WASM code will be able to interoperate in the most significant ways, just as naturally as current module interactions. You can imagine calling a JS function like `foo()` and having that actually invoke a WASM function of that name with the power to run well outside the constraints of the rest of your JS.

Things which are currently written in JS will probably continue to always be written in JS, at least for the foreseeable future. Things which are transpiled to JS will probably eventually at least consider targeting WASM instead. For things which need the utmost in performance with minimal tolerance for layers of abstraction, the likely choice will be to find a suitable non-JS language to author in, then targeting WASM.

There's a good chance this shift will be slow, and will be years in the making. WASM landing in all the major browser platforms is probably a few years out at best. In the meantime, the WASM project (<https://github.com/WebAssembly>) has an early polyfill to demonstrate proof-of-concept for its basic tenets.

But as time goes on, and as WASM learns new non-JS tricks, it's not too much a stretch of imagination to see some currently-JS things being refactored to a WASM-targetable language. For example, the performance sensitive parts of frameworks, game engines, and other heavily used tools might very well benefit from such a shift. Developers using these tools in their web applications likely won't notice much difference in usage or integration, but will just automatically take advantage of the performance and capabilities.

What's certain is that the more real WASM becomes over time, the more it means to the trajectory and design of JavaScript. It's perhaps one of the most important "beyond ES6" topics developers should keep an eye on.

Review

If all the other books in this series essentially propose this challenge, "you (may) not know JS (as much as you thought)," this book has instead suggested, "you don't know JS anymore." The book has covered a ton of new stuff added to the language in ES6. It's an exciting collection of new language features and paradigms that will forever improve our JS programs.

But JS is not done with ES6! Not even close. There's already quite a few features in various stages of development for the "beyond ES6" timeframe. In

this chapter, we briefly looked at some of the most likely candidates to land in JS very soon.

async functions are powerful syntactic sugar on top of the generators + promises pattern (see Chapter 4). `Object.observe(..)` adds direct native support for observing object change events, which is critical for implementing data binding. The `**` exponentiation operator, `...` for object properties, and `Array#includes(..)` are all simple but helpful improvements to existing mechanisms. Finally, SIMD ushers in a new era in the evolution of high performance JS.

Cliché as it sounds, the future of JS is really bright! The challenge of this series, and indeed of this book, is incumbent on every reader now. What are you waiting for? It's time to get learning and exploring!