



# Using Perl 6

Jonathan S. Duff, Moritz Lenz, Carl Mäsak,  
Patrick R. Michaud & Jonathan Worthington

## Credits

**Text** Jonathan S. Duff, Moritz Lenz, Carl Mäsak, Patrick R. Michaud, Jonathan Worthington.

**Front cover** Sebastian Riedel (<http://krai.h.com>)

**Fonts** Adobe Minion® Pro, Adobe Myriad® Pro, B&H Luxi Mono

**LaTeX-Layout** Nikolai Prokoschenko, Luc St-Louis et al. Special thanks to Konrad Mühler for his LaTeX tips collection (<http://www.kfiles.de/latex.php>)

**Editor** chromatic

**Further text contributions** Bruce Gray, Matt Kraai, Gabor Szabo, Hongwen Qiu, Alex Elsayed, Solomon Foster, Nuno Carvalho, Will Coleda, Aaron Sherman, Prakash Kailasa, Carlin Bingham, Dean Serenevy, Jason Felds, Piotr Fusik, Ralf Valerien, Tim Bunce, molecules and snarkyboojum.



Except where otherwise noted, this work is licensed under <http://creativecommons.org/licenses/by-nc-sa/3.0/>

# Contents

<b>1 Preface</b>	1
1.1 Audience . . . . .	2
1.2 Format of this book . . . . .	2
1.3 Relationship between Perl 6 and Perl 5 . . . . .	2
1.4 Perl 6 implementations . . . . .	3
1.5 Installing Rakudo . . . . .	3
1.6 Executing programs . . . . .	4
1.7 Getting involved . . . . .	4
<b>2 The Basics</b>	5
2.1 Exercises . . . . .	13
<b>3 Operators</b>	17
3.1 A Word on Precedence . . . . .	22
3.2 Comparisons and Smart Matching . . . . .	24
3.2.1 Numeric Comparisons . . . . .	25
3.2.2 String Comparisons . . . . .	25
3.2.3 Smart Matching . . . . .	27
<b>4 Subs and Signatures</b>	29
4.1 Declaring A Subroutine . . . . .	29

4.2	Adding Signatures . . . . .	32
4.2.1	The Basics . . . . .	32
4.2.2	Passing Arrays, Hashes and Code . . . . .	34
4.2.3	Interpolating Arrays and Hashes . . . . .	35
4.2.4	Optional Parameters . . . . .	35
4.2.5	Named Parameters . . . . .	36
4.2.6	Slurpy Parameters . . . . .	40
4.3	Returning Results . . . . .	41
4.4	Working With Types . . . . .	42
4.4.1	Basic Types . . . . .	43
4.4.2	Adding Constraints . . . . .	43
4.5	Captures . . . . .	44
4.5.1	Creating And Using A Capture . . . . .	45
4.5.2	Captures In Signatures . . . . .	46
4.6	Unpacking . . . . .	47
4.7	Currying . . . . .	48
4.8	Introspection . . . . .	49
<b>5</b>	<b>Multis</b>	<b>53</b>
5.1	Constraints . . . . .	55
5.2	Narrowness . . . . .	57
5.3	Multiple arguments . . . . .	58
5.4	Bindability checks . . . . .	61
5.5	Nested Signatures in Multi-dispatch . . . . .	62
5.6	Protos . . . . .	63
5.7	Multi Methods . . . . .	63
5.8	Toying with the candidate list . . . . .	64
<b>6</b>	<b>Classes and Objects</b>	<b>65</b>
6.1	Starting with class . . . . .	66

6.2	I can has state? . . . . .	67
6.3	Methods . . . . .	68
6.4	Constructors . . . . .	70
6.5	Consuming our class . . . . .	70
6.6	Exercises . . . . .	72
<b>7</b>	<b>Roles</b>	<b>75</b>
7.1	What is a role? . . . . .	78
7.2	Compile Time Composition . . . . .	79
7.2.1	Multi-methods and composition . . . . .	80
7.2.2	Calling all candidates . . . . .	81
7.2.3	Expressing requirements . . . . .	81
7.3	Runtime Application of Roles . . . . .	82
7.3.1	Differences from compile time composition . . . . .	83
7.3.2	The but operator . . . . .	83
7.4	Parametric Roles . . . . .	84
<b>8</b>	<b>Subtypes</b>	<b>85</b>
<b>9</b>	<b>Pattern matching</b>	<b>89</b>
9.1	Anchors . . . . .	94
9.2	Captures . . . . .	95
9.3	Named regexes . . . . .	97
9.4	Modifiers . . . . .	98
9.5	Backtracking control . . . . .	99
9.6	Substitutions . . . . .	100
9.7	Other Regex Features . . . . .	101
9.8	Match objects . . . . .	102
<b>10</b>	<b>Grammars</b>	<b>105</b>
10.1	Grammar Inheritance . . . . .	108
10.2	Extracting data . . . . .	110

<b>11 Built-in types, operators and methods</b>	115
11.1 Numbers . . . . .	115
11.2 Strings . . . . .	118
11.3 Bool . . . . .	119

# 1

## Preface

Perl 6 is a language specification for which multiple compilers and interpreters exist in various stages of completeness. These implementations have in turn influenced the design of the language by highlighting misfeatures, contradictions, or features of difficult implementation and little benefit. This process of iteration has produced a more cohesive and consistent language specification.

Perl 6 is versatile, intuitive, and flexible. It embraces several paradigms like procedural, object oriented, and functional programming, and offers powerful tools for parsing text.

This book is a work-in-progress. Even releases will contain some amount of TODO comments prior to the printing of the book. We've left them in because they might serve as useful cues for the reader as well as for us authors about what remains to be done. Even so, we pray the reader's indulgence and understanding.

## 1.1 Audience

This book is primarily for people who want to learn Perl 6. It is a series of tutorials, not a comprehensive reference. We expect some experience in another programming language, though no prior knowledge of Perl is necessary. After working through this book, you should have a solid grasp of the basics of Perl 6 sufficient to solve your own problems with Perl 6.

## 1.2 Format of this book

Each chapter opens with a reasonably complete example that illustrates the topic of the chapter. The example is meant to give an overall impression of how the chapter's topic is used in a real program. Following the example will be a detailed explanation of the code as well as some insight into the techniques and idioms used. This is meant to convey the “flavor” of writing Perl 6 programs so that the reader may proceed to write their own native Perl 6 programs rather than programs that resemble some other language<sup>1</sup>.

## 1.3 Relationship between Perl 6 and Perl 5

Perl 6 is the newest member of the family of languages known as Perl. It represents a major break in syntactic and semantic compatibility from Perl 5, thus the increase from 5 to 6. However, this does not mean that Perl 5 is going away. In fact, quite the opposite. Both Perl 5 and Perl 6 have active developer communities which mold the languages. Perl 5 developers try to extend the language in various ways while keeping backwards compatibility with past versions of Perl. Perl 6 developers extend the language by adding new syntactic and semantic features that enable more power and expressiveness without the restriction of backward compatibility with Perl 5 or earlier versions.

---

<sup>1</sup> Some programmers can still write Fortran in any language, however :)



Some might ask, “Why call it Perl if it’s a different language?” Perl is more than the vagaries of syntax. Perl is philosophy (there’s more than one way to do it; easy things easy, hard things possible); Perl is custom (comprehensive testing); Perl is architectural edifice (the Comprehensive Perl Archive Network); and Perl is community (perl5-porters, perl6-language). Both Perl 5 and Perl 6 share these attributes to varying degrees. As well, Perl is syncretic. Just as Perl borrows good ideas from other languages, so Perl 5 and Perl 6 share features.

## 1.4 Perl 6 implementations

Perl 6 is a specification. Any implementation that passes the official test suite can call itself “Perl 6”. Several implementations exist at various levels of maturity. All of the examples in this book will run with the *Rakudo Perl 6* Compiler, but they are in no way specific to Rakudo—any sufficiently advanced Perl 6 implementation can run them. Good luck and—as the Perl 6 community often says—have fun!

## 1.5 Installing Rakudo

For complete instructions for downloading and installing Rakudo, see <http://www.rakudo.org/how-to-get-rakudo>. Source code releases are available from <http://github.com/rakudo/rakudo/downloads>. A binary release for windows is available from <http://sourceforge.net/projects/parrotwin32/files/>.

## 1.6 Executing programs

To run a Perl 6 program with Rakudo, include the install directory in your system PATH variable and issue a command like:

```
$ perl6 hello.pl
```

If you invoke the Rakudo compiler without an explicit script to run, it enters a small interactive mode that allows Perl 6 statements to be executed from the command line.

## 1.7 Getting involved

If you are inspired by the contents of this book and want to contribute to the Perl 6 community, there are more resources available to you:

### World Wide Web

The Perl 6 homepage is <http://perl6.org/>—it contains pointers to many useful resources.

### IRC

The channel `#perl6` on `irc.freenode.net` discusses all things Perl 6.

### Mailing lists

If you need programming help with Perl 6, send an email to [perl6-users@perl.org](mailto:perl6-users@perl.org).

For issues regarding the Perl 6 language specification contact [perl6-language@perl.org](mailto:perl6-language@perl.org).  
For issues regarding Perl 6 compilers, send email to [perl6-compiler@perl.org](mailto:perl6-compiler@perl.org).

# 2

## The Basics

Perl originated as a programming language intended to gather and summarize information from text files. It's still strong in text processing, but Perl 5 is also a powerful general-purpose programming language. Perl 6 is even better.

Suppose that you host a table tennis tournament. The referees tell you the results of each game in the format `Player 1 vs Player 2 | 3:2`, which means that Player 1 won against Player 2 by 3 to 2 sets. You need a script that sums up how many matches and sets each player has won to determine the overall winner.

The input data looks like this:

```
1  Beth Ana Charlie Dave
2  Ana vs Dave | 3:0
3  Charlie vs Beth | 3:1
4  Ana vs Beth | 2:3
5  Dave vs Charlie | 3:0
```

## Chapter 2 | THE BASICS

```
6 Ana vs Charlie | 3:1
7 Beth vs Dave | 0:3
```

The first line is the list of players. Every subsequent line records a result of a match.

Here's one way to solve that problem in Perl 6:

```
1 use v6;
2
3 my $file = open 'scores';
4 my @names = $file.get.split(' ');
5
6 my %matches;
7 my %sets;
8
9 for $file.lines -> $line {
10     my ($pairing, $result) = $line.split(' | ');
11     my ($p1, $p2)          = $pairing.split(' vs ');
12     my ($r1, $r2)          = $result.split(':');
13
14     %sets{$p1} += $r1;
15     %sets{$p2} += $r2;
16
17     if $r1 > $r2 {
18         %matches{$p1}++;
19     } else {
20         %matches{$p2}++;
21     }
22 }
23
24 my @sorted = @names.sort({ %sets{$_} }).sort({ %matches{$_} }).reverse;
25
26 for @sorted -> $n {
27     say "$n has won %matches{$n} matches and %sets{$n} sets";
28 }
```

This produces the output:

```
Ana has won 2 matches and 8 sets
Dave has won 2 matches and 6 sets
Charlie has won 1 matches and 4 sets
Beth has won 1 matches and 4 sets
```

Every Perl 6 program should begin with `use v6;`. This line tells the compiler which version of Perl the program expects. Should you accidentally run the file with Perl 5, you'll get a helpful error message.

A Perl 6 program consists of zero or more statements. A *statement* ends with a semicolon or a curly bracket at the end of a line:

```
1 my $file = open 'scores';
```

`my` declares a lexical variable. Lexical variables are visible only in the current block. If there's no enclosing block, it's visible throughout the remainder of the file. A block is any part of the code enclosed between curly braces `{ }`.

A variable name begins with a *sigil*, which is non-alpha-numeric symbol such as `$`, `@`, `%`, or `&`—or occasionally the double colon `::`. The sigils usually restrict the variable to a particular type, such as a single value or a compound value. After the sigil comes an *identifier*, which may consist of letters, digits and the underscore. Between letters you can also use a dash — or an apostrophe `'`, so `isn't` and `double-click` are valid identifiers.

The `$` sigil indicates a *scalar* variable, which indicates that the variable stores a single value.

The built-in function `open` opens a file, here named *scores*, and returns a *file handle*—an object representing that file. The equality sign `=` *assigns* that file handle to the variable on the left, which means that `$file` now stores the file handle.

`'scores'` is a *string literal*. A string is a piece of text, and a string literal is a string which appears directly in the program. In this line, it's the argument provided to `open`.

```
1 my @names = $file.get.split(' ');
```

The right-hand side calls a *method* –a named group of behavior– named `get` on the file handle stored in `$file`. The `get` method reads and returns one line from the file, removing the line ending. `split` is also a method, called on the string returned from `get`. `split`'s single argument is a string containing a space character. `split` decomposes its *invocant*—the string on which it operates—into a list of strings. It turns the single string `'Beth Ana Charlie Dave'` into the list of strings `'Beth', 'Ana', 'Charlie', 'Dave'`. Finally, this list gets stored in the array `@names`. The `@` sigil marks the declared variable as an Array. Arrays store ordered lists.

Splitting on a single blank is not very robust, and breaks the code when there are trailing spaces on a line, or more than one blank in a row. You will learn a more robust method for data extraction in the chapter about regexes.

TODO: proper cross-link

```
1 my %matches;
2 my %sets;
```

These two lines of code declare two hashes. The `<%>` sigil marks each variable as a Hash. A Hash is an unordered collection of pairs of keys and values. Other programming languages call that a *hash table*, *dictionary*, or *map*. You can query a hash table for the value that corresponds to a certain `$key` with `%hash{$key}`<sup>1</sup>.

In the score counting program, `%matches` stores the number of matches each player has won. `%sets` stores the number of sets each player has won.

Sigils indicate the default access method for a variable. Variables with the `@` sigil are accessed positionally; variables with the `%` sigil are accessed by string key. The `$` sigil, however, indicates a general container that can hold anything and be accessed in any manner. A scalar can even contain a compound object like an Array or a Hash; the `$` sigil signifies that it should be treated as a single value, even in a context that expects multiple values (as with an Array or Hash).

---

<sup>1</sup> Unlike Perl 5, in Perl 6 the sigil does not change when accessing an array or hash with `[ ]` or `{ }`. This is called *sigil invariance*.

```

1 for $file.lines -> $line {
2     ...
3 }

```

`for` produces a loop that runs the *block* delimited by curly brackets and containing ... once for each item of the list, setting the variable `$line` to the current value of each iteration. `$file.lines` produces a list of the lines read from the file *scores*, starting with the line where the previous calls to `$file.get` left off, and going all the way to the end of the file.

During the first iteration, `$line` will contain the string `Ana vs Dave | 3:0`. During the second, `Charlie vs Beth | 3:1`, and so on.

```

1 my ($pairing, $result) = $line.split(' | ');

```

`my` can declare multiple variables simultaneously. The right-hand side of the assignment is again a call to `split`, this time splitting on a vertical bar surrounded by spaces. `$pairing` gets the first item of the returned list, and `$result` the second.

After processing the first line, `$pairing` will hold the string `Ana vs Dave` and `$result` `3:0`.

The next two lines follow the same pattern:

```

1 my ($p1, $p2) = $pairing.split(' vs ');
2 my ($r1, $r2) = $result.split(':');

```

The first extracts and stores the names of the two players in the variables `$p1` and `$p2`. The second extracts the results for each player and stores them in `$r1` and `$r2`.

After processing the first line of the file, the variables contain the values:

The program then counts the number of sets each player has won:

```

1 %sets{$p1} += $r1;
2 %sets{$p2} += $r2;

```

Table 2.1: Contents of Variables

Variable	Contents
\$line	'Ana vs Dave   3:0'
\$pairing	'Ana vs Dave'
\$result	'3:0'
\$p1	'Ana'
\$p2	'Dave'
\$r1	'3'
\$r2	'0'

This is a shortcut for:

```

1 %sets{$p1} = %sets{$p1} + $r1;
2 %sets{$p2} = %sets{$p2} + $r2;

```

`+= $r1` means *increase the value in the variable on the left by \$r1*. In the first iteration `%sets{$p1}` is not yet set, so it defaults to a special value called `Any`. The addition and incrementing operators treat `Any` as a number with the value of zero.

Before these two lines execute, `%sets` is empty. Adding to an entry not in the hash will cause that entry to spring into existence just-in-time, with a value starting at zero. (This is *autovivification*). After these two lines have run for the first time, `%sets` contains `'Ana' => 3`, `'Dave' => 0`. (The fat arrow `=>` separates key and value in a Pair.)

```

1 if $r1 > $r2 {
2     %matches{$p1}++;
3 } else {
4     %matches{$p2}++;
5 }

```



If `$r1` has a larger value than `$r2`, `%matches{$p1}` increments by one. If `$r1` is not larger than `$r2`, `%matches{$p2}` increments. Just as in the case of `+=`, if either hash value did not exist previously, it is autovivified by the increment operation.

`$thing++` is short for `$thing += 1` or `$thing = $thing + 1`, with the small exception that the return value of the expression is `$thing` *before* the increment, not the incremented value. If, as you can do in many other programming languages, you can use `++` as a prefix, it returns the incremented value; `my $x = 1; say ++$x` prints 2.

```
1 my @sorted = @names.sort({ %sets{$_} }).sort({ %matches{$_} }).reverse;
```

This line consists of three individually simple steps. An array's `sort` method returns a sorted version of the array's contents. However, the default sort on an array sorts by its contents. To print player names in winner-first order, the code must sort the array by the *scores* of the players, not their names. The sort method's argument is a *block* used to transform the array elements (the names of players) to the data by which to sort. The array items are passed in through the *topic variable* `$_`.

You have seen blocks before: both the for loop `-> $line { ... }` and the if statement worked on blocks. A block is a self-contained piece of Perl 6 code with an optional signature (the `-> $line` part). See signatures for more information.

The simplest way to sort the players by score would be `@names.sort({ %matches{$_} })`, which sorts by number of matches won. However Ana and Dave have both won two matches. That simple sort doesn't account for the number of sets won, which is the secondary criterion to decide who has won the tournament.

When two array items have the same value, sort leaves them in the same order as it found them. Computer scientists call this a *stable sort*. The program takes advantage of this property of Perl 6's sort to achieve the goal by sorting twice: first by the number of sets won (the secondary criterion), then by the number of matches won.

After the first sorting step, the names are in the order Beth Charlie Dave Ana. After the second sorting step, it's still the same, because no one has won fewer matches but more sets than someone else. Such a situation is entirely possible, especially at larger tournaments.

## Chapter 2 | THE BASICS

`sort` sorts in ascending order, from smallest to largest. This is the opposite of the desired order. Therefore, the code calls the `.reverse` method on the result of the second sort, and stores the final list in `@sorted`.

```
1 for @sorted -> $n {
2     say "$n has won %matches{$n} matches and %sets{$n} sets";
3 }
```

To print out the players and their scores, the code loops over `@sorted`, setting `$n` to the name of each player in turn. Read this code as “For each element of `sorted`, set `$n` to the element, then execute the contents of the following block.” `say` prints its arguments to the standard output (the screen, normally), followed by a newline. (Use `print` if you don’t want the newline at the end.)

When you run the program, you’ll see that `say` doesn’t print the contents of that string verbatim. In place of `$n` it prints the contents of `$n`— the names of players stored in `$n`. This automatic substitution of code with its contents is *interpolation*. This interpolation happens only in strings delimited by double quotes `"..."`. Single quoted strings `'...'` do not interpolate:

```
1 my $names = 'things';
2 say 'Do not call me $names'; # Do not call me $names
3 say "Do not call me $names"; # Do not call me things
```

Double quoted strings in Perl 6 can interpolate variables with the `$` sigil as well as blocks of code in curly braces. Since any arbitrary Perl code can appear within curly braces, Arrays and Hashes may be interpolated by placing them within curly braces.

Arrays within curly braces are interpolated with a single space character between each item. Hashes within curly braces are interpolated as a series of lines. Each line will contain a key, followed by a tab character, then the value associated with that key, and finally a newline.

```
1 "Math: { 1 + 2 }"                # Math: 3
2 @people = <Luke Matthew Mark>;
3 "The synoptics are: {@people}"    # The synoptics are: Luke Matthew Mark
```

```

4
5   "{%sets}";                                # From the table tennis tournament
6
7   harlie 4
8   ave    6
9   na     8
10  eth    4

```

When array and hash variables appear directly in a double-quoted string (and not inside curly brackets), they are only interpolated if their name is followed by a postcircumfix – a bracketing pair that follows a statement. It's also ok to have a method call between the variable name and the postcircumfix.

```

1  @flavours = <vanilla peach>;
2
3  "we have @flavours";           # we have @flavours
4  "we have @flavours[0]";       # we have vanilla
5  o-called "Zen slice"
6  "we have @flavours[]";        # we have vanilla peach
7
8  method calls ending in postcircumfix
9  "we have @flavours.sort()";    # we have peach vanilla
10
11  chained method calls:
12  "we have @flavours.sort.join(', ')"
13                                # we have peach, vanilla

```

## 2.1 Exercises

1. The input format of the example program is redundant: the first line containing the name of all players is not necessary, because you can find out which players participated in the tournament by looking at their names in the subsequent rows.

## Chapter 2 | THE BASICS

How can you change the program if the first input line is omitted? Hint: `%hash.keys` returns a list of all keys stored in `%hash`.

**Answer:** Remove the line `my @names = $file.get.split(' ');`, and change:

```
1 my @sorted = @names.sort({ %sets{$_} }).sort({ %matches{$_} }).reverse;
```

... into:

```
1 my @sorted = %sets.keys.sort({ %sets{$_} }).sort({ %matches{$_} }).reverse;
```

2. Instead of removing the redundancy, you can also use it to warn if a player appears that wasn't mentioned in the first line, for example due to a typo. How would you modify your program to achieve that?

**Answer:** Introduce another hash with the names of the legitimate players as keys, and look in this hash when the name of a player is read:

```
1 ...
2 my @names = $file.get.split(' ');
3 my %legitimate-players;
4 for @names -> $n {
5     %legitimate-players{$n} = 1;
6 }
7
8 ...
9
10 for $file.lines -> $line {
11     my ($pairing, $result) = $line.split(' | ');
12     my ($p1, $p2)         = $pairing.split(' vs ');
13     for $p1, $p2 -> $p {
14         if !%legitimate-players{$p} {
15             say "Warning: '$p' is not on our list!";
16         }
17     }
18 }
```

19            . . .  
20        }



# 3

## Operators

Operators are very short names for often used routines. They have special calling syntax, and can be manipulated by other operators.

Returning to the table tennis example, suppose you want to plot the number of sets that each player won in a tournament. This example uses the numbers from the previous chapter, and makes a very simple text output by just printing a number of X characters to represent horizontal bars.

```
1  use v6;
2
3  my @scores = 'Ana' => 8, 'Dave' => 6, 'Charlie' => 4, 'Beth' => 4;
4
5  my $screen-width = 30;
6
7  my $label-area-width = 1 + [max] @scores».key».chars;
8  my $max-score       = [max] @scores».value;
```

## Chapter 3 | OPERATORS

```
9 my $unit          = ($screen-width - $label-area-width) / $max-score;
10
11 for @scores {
12     my $format = '%- ' ~ $label-area-width ~ "s%\n";
13     printf $format, .key, 'X' x ($unit * .value);
14 }
```

Output:

```
Ana      XXXXXXXXXXXXXXXXXXXXXXXX
Dave     XXXXXXXXXXXXXXXXXX
Charlie  XXXXXXXXXXXX
Beth     XXXXXXXXXXXX
```

The line

```
1 my @scores = 'Ana' => 8, 'Dave' => 6, 'Charlie' => 4, 'Beth' => 4;
```

already contains three different operators: =, => and ,.

The = operator is the *assignment operator* – it takes the values from the right-hand side, and stores them in the variable on the left-hand side, here @scores.

Like other languages that have adopted a similar syntax to C, Perl 6 allows for a shorthand way to write certain assignments. Any assignment of the form `$var = $var op EXPR` can be expressed as `$var op= EXPR`. So, for instance, `~` (tilde) is the string concatenation operator; to append some text to the end of a string, you could say `$string ~= "text"` which is equivalent to `$string = $string ~ "text"`.

The `=>` operator (the fat arrow) constructs Pair objects. A Pair stores a key and a value; the key is on the left-hand side of the `=>` operator, the value on the right. It also has a specialty: if the key on the left is a bare identifier, it is taken to be a string. So one could also write the example above as



```
1 my @scores = Ana => 8, Dave => 6, Charlie => 4, Beth => 4;
```

Finally, the `,` operator constructs a `Parcel`, which is a sequence of objects. In this case the objects are pairs.

All of the three operators above are *infix* operators, which means they stand in between two *terms*. A term can be a literal like 8 or 'Dave', or a combination of other terms and operators.

The previous chapter already used other types of operators, too. It contained the statement `%games{$p1}++`; which uses the *postcircumfix* operator `{...}`. It stands behind (*post*) a term, and consists of two symbols (an opening and a closing curly bracket) which enclose (*circumfix*) another term. Behind this postcircumfix operator is an ordinary *postfix* operator with name `++`, which increments the value it qualifies. No whitespace is allowed between a term and its postfix or postcircumfix operators.

Yet another operator type is the *prefix* operator, which stands in front of a term. An example is the `-` operator, which negates the following numeric value, as in `my $x = -4`.

But the `-` operator can also mean subtraction, so say `5 - 4` will print a 1. To distinguish the prefix operator `-` from the infix operator `-`, the Perl 6 parser always keeps track of whether it expects an infix operator or a term. A term can have zero or more prefix operators, so you can actually write say `4 + -5`. After the `+` (an infix operator), the compiler expects a term, thus the `-` is interpreted as a prefix operator to the term 5.

The next line containing new features is

```
1 my $label-area-width = 1 + [max] @scores».key».chars;
```

It begins harmlessly with a variable declaration `my $label-area-width` and an assignment to it. Next comes a simple numeric addition, `1 + ....`. The right side of the `+` operator is more complicated.

## Chapter 3 | OPERATORS

In Perl 6 there is an infix `max` operator which returns the greater of two values, so `2 max 3` returns 3. Square brackets around an infix operator indicate that it is applied to a list piece by piece. So `[max] 1, 5, 3, 7` is the same as `1 max 5 max 3 max 7` and evaluates to 7.

Likewise, you can write `[+]` to get the sum of a list of values, `[*]` for the product, and use `[<=]` to check if a list is ordered by ascending values.

After the `[max]` you see the expression `@scores».key».chars`. Just like `@variable.method` calls a method on `@variable`, `@array».method` calls a method for each item in `@array`, and returns the list of the return values.

`»` is called a *hyper operator*. It is a unicode character that can be entered on most computers

<sup>1</sup> If your operating system does not make it easy to write it you can also write it using two angle brackets (`>>`).

So `@scores».key` is a list of all the keys of the pair objects in `@scores`, and `@scores».key».chars` is a list of the length of all keys in `@scores`.

The expression `[max] @scores».key».chars` gives the largest of these values. It is the same as

```
1 @scores[0].key.chars
2   max @scores[1].key.chars
3   max @scores[2].key.chars
4   max ...
```

The square brackets are called the *reduction meta operator*: it transforms the enclosed infix operator into an operator that expects a list (a *listop*), and carries out the operation between each two consecutive list items.

For plotting the names of the players and bar charts, the program needs to know how much space to allocate for the player names. Adding 1 to it leaves space for a single blank space between the name of the longest player and the left edge of the bars.

---

<sup>1</sup> Ubuntu 10.4: In System/Preferences/Keyboard/Layouts/Options/Compose Key position select one of the keys to be the “Compose” key. Then press Compose-key and the “greater than” key twice.

```
1 my $max-score = [max] @scores».value;
```

With this line the program determines the maximum score. The drawing area has the width `$screen-width - $label-area-width`, so for each score we can print

```
1 my $unit = ($screen-width - $label-area-width) / $max-score;
```

amount of X characters. This expression uses the infix operators `-` and `/` for numerical calculations.

Now all the necessary informations are in place and the chart can be printed:

```
1 for @scores {  
2     my $format = '%- ' ~ $label-area-width ~ "s%s\n";  
3     printf $format, .key, 'X' x ($unit * .value);  
4 }
```

These lines loop over the items in `@scores`, binding them to the special variable `$_` one at a time. For each such item, the program uses the `printf` built-in function to print both the name of the player and a bar.

It is similar to `printf` functions of the C and Perl 5 programming languages. It takes a format string, which specifies how to print the following parameters. If `$label-area-width` is 8, the format string is `"%- 8s%s\n"`, which means a string ('s') filled to 8 spaces (' 8') and left-justified ('-'), followed by another string and a newline.

The first string is the name of the player, the second is the bar.

The bar is generated with the infix `x` operator, the *repetition operator*. It takes a string on the left-hand side and a number on the right-hand side, and sticks the strings together as many times as the number specifies. So `'ab' x 3` returns the string `'ababab'`. `.value` returns the value of the current pair, `($unit * .value)` multiplies that values with `$unit`, and `'X' x ($unit * .value)` returns as that many X characters.

### 3.1 A Word on Precedence

The explanations of the example above have one implication, which was not yet explicitly mentioned. In the line

```
1 my @scores = 'Ana' => 8, 'Dave' => 6, 'Charlie' => 4, 'Beth' => 4;
```

The right-hand side of the assignment produces a list (because of the `,` operator) that is made of pairs (because of `=>`), and the result is then assigned to the array variable. But you could think of other ways that Perl 6 interprets this program. If you pass this line to the Perl 5 interpreter, it parses it as

```
1 (my @scores = 'Ana') => 8, 'Dave' => 6, 'Charlie' => 4, 'Beth' => 4;
```

and thus stores only one item in the variable `@scores`, the rest is parsed as a list<sup>2</sup>.

The ways in which this statement is parsed in Perl 6 is governed by *precedence rules*. For example, they state that the infix `=>` operator binds its arguments tighter than the infix `,` operator, which in turn binds tighter than the `=` assignment operator<sup>3</sup>.

The precedence rules for Perl 6 allow many commonly used idioms to be expressed naturally and without any parentheses, or even without thinking about precedence. If you want to force a different way of parsing, parentheses can be used around an expression. Then this parenthesis group has the tightest possible precedence.

```
1 say 5 - 7 / 2;      # 5 - 3.5 = 1.5
2 say (5 - 7) / 2;    # (-2) / 2 = -1
```

---

<sup>2</sup> and discarded, because it is not stored in any variable

<sup>3</sup> There are actually two assignment operators with different precedence. When the right-hand side is a scalar, the *item assignment operator* with tight precedence is used, otherwise the loose-precedence *list assignment operator* is used. This allows the two expressions `$a = 1, $b = 2` and `@a = 1, 2` to both mean something sensible: assignment to two variables in a list, and assignment of a two-item list to a single variable

Table 3.1: Precedence table

Examples	Name
	(tightest precedence)
() , 42.5	term
42.rand	method calls and postcircumfixes
\$x++	autoincrement and -decrement
\$x**2	exponentiation operator
?\$x, !\$x	boolean prefix
+\$x, ~\$x	prefix context operators
2*3, 7/5	multiplicative infix operators
1+2, 7-5	additive infix operators
\$x x 3	replication operators
\$x ~".\n"	string concatenation
1&2	conjunctive AND
1 2	conjunctive OR
abs \$x	named unary prefix
\$x cmp 3	non-chaining binary operators
\$x == 3	chaining binary operators
\$x && \$y	tight AND infix
\$x    \$y	tight OR infix
\$x > 0 ?? 1 !! -1	conditional operator
\$x = 1	item assignment
not \$x	loose unary prefix
1, 2	comma
1, 2 Z @a	list infix
@a = 1, 2	list prefix, list assignment
\$x and say "Yes"	loose AND infix
\$x or die "No"	loose OR infix
;	statement terminator
	(loosest precedence)

## 3.2 Comparisons and Smart Matching

There are several ways to compare objects in Perl. You can test for value equivalence using the `===` infix operator. For immutable objects<sup>4</sup>, this is an ordinary value comparison. `"hello" === "hello"` is true because both strings are immutable and have the same value.

For mutable objects, `===` compares their identities. Two objects only share the same identity if, in fact, they are the same object. In the following example, the two arrays `@a` and `@b`, while they *contain* the same values, are two separate array objects which have different identities and are thus *not* equivalent under `===`.

```

1  my @a = 1, 2, 3;
2  my @b = 1, 2, 3;
3
4  say @a === @a;    # 1
5  say @a === @b;    # 0
6
7  # these use identity for value
8
9  say 3 === 3;      # 1
10 say 'a' === 'a';  # 1
11
12 my $a = 'a';
13 say $a === 'a';   # 1
14
```

The `eqv` operator returns `True` only if two objects are of the same type, and of the same structure. With the variables defined above, `@a eqv @b` is true because `@a` and `@b` contain the same values each. On the other hand `'2' eqv 2` returns `False`, because the left argument is a string, the right an integer and so they are not of the same type.

TODO clarify “same structure”

---

<sup>4</sup> Objects whose value *can not* be changed; a literal value. For instance, the literal 7 will always and forever be just a 7.

### 3.2.1 Numeric Comparisons

You can ask if two objects have the same numeric value with the `==` infix operator. If one of the objects is not numeric, Perl will do its best to make it numeric before doing the comparison. If there is no good way to convert an object to a number, the default of 0 is assumed.

```
1 say 1 == 1.0;      # 1
2 say 1 == '1';      # 1
3 say 1 == '2';      # 0
4 say 3 == '3b';     # 1
```

The operators `<`, `<=`, `>=` and `>` can be used to compare the relative size of numbers, `!=` returns True if the two objects differ in their numerical value.

When you use an array or list as a number, it evaluates to the number of items in that list.

```
1 my @colors = <red blue green>;
2
3 if @colors == 3 {
4     say "It's true, @colors contains 3 items";
5 }
```

### 3.2.2 String Comparisons

Just like `==` converts its arguments to numbers before comparing, `eq` as an infix operator compares for string equality, and converts its arguments to strings if necessary.

```
1 if $greeting eq 'hello' {
2     say 'welcome';
3 }
```

Other operators compare strings lexicographically

Number Comparison	String Comparison	Stands for
==	eq	equals
!=	ne	not equal
!==	!eq	not equal
<	lt	less than
<=	le	less or equal
>	gt	greater than
>=	ge	greater or equal

For example 'a' lt 'b' is true, and likewise 'a' lt 'aa'.

!= is really just a convenience for !==, which in turn is really the ! meta operator added to the infix == operator. An equivalent explanation applies to ne and !eq.

### Three-way Comparison

The three-way comparison operators take two operands, and return `Order::Increase` if the left is smaller, `Order::Same` when both are equal, and `Order::Decrease` if the right operand is smaller<sup>5</sup>. For numbers that operator is spelled `<=>`, and for strings `leg` (from *lesser, equal, greater*). The infix `cmp` operator is a type sensitive three-way comparison operator, and compares numbers like `<=>`, string like `leg`, and for example pairs first by key, and then by values if the keys are identical.

```

1 say 10 <=> 5;      # +1
2 say 10 leg 5;      # because '1' lt '5'
3 say 'ab' leg 'a';  # +1, lexicographic comparison

```

A typical use case for three-way comparison is sorting. The `sort` method in lists can take a block or function that takes two values, compares them, and returns a value less than,

<sup>5</sup> `Order::Increase`, `Order::Same`, and `Order::Decrease` are enumerations (enums) which are further explained in subtypes



equal to or, greater than 0. The sort method then orders the values according to that return value.

```
1 say ~<abstract Concrete>.sort;
2 # output: Concrete abstract
3
4 say ~<abstract Concrete>.sort:
5     -> $a, $b { uc($a) leg uc($b) };
6 # output: abstract Concrete
```

The default comparison is case sensitive; by comparing not the values, but their upper case variant, the example above sorts case insensitively.

### 3.2.3 Smart Matching

The various comparison operators that we have seen so far all coerce their arguments to certain types before comparing them. This is useful if you wish to be very specific about what kind of comparison you want and are unsure of the types of the things that are being compared. Perl 6 also provides another operator that allows you to perform comparisons that just Do The Right Thing. It's `~~`, the smart match operator.

```
1 if $pints-drunk ~~ 8 {
2     say "Go home, you've had enough!";
3 }
4
5 if $country ~~ 'Sweden' {
6     say "Meatballs with lingonberries and potato moose, please."
7 }
8
9 unless $group-size ~~ 2..4 {
10     say "You must have between 2 and 4 people to book this tour.";
11 }
```

The smart match operator always decides what kind of comparison to do based upon the type of the value on the right hand side. In the three examples above, it would do a numeric,

## Chapter 3 | OPERATORS

a string and a range comparison respectively. While we've already seen operators to do numeric and string comparisons – `==` and `eq` – there is no operator for comparing ranges. This is part of the power of smart matching: more complex types can define interesting and useful ways to compare themselves to other things.

Smart match works by calling the `ACCEPTS` method on the operand on the right hand side and passing it the operand on the left hand side as an argument. Thus `$answer == 42` actually desugars to a method call like `42.ACCEPTS($answer)`. The upshot of this is that – after reading the chapter on writing classes and methods – you too will be able to write things that can be smart-match against just by implementing an `ACCEPTS` method to do the right thing.

TODO: explain 'desugars' terminology

TODO: add `:=` operator as that's used in `subs-n-sigs.pod`

# 4

## Subs and Signatures

A *subroutine* is a piece of code that performs a specific task. It may operate on provided data (*arguments*) and may produce results (*return values*). The *signature* of a subroutine is a description of any arguments it takes and any return values it produces.

The first chapter demonstrated simple subroutines. In one sense, the operators described in the second chapter are also subroutines that Perl 6 parses in interesting ways. However, they only scratch the surface of what's possible.

### 4.1 Declaring A Subroutine

A subroutine declaration consists of several parts. First, the subroutine declarator `sub` indicates that you are starting a subroutine declaration. Next comes an optional name and an

## Chapter 4 | SUBS AND SIGNATURES

optional signature. The body of the sub follows as a block of code enclosed in curly braces. This is what will execute every time the subroutine is called.

For example, in:

```
1 sub panic() {  
2     say "Oh no! Something has gone most terribly wrong!";  
3 }
```

... the name of the sub is `panic`. Its signature is empty. Its body consists of a single `say` statement.

By default, subroutines are lexically scoped, just like any variable declared with `my`. This means that a subroutine may only be called within the scope in which it was declared. Use the scope declarator `our` to make the subroutine available within the current package:

```
1 {  
2     our sub eat() {  
3         say "om nom nom";  
4     }  
5  
6     sub drink() {  
7         say "glug glug";  
8     }  
9 }  
10  
11 eat();    # om nom nom  
12 drink(); # fails, can't drink outside of the block
```

`our` also makes subroutines visible from the outside of a package or module:

```
1 module EatAndDrink {  
2     our sub eat() {  
3         say "om nom nom";  
4     }  
5 }
```

```

6      sub drink() {
7          say "glug glug";
8      }
9
10     atAndDrink::eat();    # om nom nom
11     atAndDrink::drink(); # fails, not declared with "our"

```

You may also export (exporting) a subroutine to make it available to another scope.

Perl 6 subroutines are objects. You can pass them around and store them in data structures just as you can do with any other piece of data. Programming language designers often call these *first-class subroutines*; they are as fundamental to and usable from the language as are hashes or arrays.

First-class subroutines can help you solve complex problems. For example, to make a little ASCII art dancing figure, you could build up a hash where the keys are names of the dance moves, and the values are anonymous subroutines. Assume that users should be able to enter a list of moves (perhaps on a dance pad or other exotic input device). How can you maintain a variable list of custom behaviors, allow user input, and restrict that input to a safe set of behaviors?

TODO this example doesn't seem like a good one for first-class subs.

```

1  my $dance = '';
2  my %moves =
3      hands-over-head => sub { $dance ~= '/o\ ' },
4      bird-arms       => sub { $dance ~= '|/o\| ' },
5      left            => sub { $dance ~= '>o ' },
6      right           => sub { $dance ~= 'o< ' },
7      arms-up         => sub { $dance ~= '\o/ ' };
8
9  my @awesome-dance = <arms-up bird-arms right hands-over-head>;
10
11  for @awesome-dance -> $move {
12      %moves{$move}();
13  }

```

```
14
15 say $dance;
```

From the output of this program, you can observe that doing the YMCA dance in ASCII art looks just as bad as in real life.

## 4.2 Adding Signatures

A subroutine signature performs two tasks. First, it declares the arguments which callers may or must pass to the subroutine. Second, it declares the variables in the subroutine to which the arguments are bound. These variables are called *parameters*. Perl 6 signatures go further; they allow you to constrain the values of arguments and to match against and extract parts of complex data structures.

TODO A third task: specifying a return type.

### 4.2.1 The Basics

In its simplest form, a signature is a comma separated list of variable names to which to bind incoming arguments.

```
1 sub order-beer($type, $pints) {
2     say ($pints == 1 ?? 'A pint' !! "$pints pints") ~ " of $type, please."
3 }
4
5 order-beer('Hobgoblin', 1);    # A pint of Hobgoblin, please.
6 order-beer('Zlatý Bažant', 3); # 3 pints of Zlatý Bažant, please.
```

The use of the term *bound* instead of *assigned* is significant. The variables in your signature are read-only references to the passed arguments. You cannot modify passed-in values within the subroutine.

If read-only binding is too limiting, you can relax this restriction. A parameter marked `is rw` means that you can modify the passed argument within the subroutine. Any modification will modify the original in place. If you attempt to pass a literal or some other constant value for an `rw` parameter, the binding will fail at the point of the call, throwing an exception:

```
1 sub make-it-more-so($it is rw) {
2     $it =~ substr($it, $it.chars - 1) x 5;
3 }
4
5 my $happy = "yay!";
6 make-it-more-so($happy);
7 say $happy;           # yay!!!!!!
8 make-it-more-so("uh-oh"); # Fails; can't modify a constant
```

If, instead, you want your own copy of the argument to work with inside the subroutine—to leave the original untouched—mark the parameter `is copy`:

```
1 sub say-it-one-higher($it is copy) {
2     $it++;
3     say $it;
4 }
5
6 my $unanswer = 41;
7 say-it-one-higher($unanswer); # 42
8 say-it-one-higher(41);       # 42
```

The extra verbosity of marking parameters as mutable may seem excessive, but it's likely you won't use these modifiers often. While certain languages require you to mark parameters as `rw` to emulate returning multiple results from a single subroutine, Perl allows you to return multiple values directly.

### 4.2.2 Passing Arrays, Hashes and Code

A variable's sigil indicates its intended use. In a signature, a variable's sigil acts as a constraint on the type of argument passed. The @ sigil, for example, checks that the passed value conforms to the `Positional` type, which encompasses types like `Array` and `list`. Failing to pass something that matches this constraint will cause the call to fail.

```

1 sub shout-them(@words) {
2     for @words -> $w {
3         print uc("$w ");
4     }
5 }
6
7 my @last_words = <do not want>;
8 shout-them(@last_words); # DO NOT WANT
9 shout-them('help');      # Fails; a string is not Positional

```

Similarly, the % sigil implies that the caller must pass something that is `Associative`—something which allows indexing through the `<...>` or `{...}` operations. The & sigil requires that the caller pass something callable, such as an anonymous subroutine. In that case, you may also call the callable parameter without the & sigil:

```

1 sub do-it-lots(&it, $how-many-times) {
2     for 1..$how-many-times {
3         it();
4     }
5 }
6
7 do-it-lots(sub { say "Eating a stroopwafel" }, 10);

```

A scalar (the \$ sigil) implies no constraints. Anything may bind to it, even if it could also bind to one of the other sigils.



### 4.2.3 Interpolating Arrays and Hashes

Sometimes you want to fill positional arguments from an array. Instead of writing `eat(@food[0], @food[1], @food[2], ...)` and so on, you can *flatten* them into the argument list by prepending a vertical bar: `eat(|@food)`.

Likewise, you can interpolate hashes into named arguments:

```
1 sub order-shrimps($count, $from) {
2     say "I'd like $count pieces of shrimp from the $from, please";
3 }
4
5 my %user-preferences = ( from => 'Northern Sea' );
6 order-shrimps(3, |%user-preferences)
```

### 4.2.4 Optional Parameters

Sometimes, passing an argument may be unnecessary. Other parameters can have sensible default values. In these cases, it is possible to mark such parameters as optional, so those calling the subroutine can choose whether to pass an argument.

Either assign a default value to the parameter in the signature:

```
1 sub order-steak($how = 'medium') {
2     say "I'd like a steak, $how";
3 }
4
5 order-steak();
6 order-steak('well done');
```

... or append a question mark to the parameter's name, in which case the parameter binds to an undefined value if no argument is passed:

## Chapter 4 | SUBS AND SIGNATURES

```
1 sub order-burger($type, $side?) {
2     say "I'd like a $type burger" ~
3     ( defined($side) ?? " with a side of $side" !! "" );
4 }
5
6 order-burger("triple bacon", "deep fried onion rings");
```

### 4.2.5 Named Parameters

When a subroutine has many parameters, it is often easier to pass parameters by name instead of trying to remember the correct order of parameters. Note that when you do so, the order in which they appear as arguments does not matter:

```
1 sub order-beer($type, $pints) {
2     say ($pints == 1 ?? 'A pint' !! "$pints pints") ~ " of $type, please."
3 }
4
5 order-beer(type => 'Hobgoblin', pints => 1);
6 # A pint of Hobgoblin, please.
7
8 order-beer(pints => 3, type => 'Zlatý Bažant');
9 # 3 pints of Zlatý Bažant, please.
```

You may also specify that an incoming argument may only fill a parameter when passed by name, such that no positional argument may successfully bind to it. To do this, precede the name of the parameter with a colon:

```
1 sub order-shrimps($count, :$from = 'North Sea') {
2     say "I'd like $count pieces of shrimp from the $from, please";
3 }
4
5 order-shrimps(6); # takes 'North Sea'
6 order-shrimps(4, from => 'Atlantic Ocean');
7 order-shrimps(22, 'Mediterranean Sea'); # not allowed, :$from is named only
```

Unlike positional parameters, named parameters are optional by default. Appending a `!` makes it mandatory.

```
1 sub design-ice-cream-mixture(:$name!, $base = 'Vanilla') {
2     say "Creating a new recipe named $name!"
3 }
4
5 design-ice-cream-mixture(name => 'Plain');
6 design-ice-cream-mixture(base => 'Strawberry chip'); # missing $name
```

## Renaming Parameters

Because you must use their names when passing named parameters, parameter names are part of a subroutine's public API. Choose them carefully! Sometimes it may be convenient to expose a parameter with one name while binding to a variable of a different name:

```
1 sub announce-time(:dinner($supper) = '8pm') {
2     say "We eat dinner at $supper";
3 }
4
5 announce-time(dinner => '9pm');      # We eat dinner at 9pm
```

Parameters can also have multiple names. If some of your users are British and others are Americans, you might write:

```
1 sub paint-rectangle(
2     :$x      = 0,
3     :$y      = 0,
4     :$width  = 100,
5     :$height = 50,
6     :color(:colour($c))) {
7
8     # print a piece of SVG that represents a rectangle
9     say qq[<rect x="$x" y="$y" width="$width" height="$height"
10         style="fill: $c" />]
```

## Chapter 4 | SUBS AND SIGNATURES

```
11 }
12
13 # both calls work the same
14 paint-rectangle :color<Blue>;
15 paint-rectangle :colour<Blue>;
16
17 # of course you can still fill the other options
18 paint-rectangle :width(30), :height(10), :colour<Blue>;
```

### Alternative Named Argument Syntaxes

Named arguments are actually Pairs (of keys and values). There are multiple ways to write Pairs. The difference between the approaches is primarily one of clarity, as each alternative provides a different quoting mechanism. These three calls all mean the same thing:

```
1 announce-time(dinner => '9pm');
2 announce-time(:dinner('9pm'));
3 announce-time(:dinner<9pm>);
```

If you're passing a boolean value, you may omit the value portion of the pair:

```
1 toggle-blender( :enabled); # enables the blender
2 toggle-blender(:!enabled); # disables the blender
```

A named argument of the form `:name` with no value has an implicit value of `Bool::True`. The negated form of this, `:!name`, has an implicit value of `Bool::False`.

If you use a variable to create a pair, you can reuse the variable name as the key of the pair.

```
1 my $dinner = '9pm';
2 announce-dinner :$dinner; # same as dinner => $dinner;
```

`pair_forms` lists possible Pair forms and their meanings.

Table 4.1: C<Pair> forms and their meanings

Shorthand	Long form	Description
:allowed	allowed => Bool::True	Boolean flag
:!allowed	allowed => Bool::False	Boolean flag
:bev<tea coffee>	bev => ('tee', 'coffee')	List
:times[1, 3]	times => [1, 3]	Array
:opts{ a => 2 }	opts => { a => 2 }	Hash
:\$var	var => \$var	Scalar variable
:@var	var => @var	Array variable
:%var	var => %var	Hash variable

You can use any of these forms in any context where you can use a Pair object. For example, when populating a hash:

```

1  # TODO: better example
2  my $black = 12;
3  my %color-popularities = :$black, :blue(8),
4                           red => 18, :white<0>;
5  # same as
6  # my %color-popularities =
7  #     black => 12,
8  #     blue  => 8,
9  #     red   => 18,
10 #     white => 0;
```

Finally, to pass an existing Pair object to a subroutine by position, not name, either put it in parentheses (like (: \$thing)), or use the => operator with a quoted string on the left-hand side: "thing" => \$thing.

## 4.2.6 Slurpy Parameters

In an earlier example the function `shout-it` accepted an array argument. This prevented users from passing in a single argument. To enable both possibilities, or to allow multiple positional arguments and even multiple array arguments all of which will flatten into a single array parameter in the subroutine, prepend the *slurpy* prefix (`*`) to the parameter name:

```

1  sub shout-them(*@words) {
2      for @words -> $w {
3          print uc("$w ");
4      }
5  }
6
7  # now you can pass items
8  shout-them('go');           # GO
9  shout-them('go', 'home');   # GO HOME
10
11 my @words = ('go', 'home');
12 shout-them(@words);         # still works

```

A slurpy parameter—a parameter preceded by an asterisk (`*`)—stores all remaining unbound positional arguments in an array. Likewise, `%%hash` slurps all the remaining unbound named arguments into a hash.

Slurpy arrays and hashes allow you to pass all positional and named arguments to another routine:

```

1  sub debug-wrapper(&code, *@positional, %%named) {
2      warn "Calling '&code.name()' with arguments "
3          ~ "@positional.perl(), %named.perl()\n";
4      code(|@positional, |%named);
5      warn "... back from '&code.name()'\n";
6  }
7
8  debug-wrapper(&order-shrimps, 4, from => 'Atlantic Ocean');

```

## 4.3 Returning Results

Subroutines can also return values. The ASCII art dancing example from earlier in this chapter is simpler when each subroutine returns a new string:

```
1 my %moves =
2     hands-over-head => sub { return '/o\ ' },
3     bird-arms       => sub { return '|/o\| ' },
4     left            => sub { return '>o ' },
5     right           => sub { return 'o< ' },
6     arms-up         => sub { return '\o/ ' };
7
8 my @awesome-dance = <arms-up bird-arms right hands-over-head>;
9
10 for @awesome-dance -> $move {
11     print %moves{$move}();
12 }
13
14 print "\n";
```

A Perl subroutine can return multiple values:

```
1 sub menu {
2     if rand < 0.5 {
3         return ('fish', 'white wine')
4     } else {
5         return ('steak', 'red wine');
6     }
7 }
8
9 my ($food, $beverage) = menu();
```

If you exclude the return statement, Perl will return the value produced by the last statement run inside the subroutine. This simplifies the previous example:

## Chapter 4 | SUBS AND SIGNATURES

```
1 sub menu {
2     if rand < 0.5 {
3         'fish', 'white wine'
4     } else {
5         'steak', 'red wine';
6     }
7 }
8
9 my ($food, $beverage) = menu();
```

Be wary of relying on this: when the flow of control within a subroutine is sufficiently complex, adding an explicit return will clarify the code. As a general rule, only the simplest subroutines benefit from implicit return.

return has the additional effect of immediately exiting the subroutine:

```
1 sub create-world(*%characteristics) {
2     my $world = World.new(%characteristics);
3     return $world if %characteristics<temporary>;
4
5     save-world($world);
6 }
```

... and you'd better not misplace your new \$world if it's temporary, as it's the only one you're going to get.

### 4.4 Working With Types

Many subroutines cannot meaningfully work with arbitrary parameters, but require that the parameters support certain methods or have other properties. In these cases, it makes sense to restrict the types of parameters, such that attempts to pass incorrect values as arguments will cause Perl to raise an error at the time of calling the subroutine.



### 4.4.1 Basic Types

The easiest way to restrict the possible values that a subroutine accepts is by writing a type name before a parameter. For example, a subroutine that performs numeric calculations on its parameters could require that its arguments are of the type `Numeric`:

```
1 sub mean(Numeric $a, Numeric $b) {  
2     return ($a + $b) / 2;  
3 }  
4  
5 say mean 2.5, 1.5;  
6 say mean 'some', 'strings';
```

This produces the output:

```
2  
Nominal type check failed for parameter '$a';  
  expected Numeric but got Str instead
```

If multiple parameters have type constraints, each argument must fulfill the type constraint of the parameter to which it binds.

### 4.4.2 Adding Constraints

Sometimes a type name is insufficient to describe the requirements for an argument. In this case, you may add an additional *constraint* to the parameter with a `where` block:

```
1 sub circle-radius-from-area(Numeric $area where { $area >= 0 }) {  
2     ($area / pi).sqrt  
3 }  
4  
5 say circle-radius-from-area(3);    # OK  
6 say circle-radius-from-area(-3);  # Error
```

## Chapter 4 | SUBS AND SIGNATURES

Because the calculation is meaningful only for non-negative area values, the parameter includes a constraint which returns True for non-negative values. If this constraint returns a false value, the type check will fail when something calls this subroutine.

The block after the where optional; Perl performs the check by smart matching the argument against whatever follows the where. It is possible to accept arguments in a certain range by writing:

```
1 sub set-volume(Numeric $volume where 0..11) {
2     say "Turning it up to $volume";
3 }
```

To constrain arguments to those existing keys of a hash:

```
1 my %in-stock = 'Staropramen' => 8, 'Mori' => 5, 'La Trappe' => 9;
2
3 sub order-beer(Str $name where %in-stock) {
4     say "Here's your $name";
5     %in-stock{$name}--;
6     if %in-stock{$name} == 0 {
7         say "OH NO! That was the last $name, folks! :'(";
8         %in-stock.delete($name);
9     }
10 }
```

### 4.5 Captures

In one sense, a signature is a collection of parameters. Captures fill the same niche for arguments. Just as you rarely think of a signature as a whole—instead focusing on individual parameters—you rarely have to think about captures. When you do, Perl 6 allows you to manipulate captures directly.

Captures have both positional and named parts which act like lists and hashes, respectively. The list-like parts contain positional arguments and the hash-like parts contain named arguments.

### 4.5.1 Creating And Using A Capture

To build a capture, use the `\(...)` syntax. Like arrays and hashes, you can interpolate a capture into an argument by using `|`:

```
1 sub act($left, $right, :$action) {
2     $action($left, $right);
3 }
4
5 my @tasks = \(39, 3, action => { say $^a + $^b }),
6             \(6, 7, action => { say $^a * $^b });
7
8 for @tasks -> $task-args {
9     act(|$task-args);
10 }
```

This program creates an array of captures, each of which contains two positional arguments and one named argument. It then iterates over the array, making a call to `act` with each argument set. Perl 6 allows you to specify the arguments for a call and the call itself separately, so as to apply the same arguments over many calls, or the same call to many sets of arguments. The code that performs the application need not know whether any of the arguments are named or positional.

Unlike signatures, captures work like references. Any variable mentioned in a capture exists in the capture as a *reference* to the variable. Thus `rw` parameters still work with captures involved.

```
1 my $value = 7;
2 my $to-change = \($value);
3
```

```
4 sub double($x is rw) {  
5     $x *= 2;  
6 }  
7  
8 sub triple($x is rw) {  
9     $x *= 3;  
10 }  
11  
12 triple(|$to-change);  
13 double(|$to-change);  
14  
15 say $value; # 42
```

Perl types with both positional and named parts also show up in various other situations. For example, regex matches have both positional and named matches—Match objects themselves are a type of capture. It's also possible to conceive of an XML node type that is a type of capture, with named attributes and positional children. Binding this node to a function could use the appropriate parameter syntax to work with various children and attributes.

### 4.5.2 Captures In Signatures

All calls build a capture on the caller side and unpack it according to the signature on the callee side<sup>1</sup>. It is also possible to write a signature that binds the capture itself into a variable. This is especially useful for writing routines that delegate to other routines with the same arguments.

```
1 sub visit-czechoslovakia(|$plan) {  
2     warn "Sorry, this country has been deprecated.";  
3     visit-slovakia(|$plan);  
4     visit-czech-republic(|$plan);  
5 }
```

---

<sup>1</sup> An optimizing Perl 6 compiler may, of course, be able to optimize away part or all of this process, depending on what it knows at compilation time.

The benefit of using this over a signature like `:(*@pos, *%named)` is that these both enforce some context on the arguments, which may be premature. For example, if the caller passes two arrays, they would flatten into `@pos`. This means that the two nested arrays could not be recovered at the point of delegation. A capture preserves the two array arguments, so that the final callee's signature may determine how to bind them.

## 4.6 Unpacking

Sometimes you need to work with only part of an array or a hash. You can do that with ordinary slicing access, or you can use signature binding:

```
1 sub first-is-largest(@a) {
2     my $first = @a.shift;
3     # TODO: either explain junctions, or find a
4     # concise way to write without them
5     return $first >= all(@a);
6 }
7
8 # same thing:
9 sub first-is-largest(@a) {
10     my :($first, *@rest) := \(@a)
11     return $first >= all(@rest);
12 }
```

The signature binding approach might seem clumsy, but when you use it in the main signature of a subroutine, you get tremendous power:

```
1 sub first-is-largest([$first, *@rest]) {
2     return $first >= all(@rest);
3 }
```

The brackets in the signature tell the compiler to expect a list-like argument. Instead of binding to an array parameter, it instead *unpacks* its arguments into several parameters—in this case, a scalar for the first element and an array for the rest. This *subsignature* also acts

## Chapter 4 | SUBS AND SIGNATURES

as a constraint on the array parameter: the signature binding will fail unless the list in the capture contains at least one item.

Likewise you can unpack a hash by using `%(...)` instead of square brackets, but you must access named parameters instead of positional.

```
1 sub create-world(%(:$temporary, *%characteristics)) {
2     my $world = World.new(%characteristics);
3     return $world if $temporary;
4
5     save-world($world);
6 }
```

# TODO: come up with a good example # maybe steal something from <http://jnthn.net/papers/2010-yapc-eu-signatures.pdf>

# TODO: generic object unpacking

### 4.7 Currying

Consider a module that provided the example from the “Optional Parameters” section:

```
1 sub order-burger( $type, $side? ) { ... };
```

If you used `order-burger` repeatedly, but often with a side of french fries, you might wish that the author had also provided a `order-burger-and-fries` sub. You could easily write it yourself:

```
1 sub order-burger-and-fries ( $type ) {
2     order-burger( $type, side => 'french fries' );
3 }
```

If your personal order is always vegetarian, you might instead wish for a `order-the-usual` sub. This is less concise to write, due to the optional second parameter:

```
1 sub order-the-usual ( $side? ) {
2     if ( $side.defined ) {
3         order-burger( 'veggie', $side );
4     }
5     else {
6         order-burger( 'veggie' );
7     }
8 }
```

Currying gives you a shortcut for these exact cases; it creates a new sub from an existing sub, with parameters already filled in. In Perl 6, curry with the `.assuming` method:

```
1 &order-the-usual      := &order-burger.assuming( 'veggie' );
2 &order-burger-and-fries := &order-burger.assuming( side => 'french fries' );
```

The new sub is like any other sub, and works with all the various parameter-passing schemes already described.

```
1 order-the-usual( 'salsa' );
2 order-the-usual( side => 'broccoli' );
3
4 order-burger-and-fries( 'plain' );
5 order-burger-and-fries( :type<<double-beef>> );
```

## 4.8 Introspection

Subroutines and their signatures are objects like any other. Besides calling them, you can learn things about them, including the details of their parameters:

```
1 sub logarithm(Numeric $x, Numeric :$base = exp(1)) {
2     log($x) / log($base);
```

## Chapter 4 | SUBS AND SIGNATURES

```
3 }
4
5 my @params = &logarithm.signature.params;
6 say @params.elems, ' parameters';
7
8 for @params {
9     say "Name:      ", .name;
10    say "  Type:      ", .type;
11    say "  named?     ", .named    ?? 'yes' !! 'no';
12    say "  slurpy?    ", .slurpy   ?? 'yes' !! 'no';
13    say "  optional? ", .optional  ?? 'yes' !! 'no';
14 }
```

```
2 parameters
Name:      $x
Type:      Numeric()
named?     no
slurpy?    no
optional?  no
Name:      $base
Type:      Numeric()
named?     yes
slurpy?    no
optional?  yes
```

The `&` sigil followed by a subroutine name gets the object representing that subroutine. `&logarithm.signature` returns the signature associated with the subroutine, and calling `.params` on the signature returns a list of `Parameter` objects. Each of these objects describes one parameter in detail.

# TODO: talk about `&signature.cando` once that's implemented

Signature introspection allows you to build interfaces that can obtain and then pass the right data to a subroutine. For example, you could build a web form generator that knew how to get input from a user, validate it, and then call a routine with it based upon the information obtained through introspection. A similar approach might generate a command line interface along with some basic usage instructions.



Table 4.2: Methods in the Parameter class

# stolen straight from S06, adapted a bit

method	description
name	The name of the lexical variable to bind to, if any
type	The nominal type
constraints	Any further type constraints
readonly	True if the parameter has is_readonly trait
rw	True if the parameter has is_rw trait
copy	True if the parameter has is_copy trait
named	True if the parameter is to be passed by name
named_names	List of names a named parameter can be passed as
slurpy	True if the parameter is slurpy
optional	True if the parameter is optional
default	A closure returning the default value
signature	A nested signature to bind the argument against

Beyond this, traits (traits) allow you to associate extra data with parameters. This metadata can go far beyond that which subroutines, signatures, and parameters normally provide.



# 5

## Multis

Perl usually decides which function to call based on the name of the function or the contents of a function reference. This is simple to understand. Perl can also examine the contents of the arguments provided to decide which of several variants of a function—variants each with the same name—to call. In this case, the amount and types of the function’s arguments help to distinguish between multiple variants of a function. This is *multidispatch*, and the functions to which Perl can dispatch in this case are *multis*.

Javascript Object Notation (*JSON*) is a simple data exchange format often used for communicating with web services. It supports arrays, hashes, numbers, strings, boolean values, and `null`, the undefined value.

`JSON::Tiny` is a minimal library used to convert Perl 6 data structures to JSON. See `grammars` for the other part of that module, which parses JSON and turns it into Perl 6 data structures. The full code, containing additional documentation and tests, is available from

<http://github.com/moritz/json/>. This snippet demonstrates how multis make the code simpler and more obvious:

```

1 multi to-json(Real $d) { ~$d }
2 multi to-json(Bool $d) { $d ?? 'true' !! 'false'; }
3 multi to-json(Str $d) {
4     ''
5     ~ $d.trans(['"', '\\', '\b', '\f', '\n', '\r', '\t']
6               => ['\"', '\\\\', '\\b', '\\f', '\\n', '\\r', '\\t'])
7     ~ ''
8 }
9
10 multi to-json(Array $d) {
11     return '['
12         ~ $d.values.map({ to-json($_) }).join(', ')
13         ~ ' ]';
14 }
15
16 multi to-json(Hash $d) {
17     return '{ '
18         ~ $d.pairs.map({ to-json(.key)
19                       ~ ' : '
20                       ~ to-json(.value) }).join(', ')
21         ~ ' }';
22 }
23
24 multi to-json($d where {!defined $d}) { 'null' }
25
26 multi to-json($d) {
27     die "Can't serialize an object of type " ~ $d.WHAT.perl
28 }

```

This code defines a single multi sub named `to-json`, which takes one argument and turns that into a string. `to-json` has many *candidates*; these subs all have the name `to-json` but differ in their signatures. Every candidate resembles:

```

1 multi to-json(Bool $data) { ... }
2 multi to-json(Real $data) { ... }

```

Which one is actually called depends on the type of the data passed to the subroutine. A call such as `to-json(Bool::True)` invokes the first candidate. Passing a numeric value of type `Real` instead invokes the second.

The candidate for handling `Real` is very simple; because JSON's and Perl 6's number formats coincide, the JSON converter can rely on Perl's conversion of these numbers to strings. The `Bool` candidate returns a literal string `'true'` or `'false'`.

The `Str` candidate does more work: it wraps its parameter in quotes and escapes literal characters that the JSON spec does not allow in strings—a tab character becomes `\t`, a newline `\n`, and so on.

The `to-json(Array $d)` candidate converts all elements of the array to JSON with recursive calls to `to-json`, joins them with commas, and surrounds them with square brackets. The recursive calls demonstrate a powerful truth of multidispatch: these calls do not necessarily recurse to the `Array` candidate, but dispatch to the appropriate candidate based on the types of *their* arguments.

The candidate that processes hashes turns them into the form `{ "key1" : "value1", "key2" : [ "second", "value" ] }`. It does this again by recursing into `to-json`.

## 5.1 Constraints

Candidates can specify more complex signatures:

```

1 multi to-json($d where {!defined $d}) { 'null' }

```

This candidate adds two new twists. It contains no type definition, in which case the type of the parameter defaults to `Any`, the root of the normal branch of the type hierarchy. More interestingly, the `where {!defined $d}` clause is a *constraint*, which defines a so-called *subset*

*type*. This candidate will match only *some* values of the type Any—those where the value is undefined.

Whenever the compiler performs a type check on the parameter \$d, it first checks the *nominal* type (here, Any). If that check succeeds, it calls the code block. The entire type check can only succeed if the code block returns a true value.

The curly braces for the constraint can contain arbitrary code. You can abuse this to count how often a type check occurs:

```

1  my $counter = 0;
2
3  multi a(Int $x) { };
4  multi a($x)     { }
5  multi a($x where { $counter++; True }) { };
6
7  a(3);
8  say $counter;      # says 0
9  a('str');
10 say $counter;      # says 2

```

This code defines three multis, one of which increases a counter whenever its where clause executes. Any Perl 6 compiler is free to optimize away type checks it knows will succeed. In the current Rakudo implementation, the second line with say will print a higher number than the first.

In the first call of a(3), the nominal types alone already determine the best candidate match, so the where block never executes and the first \$counter output is always 0.

The output after the second call is at least 1. The compiler has to execute the where-block at least once to check if the third candidate is the best match, but the specification does not require the *minimal* possible number of runs. This is illustrated in the second \$counter output. The specific implementation used to run this test actually executes the where-block twice. Keep in mind that the number of times the subtype checks blocks execute is specific to any particular implementation of Perl 6.

Avoid writing code like this in anything other than example code. Relying on the side effects of type checks produces unreliable code.

## 5.2 Narrowness

One candidate remains from the JSON example:

```
1 multi to-json($d) {  
2     die "Can't serialize an object of type " ~ $d.WHAT.perl  
3 }
```

With no explicit type or constraint on the parameter `$d`, its type defaults to `Any`—and thus it matches any passed object. The body of this function complains that it doesn't know what to do with the argument. This works for the example, because JSON's specification covers only a few basic structures.

The declaration and intent may seem simple at first, but look closer. This final candidate matches not only objects for which there is no candidate defined, but it can match for *all* objects, including `Int`, `Bool`, `Num`. A call like `to-json(2)` has *two* matching candidates—`Int` and `Any`.

If you run that code, you'll discover that the `Int` candidate gets called. Because `Int` is a type that conforms to `Any`, it is a *narrower* match for an integer. Given two types `A` and `B`, where `A` conforms to `B` (`A ~~ B`, in Perl 6 code), an object which conforms to `A` does so more narrowly than to `B`. In the case of multi dispatch, the narrowest match always wins.

A successfully evaluated constraint makes a match narrower than a similar signature without a constraint. In the case of:

```
1 multi to-json($d) { ... }  
2 multi to-json($d where {!defined $d}) { ... }
```

... an undefined value dispatches to the second candidate.

However, a matching constraint always contributes less to narrowness than a more specific match in the nominal type.

```

1  TODO: Better example
2
3  multi a(Any $x where { $x > 0 }) { 'Constraint' }
4  multi a(Int $x)                  { 'Nominal type' }
5
6  say a(3), ' wins';              # says Nominal type wins

```

This restriction allows a clever compiler optimization: it can sort all candidates by narrowness once to find the candidate with the best matching signature by examining nominal type constraints. These are far cheaper to check than constraint checks. Constraint checking occurs next, then the compiler considers the nominal types of candidates.

With some trickery it is possible to get an object which conforms to a built-in type (Num, for example) but which is also an undefined value. In this case the candidate that is specific to Num wins, because the nominal type check is narrower than the where {!defined \$d} constraint.

## 5.3 Multiple arguments

Candidate signatures may contain any number of positional and named arguments, both explicit and slurpy. However only positional parameters contribute to the narrowness of a match:

```

1  # RAKUDO has problems with an enum here,
2  # it answers with "Player One wins\nDraw\nDraw"
3  # using separate classes would fix that,
4  # but is not as pretty.
5  enum Symbol <Rock Paper Scissors>;

```



```

6 multi wins(Scissors $, Paper    $) { +1 }
7 multi wins(Paper    $, Rock     $) { +1 }
8 multi wins(Rock     $, Scissors $) { +1 }
9 multi wins(::T      $, T        $) { 0 }
10 multi wins(        $,          $) { -1 }
11
12 sub play($a, $b) {
13   given wins($a, $b) {
14     when +1 { say 'Player One wins' }
15     when 0 { say 'Draw' }
16     when -1 { say 'Player Two wins' }
17   }
18 }
19
20 play(Scissors, Paper);
21 play(Paper,    Paper);
22 play(Rock,     Paper);

```

	Rock	Paper	Scissors
Rock	0	-1	+1
Paper	+1	0	-1
Scissors	-1	+1	0

Figure 5.1: Who wins the *Rock, Paper, Scissors* game?

This example demonstrates how multiple dispatch can encapsulate all of the rules of a popular game. Both players independently select a symbol (rock, paper, or scissors). Scissors win against paper, paper wraps rock, and scissors can't cut rock, but go blunt trying. If both players select the same item, it's a draw.

The code creates a type for each possible symbol by declaring an enumerated type, or *enum*. For each combination of chosen symbols for which Player One wins there's a candidate of the form:

## Chapter 5 | MULTIS

```
1 multi wins(Scissors $, Paper $) { +1 }
```

Because the bodies of the subs here do not use the parameters, there's no reason to force the programmer to name them; they're *anonymous parameters*. A single \$ in a signature identifies an anonymous scalar variable.

The fourth candidate, `multi wins(::T $, T $) { 0 }` uses `::T`, which is a *type capture* (similar to *generics* or *templates* in other programming languages). It binds the nominal type of the first argument to `T`, which can then act as a type constraint. If you pass a `Rock` as the first argument, `T` acts as an alias for `Rock` inside the rest of the signature and the body of the routine. The signature `(::T $, T $)` will bind only two objects of the same type, or where the second is of a subtype of the first.

In this game, that fourth candidate matches only for two objects of the same type. The routine returns 0 to indicate a draw.

The final candidate is a fallback for the cases not covered yet—every case in which Player Two wins.

If the `(Scissors, Paper)` candidate matches the supplied argument list, it is two steps narrower than the `(Any, Any)` fallback, because both `Scissors` and `Paper` are direct subtypes of `Any`, so both contribute one step.

If the `(::T, T)` candidate matches, the type capture in the first parameter does not contribute any narrowness—it is not a constraint, after all. However `T` is a constraint for the second parameter which accounts for as many steps of narrowness as the number of inheritance steps between `T` and `Any`. Passing two `Rocks` means that `::T, T` is one step narrower than `Any, Any`. A possible candidate:

```
1 multi wins(Rock $, Rock $) {  
2     say "Two rocks? What is this, 20,000 years ago?"  
3 }
```

... would win against `(::T, T)`.

## 5.4 Bindability checks

Traits can apply *implicit constraints*:

```
1 multi swap($a is rw, $b is rw) {  
2     ($a, $b) = ($b, $a);  
3 }
```

This routine exchanges the contents of its two arguments. It must bind the two arguments as `rw`—both readable and writable. Calling the `swap` routine with an immutable value (for example a number literal) will fail.

The built-in function `substr` can not only extract parts of strings, but also modify them:

```
1 # substr(String, Start, Length)  
2 say substr('Perl 5', 0, 4);          # prints Perl  
3  
4 my $p = 'Perl 5';  
5 # substr(String, Start, Length, Substitution)  
6 substr($p, 6, 1, '6');  
7 # now $p contains the string Perl 6
```

You already know that the three-argument version and the four-argument version have different candidates: the latter binds its first argument as `rw`:

```
1 multi substr($str, $start = 0, $length = *) { ... }  
2 multi substr($str is rw, $start, $length, $substitution) { ... }
```

This is also an example of candidates with different *arity* (number of expected arguments). This is seldom really necessary, because it is often a better alternative to make parameters optional. Cases where an arbitrary number of arguments are allowed are handled with slurpy parameters instead:

```

1 sub mean(*@values) {
2     ([+] @values) / @values;
3 }

```

## 5.5 Nested Signatures in Multi-dispatch

An earlier chapter showed how to use nested signatures to look deeper into data structures and extract parts of them. In the context of multiple dispatch, nested signatures take on a second task: they act as constraints to distinguish between the candidates. This means that it is possible to dispatch based upon the shape of a data structure. This brings Perl 6 a lot of the expressive power provided by pattern matching in various functional languages.

Some algorithms have very tidy and natural expressions with this feature, especially those which recurse to a simple base case. Consider quicksort. The base case is that of the empty list, which trivially sorts to the empty list. A Perl 6 version might be:

```

1 multi quicksort([]) { () }

```

The `[]` declares an empty nested signature for the first positional parameter. Additionally, it requires that the first positional parameter be an indexable item—anything that would match the `@` sigil. The signature will only match if the multi has a single parameter which is an empty list.

The other case is a list which contains at least one value—the pivot—and possibly other values to partition according to the pivot. The rest of quicksort is a couple of recursive calls to sort both partitions:

```

1 multi quicksort([$pivot, *@rest]) {
2     my @before = @rest.grep({ $_ <= $pivot });
3     my @after  = @rest.grep({ $_ >  $pivot });
4
5     return quicksort(@before), $pivot, quicksort(@after);
6 }

```

## 5.6 Protos

You have two options to write multi subs: either you start every candidate with `multi sub ...` or `multi ...`, or you declare once and for all that the compiler shall view every sub of a given name as a multi candidate. Do the latter by installing a *proto* routine:

```
1 proto to-json($) { ... }      # literal ... here
2
3 # automatically a multi
4 sub to-json(Bool $d) { $d ?? 'true' !! 'false' }
```

Nearly all Perl 6 built-in functions and operators export a proto definition, which prevents accidental overriding of built-ins<sup>1</sup>.

To hide all candidates of a multi and replace them by another sub, declare it as `only sub YourSub`. At the time of writing, no compiler supports this.

## 5.7 Multi Methods

Methods can participate in dispatch just as do subroutines. For multi method dispatch the invocant acts as a positional parameter.

The main difference between sub and method calls is where the dispatcher searches for the routines: it looks for subroutines in the current and outer lexical scopes, whereas it looks for methods in the class of the invocant and recursively in any parent classes.

Unlike subroutine dispatch, you can dispatch to multiple candidates with multimethods. The `$object.?method` syntax dispatches to zero or one matching candidates; it is no error if there is no matching candidate. `$object.*method` calls *all* matching candidates, but it is

---

<sup>1</sup> One of the very rare exceptions is the smart match operator `infix:<~~>` which is not easily overloadable. Instead it redispaches to overloadable multi methods.

no error if there are no matching candidates. `$object.+method` calls at least one matching candidate.

## 5.8 Toying with the candidate list

Each multi dispatch builds a list of candidates, all of which satisfy the nominal type constraints. For a normal sub or method call, the dispatcher invokes the first candidate which passes any additional constraint checks.

A routine can choose to delegate its work to other candidates in that list. The `callsame` primitive calls the next candidate, passing along the arguments received. The `callwith` primitive calls the next candidate with different (and provided) arguments. After the called routine has done its work, the callee can continue its work.

If there's no further work to do, the routine can decide to hand control completely to the next candidate by calling `nextsame` or `nextwith`. The former reuses the argument list and the latter allows the use of a different argument list. This delegation is common in object destructors, where each subclass may perform some cleanup for its own particular data. After it finishes its work, it can delegate to its parent class `meethod` by calling `nextsame`.

# 6

## Classes and Objects

TODO: start with a much simpler bare-bones example!

The following program shows how a dependency handler might look in Perl 6. It showcases custom constructors, private and public attributes, methods and various aspects of signatures. It's not very much code, and yet the result is interesting and, at times, useful.

```
1 class Task {
2     has      &!callback;
3     has Task @!dependencies;
4     has Bool $.done;
5
6     # RAKUDO: Should really be '&callback' [perl #69766]
7     method new(Callable $callback, Task *@dependencies) {
8         return self.bless(*, :$callback, :@dependencies);
9     }
10 }
```

```

11     method add-dependency(Task $dependency) {
12         push @!dependencies, $dependency;
13     }
14
15     method perform() {
16         unless $!done {
17             .perform() for @!dependencies;
18             &!callback();
19             $!done = True;
20         }
21     }
22 }
23
24 my $eat =
25     Task.new({ say 'eating dinner. NOM!' },
26     Task.new({ say 'making dinner' },
27     Task.new({ say 'buying food' },
28     Task.new({ say 'making some money' })),
29     Task.new({ say 'going to the store' })
30     ),
31     Task.new({ say 'cleaning kitchen' })
32     )
33 );
34
35 $eat.perform();

```

## 6.1 Starting with class

Perl 6, like many other languages, uses the `class` keyword to introduce a new class. Anything inside of the block that follows is part of the class definition. You may place arbitrary code there, just as you can with any other block, but classes commonly contain declarations. The example code includes declarations relating to state (attributes, introduced through the `has` keyword) and behavior (methods, through the `method` keyword).



Declaring a class creates a *type object*, which by default gets installed into the package (just like a variable declared with our scope). This type object is an “empty instance” of the class. You’ve already seen these. For example, types such as `Int` and `Str` refer to the type object of one of the Perl 6 built-in classes. The example uses the class name `Task` so that other code can refer to it later, such as to create class instances by calling the `new` method.

## 6.2 I can has state?

The first three lines inside the class block all declare attributes (called *fields* or *instance storage* in other languages). These are storage locations that every instance of a class gets. Just as a `my` variable can not be accessed from the outside of its declared scope, attributes are not accessible outside of the class. This *encapsulation* is one of the key principles of object oriented design.

The first declaration specifies instance storage for a callback – a bit of code to invoke in order to perform the task that an object represents:

```
1 has &!callback;
```

The `&` sigil indicates that this attribute represents something invocable. The `!` character is a *twigil*, or secondary sigil. A twigil forms part of the name of the variable. In this case, the `!` twigil emphasizes that this attribute is private to the class.

The second declaration also uses the private twigil:

```
1 has Task @!dependencies;
```

However, this attribute represents an array of items, so it requires the `@` sigil. These items each specify a task that must be completed before the present one can complete. Furthermore, the type declaration on this attribute indicates that the array may only hold instances of the `Task` class (or some subclass of it).

The third attribute represents the state of completion of a task:

## Chapter 6 | CLASSES AND OBJECTS

```
1 has Bool $.done;
```

This scalar attribute (with the \$ sigil) has a type of Bool. Instead of the ! twigil, this twigil is .. While Perl 6 does enforce encapsulation on attributes, it also saves you from writing accessor methods. Replacing the ! with a . both declares the attribute \$!done and an accessor method named done. It's as if you had written:

```
1 has Bool $!done;  
2 method done() { return $!done }
```

Note that this is not like declaring a public attribute, as some languages allow; you really get *both* a private storage location and a method, without having to write the method by hand. You are free instead to write your own accessor method, if at some future point you need to do something more complex than return the value.

Note that using the . twigil has created a method that will provide with readonly access to the attribute. If instead the users of this object should be able to reset a task's completion state (perhaps to perform it again), you can change the attribute declaration:

```
1 has Bool $.done is rw;
```

The is rw trait causes the generated accessor method to return something external code can modify to change the value of the attribute.

### 6.3 Methods

While attributes give objects state, methods give objects behaviors. Ignore the new method temporarily; it's a special type of method. Consider the second method, add-dependency, which adds a new task to this task's dependency list.

```
1 method add-dependency(Task $dependency) {  
2     push @!dependencies, $dependency;  
3 }
```

In many ways, this looks a lot like a sub declaration. However, there are two important differences. First, declaring this routine as a method adds it to the list of methods for the current class. Thus any instance of the Task class can call this method with the `.` method call operator. Second, a method places its invocant into the special variable `self`.

The method itself takes the passed parameter—which must be an instance of the Task class—and pushes it onto the invocant's `@!dependencies` attribute.

The second method contains the main logic of the dependency handler:

```
1  method perform() {  
2      unless $!done {  
3          .perform() for @!dependencies;  
4          &!callback();  
5          $!done = True;  
6      }  
7  }
```

It takes no parameters, working instead with the object's attributes. First, it checks if the task has already completed by checking the `$!done` attribute. If so, there's nothing to do.

Otherwise, the method performs all of the task's dependencies, using the `for` construct to iterate over all of the items in the `@!dependencies` attribute. This iteration places each item—each a Task object—into the topic variable, `$.` Using the `.` method call operator without specifying an explicit invocant uses the current topic as the invocant. Thus the iteration construct calls the `.perform()` method on every Task object in the `@!dependencies` attribute of the current invocant.

After all of the dependencies have completed, it's time to perform the current Task's task by invoking the `&!callback` attribute directly; this is the purpose of the parentheses. Finally, the method sets the `$!done` attribute to `True`, so that subsequent invocations of `perform` on this object (if this Task is a dependency of another Task, for example) will not repeat the task.

## 6.4 Constructors

Perl 6 is rather more liberal than many languages in the area of constructors. A constructor is anything that returns an instance of the class. Furthermore, constructors are ordinary methods. You inherit a default constructor named `new` from the base class `Object`, but you are free to override `new`, as this example does:

```
1 # RAKUDO: Should really be '&callback' [perl #69766]
2 method new(Callable $callback, Task *@dependencies) {
3     return self.bless(*, :$callback, :@dependencies);
4 }
```

The biggest difference between constructors in Perl 6 and constructors in languages such as C# and Java is that rather than setting up state on a somehow already magically created object, Perl 6 constructors actually create the object themselves. The easiest way to do this is by calling the `bless` method, also inherited from `Object`. The `bless` method expects a positional parameter—the so-called “candidate”—and a set of named parameters providing the initial values for each attribute.

The example’s constructor turns positional arguments into named arguments, so that the class can provide a nice constructor for its users. The first parameter is the callback (the thing to do to execute the task). The rest of the parameters are dependent Task instances. The constructor captures these into the `@dependencies` slurpy array and passes them as named parameters to `bless` (note that `:$callback` uses the name of the variable—minus the sigil—as the name of the parameter).

## 6.5 Consuming our class

After creating a class, you can create instances of the class. Declaring a custom constructor provides a simple way of declaring tasks along with their dependencies. To create a single task with no dependencies, write:

```
1 my $eat = Task.new({ say 'eating dinner. NOM!' });
```

An earlier section explained that declaring the class `Task` installed a type object had been installed in the namespace. This type object is a kind of “empty instance” of the class, specifically an instance without any state. You can call methods on that instance, as long as they do not try to access any state; `new` is an example, as it creates a new object rather than modifying or accessing an existing object.

Unfortunately, dinner never magically happens. It has dependent tasks:

```
1 my $eat =  
2     Task.new({ say 'eating dinner. NOM!' },  
3         Task.new({ say 'making dinner' },  
4             Task.new({ say 'buying food' },  
5                 Task.new({ say 'making some money' }),  
6                 Task.new({ say 'going to the store' })  
7             ),  
8         Task.new({ say 'cleaning kitchen' })  
9     )  
10 );
```

Notice how the custom constructor and sensible use of whitespace allows a layout which makes task dependencies clear.

Finally, the `perform` method call recursively calls the `perform` method on the various other dependencies in order, giving the output:

```
1 making some money  
2 going to the store  
3 buying food  
4 cleaning kitchen  
5 making dinner  
6 eating dinner. NOM!
```

## 6.6 Exercises

1. The method `add-dependency` in `Task` permits the creation of *cycles* in the dependency graph. That is, if you follow dependencies, you can eventually return to the original `Task`. Show how to create a graph with cycles and explain why the `perform` method of a `Task` whose dependencies contain a cycle would never terminate successfully.

**Answer:** You can create two tasks, and then “short-circuit” them with `add-dependency`:

```
1 my $a = Task.new({ say 'A' });
2 my $b = Task.new({ say 'B' }, $a);
3 $a.add-dependency($b);
```

The `perform` method will never terminate because the first thing the method does is to call all the `perform` methods of its dependencies. Because `$a` and `$b` are dependencies of each other, none of them would ever get around to calling their callbacks. The program will exhaust memory before it ever prints 'A' or 'B'.

2. Is there a way to detect the presence of a cycle during the course of a `perform` call? Is there a way to prevent cycles from ever forming through `add-dependency`?

**Answer:** To detect the presence of a cycle during a `perform` call, keep track of which `Tasks` have started; prevent a `Task` from starting twice before finishing:

```
1 augment class Task {
2     has Bool $!started = False;
3
4     method perform() {
5         if $!started++ && !$!done {
6             die "Cycle detected, aborting";
7         }
8
9         unless $!done {
10             .perform() for @!dependencies;
11             &!callback();
12         }
13     }
14 }
```

```

12         $!done = True;
13     }
14 }
15 }

```

Another approach is to stop cycles from forming during add-dependency by checking whether there's already a dependency running in the other direction. (This is the only situation in which a cycle can occur.) This requires the addition of a helper method `depends-on`, which checks whether a task depends on another one, either directly or transitively. Note the use of `»` and `[||]` to write succinctly what would otherwise have involved looping over all the dependencies of the Task:

```

1  augment class Task {
2      method depends-on(Task $some-task) {
3          $some-task === any(@!dependencies)
4          [||] @!dependencies».depends-on($some-task)
5      }
6
7      method add-dependency(Task $dependency) {
8          if $dependency.depends-on(self) {
9              warn 'Cannot add that task, since it would introduce a cycle.';
10             return;
11         }
12         push @!dependencies, $dependency;
13     }
14 }

```

3. How could Task objects execute their dependencies in parallel? (Think especially about how to avoid collisions in “diamond dependencies”, where a Task has two different dependencies which in turn have the same dependency.)

**Answer:** Enabling parallelism is easy; change the line `.perform()` for `@!dependencies;` into `@!dependencies».perform()`. However, there may be race conditions in the case of diamond dependencies, wherein Tasks A starts B and C in parallel, and both start a copy of D, making D run twice. The solution to this is the same as with the cycle-detection in Question 2: introducing an attribute `$!started`. Note that it's impolite to die if a Task has

## Chapter 6 | CLASSES AND OBJECTS

started but not yet finished, because this time it might be due to parallelism rather than cycles:

```
1  augment class Task {
2      has Bool $!started = False;
3
4      method perform() {
5          unless $!started++ {
6              @!dependencies».perform();
7              &!callback();
8              $!done = True;
9          }
10     }
11 }
```



# 7

## Roles

A *role* is a standalone, named, reusable unit of behavior. You can compose a role into a class at compile time or add it to an individual object at runtime.

That's an abstract definition best explained by an example. This program demonstrates a simple and pluggable IRC bot framework which understands a few simple commands.

```
1 # XXX This is VERY preliminary code and needs filling out. But it
2 # does provide opportunities to discuss runtime mixins, compile time
3 # composition, requirements and a few other bits.
4
5 my regex nick { \w+ }
6 my regex join-line { ... <nick> ... }
7 my regex message-line { $<sender>=[...] $<message>=[...] }
8
9 class IRCBot {
10     has $.bot-nick;
```

## Chapter 7 | ROLES

```
11     method run($server) {
12         ...
13     }
14 }
15
16 role KarmaTracking {
17     has %!karma-scores;
18
19     multi method on-message($sender, $msg where /^karma <ws> <nick>/) {
20         if %!karma-scores{$<nick>} -> $karma {
21             return $<nick> ~ " has karma $karma";
22         }
23         else {
24             return $<nick> ~ " has neutral karma";
25         }
26     }
27
28     multi method on-message($sender, $msg where /<nick> '++'/) {
29         %!karma-scores{$<nick>}++;
30     }
31
32     multi method on-message($sender, $msg where /<nick> '--'/) {
33         %!karma-scores{$<nick>}--;
34     }
35 }
36
37 role Oping {
38     has @!whoz-op;
39
40     multi method on-join($nick) {
41         if $nick eq any(@!whoz-op) {
42             return "/mode +o $nick";
43         }
44     }
45
46     multi method on-message($sender, $msg where /^trust <ws> <nick>/) {
47         if $sender eq any(@!whoz-op) {
48             push @!whoz-op, $<nick>;
```

```

49         return "I now trust " ~ $<nick>;
50     }
51     else {
52         return "But $sender, I don't trust you";
53     }
54 }
55 }
56
57 role AnswerToAll {
58     method process($raw-in) {
59         if $raw-in ~~ /<on-join>/ {
60             self.*on-join($<nick>);
61         }
62         elsif $raw-in ~~ /<on-message>/ {
63             self.*on-message($<sender>, $<message>)
64         }
65     }
66 }
67
68 role AnswerIfTalkedTo {
69     method bot-nick() { ... }
70
71     method process($raw-in) {
72         if $raw-in ~~ /<on-join>/ {
73             self.*on-join($<nick>);
74         }
75         elsif $raw-in ~~ /<on-message>/ -> $msg {
76             my $my-nick = self.bot-nick();
77             if $msg<msg> ~~ /^ $my-nick ':'/ {
78                 self.*on-message($msg<sender>, $msg<message>)
79             }
80         }
81     }
82 }
83
84 my %pluggables =
85     karma => KarmaTracking,
86     op => Oping;

```

```

87
88  role Plugins {
89      multi method on-message($self is rw: $sender, $msg where /^youdo <ws> (\w+)/) {
90          if %pluggables{$0} -> $plug-in {
91              $self does $plug-in;
92              return "Loaded $0";
93          }
94      }
95  }
96
97  class KarmaKeeper is IRCBot does AnswerToAll does KarmaTracking {
98  }
99
100 class NothingBot is IRCBot does AnswerIfTalkedTo does Plugins {
101 }

```

## 7.1 What is a role?

Previous chapters have explained classes and grammars. A role is another type of package. Like classes and grammars, a role can contain methods (including named regexes) and attributes. However, a role cannot stand on its own. To use a role, you must incorporate it into an object, a class, or a grammar.

In traditional object orientation, classes perform two tasks. They represent entities in the system by providing models from which to create instances. They also provide a mechanism for code re-use. These two tasks are somewhat in tension. For optimal re-use, classes should be small, but in order to represent a complex entity with many behaviors, classes tend to grow large.

Perl 6 classes retain the responsibility for instance management. Roles handle the task of code reuse. A role contains the methods and attributes required to provide a named, reusable unit of behavior. Building a class out of roles uses a safe mechanism called *flattening composition*. You may also apply a role to an individual object. Both of these design techniques appear in the example code.

Roles may also support parameters. This helps Perl 6 provide generic programming, along the lines of generics in C# and Java, or templates in C++.

## 7.2 Compile Time Composition

Take a look at the `KarmaKeeper` class definition. The body is empty; the class defines no attributes or methods of its own. The class inherits from `IRCBot`, using the `is` trait modifier – something familiar from earlier chapters – but it also uses the `does` trait modifier to compose two roles into the class.

The process of role composition is simple. Perl takes the attributes and methods defined in each role and copies them into the class. After composition, the class appears as if it had defined those behaviors itself. This is part of the flattening property: after composing a role into the class, the roles in and of themselves are only important when querying the class to determine *if* it performs the role. Querying the methods of the `KarmaKeeper` class through introspection will report that the class has both a `process` method and an `on-message` multi method.

If this were all that roles provided, they'd have few advantages over inheritance or mixins. Roles get much more interesting in the case of a conflict. Consider the class definition:

```
1 class MyBot is IRCBot does AnswerToAll does AnswerIfTalkedTo {  
2 }
```

Both the `AnswerToAll` and `AnswerIfTalkedTo` roles provide a method named `process`. Even though they share a name, the methods perform semantically different behaviors—behaviors which conflict. The role composer will produce a compile-time error about this conflict, asking the programmer to provide a resolution.

Multiple inheritance and mixin mechanisms rarely provide this degree of conflict resolution. In those situations, the order of inheritance or mixin decides which method wins. All possible roles are equal in role composition.

What can you do if there is a conflict? In this case, it makes little sense to compose both of the roles into a class. The programmer here has made a mistake and should choose to compose only one role to provide the desired behavior. An alternative way to resolve a conflict is to write a method with the same name in the class body itself:

```

1  class MyBot is IRCBot does AnswerToAll does AnswerIfTalkedTo {
2      method process($raw-in) {
3          # Do something sensible here...
4      }
5  }
```

If the role composer detects a method with the same name in the class body, it will then disregard all of the (possibly conflicting) ones from the roles. Put simply, methods in the class always win.

### 7.2.1 Multi-methods and composition

Sometimes it's okay to have multiple methods of the same name, provided they have different signatures such that the multidispatch mechanism can distinguish between them. Multi methods with the same name from different roles will not be in conflict. Instead, the candidates from all of the roles will combine together during composition into the class.

If the class provides a method of the same name that is also multi, then the multi candidates from the class will be included. On the other hand, if the class has a method of the same name that is *not* declared as a multi, then the method in the class alone—as usual—will win.

This is the mechanism by which a class that composes both, for example, the KarmaTracking role and the Oping role would end up having the candidates that both roles provide for the on-message method. As a class ends up composing more than a couple of roles, it may be preferable to use an alternative syntax that allows you to list the roles in the class body:

```

1  class KarmaKeeper is IRCBot {
2      does AnswerToAll;
3      does KarmaTracking;
```

```
4     does Oping;
5 }
```

## 7.2.2 Calling all candidates

The process methods of the roles `AnswerToAll` and `AnswerIfTalkedTo` use a modified syntax for calling methods:

```
1 self.*on-message($msg<sender>, $msg<message>)
```

The use of `.*` instead of `.` changes the semantics of the method dispatch. Just as the `*` quantifier in regexes means “zero or more”, the `.*` dispatch operator will call zero or more matching methods. If no `on-message` multi candidates match, the call will not produce an error. If more than one `on-message` multi candidate matches, Perl will call all of them, whether found by multiple dispatch, searching the inheritance hierarchy, or both.

There are two other variants. `.+` greedily calls all methods but dies if there is not at least one method that matches. `.*?`, which tries to call one method, but returns a `Failure` rather than throwing an exception. The zero-or-more and one-or-more dispatch forms are somewhat rare, but very useful in event driven programming. One-or-failure is very useful when dealing with per-object role application.

## 7.2.3 Expressing requirements

The role `AnswerIfTalkedTo` declares a stub for the method `bot-nick`, but never gives an implementation.

```
1 method bot-nick() { ... }
```

In the context of a role, this technique declares that any class which composes this role must somehow provide a method named `bot-nick`. The class itself may provide it, another

role must provide it, or a parent class must provide it. IRCBot does the latter; it IRCBot defines an attribute `$!bot-nick` along with an accessor method.

You are not *required* to make explicit the methods on which your role depends. If you do so, the role composer can detect any errors at compile time. If you do not, the problem will only appear at runtime, when and if something attempts to call the missing method.

## 7.3 Runtime Application of Roles

Class-based OO usually works, where instances have the methods and attributes of the classes they instantiate. Sometimes, something more dynamic is useful. Perl 6 allows you to add extra behaviors to individual objects by applying roles to individual objects at runtime.

The example in this chapter uses this in order to give bots new abilities during its lifetime. The `Plugins` role is at the heart of this. Note the signature of the method `on-message`. It captures the invocant into a variable `$self` marked `rw`, which indicates that the invocant may be modified. Inside the method, that happens:

```
1  if %pluggables{$0} -> $plug-in {
2      $self does $plug-in;
3      return "Loaded $0";
4  }
```

Like classes, roles are first-class in Perl 6; you can pass them around just like any other object. The `%pluggables` hash maps names of plug-ins to Role objects. Thus this lookup stores a Role in `$plug-in`. The `does` operator adds this role to `$self`—not the *class* of `$self`, but the instance itself. From this point on, `$self` now has all of the methods from the role, in addition to all of the ones that it had before. Importantly, this does not have any influence on any other instances of the same class; only this one instance has changed.



### 7.3.1 Differences from compile time composition

Runtime application differs from compile time composition in that methods in the applied role in will automatically override any of the same name within the class of the object. It's as if you had written an anonymous subclass of the current class of the object that composed the role into it. This means that `. *` will find both those methods that mixed into the object from one or more roles along with any that already existed in the class.

If you wish to apply multiple roles at a time, list them all with `does`. This case behaves the same way as compile-time composition, in that the role composer will compose them all into the imaginary anonymous subclass. As you might expect, any conflicts will occur at this point.

This gives a degree of safety, but it happens at runtime and is thus not as safe as compile time composition. For safety, perform your compositions at compile time. For example, instead of applying multiple roles to an instance, compose them into a new role at compile time and apply that role to the instance.

### 7.3.2 The `but` operator

Runtime role application with `does` modifies an object in-place: `$x does SomeRole` modifies the object stored in `$x`. When this is not desirable, the `but` operator can be used instead. It returns a clone of the object, and applies the role composition only to the clone – the original object stays the same.

TODO: example

## 7.4 Parametric Roles

# 8

## Subtypes

```
1  enum Suit <spades hearts diamonds clubs>;
2  enum Rank (2, 3, 4, 5, 6, 7, 8, 9, 10,
3      'jack', 'queen', 'king', 'ace');
4
5  class Card {
6      has Suit $.suit;
7      has Rank $.rank;
8
9      method Str {
10         $.rank.name ~ ' of ' ~ $.suit.name;
11     }
12 }
13
14 subset PokerHand of List where { .elems == 5 && all(|$_) ~~ Card }
15
16 sub n-of-a-kind($n, @cards) {
17     for @cards>>.rank.uniq -> $rank {
```

## Chapter 8 | SUBTYPES

```
18         return True if $n == grep $rank, @cards>>.rank;
19     }
20     return False;
21 }
22
23 subset Quad          of PokerHand where { n-of-a-kind(4, $_) }
24 subset ThreeOfAKind of PokerHand where { n-of-a-kind(3, $_) }
25 subset OnePair      of PokerHand where { n-of-a-kind(2, $_) }
26
27 subset FullHouse of PokerHand where OnePair & ThreeOfAKind;
28
29 subset Flush of PokerHand where -> @cards { [==] @cards>>.suit }
30
31 subset Straight of PokerHand where sub (@cards) {
32     my @sorted-cards = @cards.sort({ .rank });
33     my ($head, @tail) = @sorted-cards;
34     for @tail -> $card {
35         return False if $card.rank != $head.rank + 1;
36         $head = $card;
37     }
38     return True;
39 }
40
41 subset StraightFlush of Flush where Straight;
42
43 subset TwoPair of PokerHand where sub (@cards) {
44     my $pairs = 0;
45     for @cards>>.rank.uniq -> $rank {
46         ++$pairs if 2 == grep $rank, @cards>>.rank;
47     }
48     return $pairs == 2;
49 }
50
51 sub classify(PokerHand $_) {
52     when StraightFlush { 'straight flush', 8 }
53     when Quad          { 'four of a kind', 7 }
54     when FullHouse     { 'full house',    6 }
55     when Flush         { 'flush',        5 }
```

```

56     when Straight      { 'straight',      4 }
57     when ThreeOfAKind { 'three of a kind', 3 }
58     when TwoPair      { 'two pair',       2 }
59     when OnePair      { 'one pair',       1 }
60     when *             { 'high cards',    0 }
61 }
62
63 my @deck = map -> $suit, $rank { Card.new(:$suit, :$rank) },
64     (Suit.pick(*) X Rank.pick(*));
65
66 @deck .= pick(*);
67
68 my @hand1;
69 @hand1.push(@deck.shift()) for ^5;
70 my @hand2;
71 @hand2.push(@deck.shift()) for ^5;
72
73 say 'Hand 1: ', map { "\n $_" }, @hand1>>.Str;
74 say 'Hand 2: ', map { "\n $_" }, @hand2>>.Str;
75
76 my ($hand1-description, $hand1-value) = classify(@hand1);
77 my ($hand2-description, $hand2-value) = classify(@hand2);
78
79 say sprintf q[The first hand is a '%s' and the second one a '%s', so %s.],
80     $hand1-description, $hand2-description,
81     $hand1-value > $hand2-value
82     ?? 'the first hand wins'
83     !! $hand2-value > $hand1-value
84     ?? 'the second hand wins'
85     !! "the hands are of equal value"; # XXX: this is wrong

```



# 9

## Pattern matching

Regular expressions are a concept from computer science where simple patterns are used to describe the format of text. Pattern matching is applying these patterns to actual strings to see if they ... well, match. Most modern regular expression facilities are more powerful than traditional regular expressions due to the influence of languages such as Perl, but the short-hand term `regex` has stuck and continues to mean “regular expression like pattern matching”. In Perl 6, though the specific syntax used to describe the patterns is different from PCRE<sup>1</sup> and POSIX<sup>2</sup>, we continue to call them `regex`.

A common writing error is to duplicate a word by accident. It is hard to catch such errors by rereading your own text, but Perl can do it for you using `regex`:

```
1 my $s = 'the quick brown fox jumped over the the lazy dog';  
2
```

---

<sup>1</sup> Perl Compatible Regular Expressions

<sup>2</sup> Portable Operating System Interface for Unix. See IEEE standard 1003.1-2001

## Chapter 9 | PATTERN MATCHING

```
3 if $s ~~ m/ « (\w+) \W+ $0 » / {  
4     say "Found '$0' twice in a row";  
5 }
```

In the simplest case a regex consists of a constant string. Matching a string against that regex searches for that string:

```
1 if 'properly' ~~ m/ perl / {  
2     say "'properly' contains 'perl'";  
3 }
```

The construct `m/ ... /` builds a regex. A regex on the right hand side of the `~~` smart match operator applies against the string on the left hand side. By default, whitespace inside the regex is irrelevant for the matching, so writing the regex as `m/ perl /`, `m/perl/` or `m/ p e r l/` all produce the exact same semantics—although the first way is probably the most readable one.

Only word characters, digits, and the underscore cause an exact substring search. All other characters may have a special meaning. If you want to search for a comma, an asterisk, or another non-word character, you must quote or escape it<sup>3</sup>:

```
1 my $str = "I'm *very* happy";  
2  
3 # quoting  
4 if $str ~~ m/ '*very*' / { say '\o/' }  
5 # escaping  
6 if $str ~~ m/ \* very \* / { say '\o/' }
```

However searching for literal strings gets boring pretty quickly. Regex support special (also called *metasyntactic*) characters. The dot (`.`) matches a single, arbitrary character:

```
1 my @words = <spell superlative openly stuff>;  
2 for @words -> $w {
```

---

<sup>3</sup> If you're just searching for literal text and not actually utilizing the pattern matching features of regex, consider using the `index` or `rindex` subroutines instead.



```

3      if $w =~ m/ pe.l / {
4          say "$w contains $/";
5      } else {
6          say "no match for $w";
7      }
8  }

```

This prints

```

spell contains pell
superlative contains perl
openly contains penl
no match for stuff

```

The dot matched an l, r, and n, but it would also match a space in the sentence *the spectroscope lacks resolution*—regexes don’t care about word boundaries at all. The special variable `$/` stores (among other things) only the part of the string that matched the regular expression. `$/` holds the so-called *match object*.

Suppose you have a big chunk of text. For solving a crossword puzzle you are looking for words containing pe, then an arbitrary letter, and then an l (but not a space, as your puzzle has extra markers for those). The appropriate regex for that is `m/pe \w l/`. The `\w` control sequence stands for a “Word” character—a letter, digit, or an underscore. In the example at the beginning of this chapter `\w` is used to build the definition of a “word”.

Several other common control sequences each match a single character:

Each of these backslash sequence means the complete opposite if you convert the letter to upper case: `\W` matches a character that’s not a word character and `\N` matches a single character that’s not a newline.

These matches are not limited to the ASCII range—`\d` matches Latin, Arabic-Indic, Devanagari and other digits, `\s` matches non-breaking whitespace and so on. These *character classes* follow the Unicode definition of what is a letter, a number, and so on. Define custom character classes by listing them inside nested angle and square brackets `<[ ... ]>`.

Table 9.1: Backslash sequences and their meaning

Symbol	Description	Examples
\w	word character	l, ö, 3, _
\d	digit	0, 1
\s	whitespace	(tab), (blank), (newline)
\t	tabulator	(tab)
\n	newline	(newline)
\h	horizontal whitespace	(space), (tab)
\v	vertical whitespace	(newline), (vertical tab)

```

1  if $str ~~ / <[aeiou]> / {
2      say "'$str' contains a vowel";
3  }
4
5  # negation with a -
6  if $str ~~ / <-[aeiou]> / {
7      say "'$str' contains something that's not a vowel";
8  }

```

Rather than listing each character in the character class individually, you may specify a range of characters by placing the range operator `..` between the character that starts the range and the character that ends the range:

```

1  # match a, b, c, d, ..., y, z
2  if $str ~~ / <[a..z]> / {
3      say "'$str' contains a lower case Latin letter";
4  }

```

Added to or subtract from character classes with the `+` and `-` operators:

```

1  if $str ~~ / <[a..z]+[0..9]> / {
2      say "'$str' contains a letter or number";
3  }

```

```

4
5 if $str =~ / <[a..z]-[aeiou]> / {
6     say "'$str' contains a consonant";
7 }

```

The negated character class is a special application of this idea.

A *quantifier* can specify how often something has to occur. A question mark `?` makes the preceding unit (be it a letter, a character class, or something more complicated) optional, meaning it can either be present either zero or one times in the string being matched. So `m/ho u? se/` matches either house or hose. You can also write the regex as `m/hou?se/` without any spaces, and the `?` still quantifies only the `u`.

The asterisk `*` stands for zero or more occurrences, so `m/z\w*o/` can match `zo`, `zoo`, `zero` and so on. The plus `+` stands for one or more occurrences, `\w+ usually` matches what you might consider a word (though only matches the first three characters from `isn't` because `'` isn't a word character).

The most general quantifier is `**`. If followed by a number it matches that many times, and if followed by a range, it can match any number of times that the range allows:

```

1 # match a date of the form 2009-10-24:
2 m/ \d**4 '-' \d\d '-' \d\d /
3
4 # match at least three 'a's in a row:
5 m/ a ** 3..* /

```

If the right hand side is neither a number nor a range, it becomes a delimiter, which means that `m/ \w ** ', '/` matches a list of characters separated by a comma and a whitespace each.

If a quantifier has several ways to match, Perl will choose the longest one. This is *greedy* matching. Appending a question mark to a quantifier makes it non-greedy<sup>45</sup>:

---

<sup>4</sup> The non-greedy general quantifier is `$thing **? $count`, so the question mark goes directly after the second asterisk.

<sup>5</sup> This example is a very poor way to parse HTML; using a proper parser is always preferable.

## Chapter 9 | PATTERN MATCHING

```
1 my $html = '<p>A paragraph</p> <p>And a second one</p>';
2 if $html =~ m/ '<p>' .* '</p>' / {
3     say 'Matches the complete string!';
4 }
5
6 if $html =~ m/ '<p>' .*? '</p>' / {
7     say 'Matches only <p>A paragraph</p>!';
8 }
```

To apply a modifier to more than just one character or character class, group items with square brackets:

```
1 my $ingredients = 'milk, flour, eggs and sugar';
2 # prints "milk, flour, eggs"
3 $ingredients =~ m/ [\w+] ** [\s*,\s*] / && say $/;
```

Separate *alternations*—tokens and units of which *any* can match— with vertical bars. One vertical bar between two parts of a regex means that the longest alternative wins. Two bars make the first matching alternative win.

```
1 $string =~ m/ \d**4 '-' \d\d '-' \d\d | 'today' | 'yesterday' /
```

### 9.1 Anchors

So far every regex could match anywhere within a string. Often it is desirable to limit the match to the start or end of a string or line, or to word boundaries.

A single caret `^` anchors the regex to the start of the string, a dollar `$` to the end. `m/ ^a /` matches strings beginning with an `a`, and `m/ ^ a $ /` matches strings that consist only of an `a`.

Table 9.2: Regex anchors

Anchor	Meaning
<code>^</code>	start of string
<code>\$</code>	end of string
<code>^^</code>	start of a line
<code>\$\$</code>	end of a line
<code>&lt;&lt;</code>	left word boundary
<code>«</code>	left word boundary
<code>&gt;&gt;</code>	right word boundary
<code>»</code>	right word boundary

## 9.2 Captures

Regexes are useful to check if a string is in a certain format, and to search for patterns within a string. With some more features they can be very good for *extracting* information too.

Surrounding part of a regex with round brackets (aka parentheses) `(...)` makes Perl *capture* the string it matches. The string matched by the first group of parentheses is available in `$/[0]`, the second in `$/[1]`, etc. `$/` acts as an array containing the captures from each parentheses group.

```

1 my $str = 'Germany was reunited on 1990-10-03, peacefully';
2 if $str =~ m/ (\d**4) \- (\d\d) \- (\d\d) / {
3     say 'Year: ', $/[0];
4     say 'Month: ', $/[1];
5     say 'Day: ', $/[2];
6     # usage as an array:
7     say $/.join('-');      # prints 1990-10-03
8 }

```

If you quantify a capture, the corresponding entry in the match object is a list of other match objects:

## Chapter 9 | PATTERN MATCHING

```
1 my $ingredients = 'eggs, milk, sugar and flour';
2 if $ingredients =~ m/(\w+) ** [\s,\s*] \s* 'and' \s* (\w+)/ {
3     say 'list: ', $/[0].join(' | ');
4     say 'end: ', $/[1];
5 }
```

This prints

```
list: eggs | milk | sugar
end: flour
```

The first capture, `(\w+)`, was quantified, and thus `$/[0]` is a list on which the code calls the `.join` method. Regardless of how many times the first capture matches, the second is still available in `$/[1]`.

As a shortcut, `$/[0]` is also available under the name `$0`, `$/[1]` as `$1`, and so on. These aliases are also available inside the regex. This allows you to write a regex that detects that common error of duplicated words, just like the example at the beginning of this chapter:

```
1 my $s = 'the quick brown fox jumped over the the lazy dog';
2
3 if $s =~ m/ « (\w+) \W+ $0 » / {
4     say "Found '$0' twice in a row";
5 }
```

The regex first anchors to a left word boundary with `«` so that it doesn't match partial duplication of words. Next, the regex captures a word `((\w+))`, followed by at least one non-word character `\W+`. This implies a right word boundary, so there is no need to use an explicit boundary. Then it matches the previous capture followed by a right word boundary.

Without the first word boundary anchor, the regex would for example match *strand and beach* or *lathe the table leg*. Without the last word boundary anchor it would also match *the theory*.

## 9.3 Named regexes

You can declare regexes just like subroutines and even name them. Suppose you found the example at the beginning of this chapter useful and want to make it available easily. Suppose also you want to extend it to handle contractions such as *doesn't* or *isn't*:

```
1 my regex word { \w+ [ \\' \w+]? }
2 my regex dup  { « <word=&word> \W+ $<word> » }
3
4 if $s ~~ m/ <dup=&dup> / {
5     say "Found '{$<dup><word>}' twice in a row";
6 }
```

This code introduces a regex named `word`, which matches at least one word character, optionally followed by a single quote. Another regex called `dup` (short for *duplicate*) is anchored at a word boundary.

Since named regex are very much like subroutines, within a regex, the syntax `<&word>` locates the regex `word` within the current lexical scope and matches as if the regex were used in its place. The `<name=&regex>` syntax creates a capture named `name`, which records what `&regex` matched in the match object.

In our example, `dup` calls the `word` regex, then matches at least one non-word character, and then matches the same string as previously matched by the regex `word`. It ends with another word boundary. The syntax for This *backreference* is a dollar sign followed by the name of the capture in angle brackets.<sup>6</sup> `»` simply looks up a regex named `word` in the current grammar and parent grammars, and creates a capture of the same name.>

Within the `if` block, `$<dup>` is short for `$/{'dup'}`. It accesses the match object that the regex `dup` produced. `dup` also has a subrule called `word`, and the match object produced from that call is accessible as `$<dup><word>`.

---

<sup>6</sup> In grammars, which are introduced in the next chapter, `c« <word`

Just as subroutines allow for ordinary code, named regexes make it easy to organize complex regexes in smaller pieces.

## 9.4 Modifiers

The previous example to match a list of words was:

```
1 m/(\w+) ** [\s*] \s* 'and' \s* (\w+)/
```

This works, but the repeated “I don’t care about whitespace” units are clumsy. The desire to allow whitespace *anywhere* in a string is common, and Perl 6 regexes provide such an option: the `:sigspace` modifier (shortened to `:s`):

```
1 my $ingredients = 'eggs, milk, sugar and flour';
2
3 if $ingredients ~~ m/:s ( \w+ ) ** \s* 'and' (\w+)/ {
4     say 'list: ', $/[0].join(' | ');
5     say 'end: ', $/[1];
6 }
```

This modifier allows optional whitespace in the text wherever there is one or more whitespace character in the pattern. It’s even a bit cleverer than that: between two word characters whitespace is mandatory. The regex does *not* match the string `eggs, milk, sugarand-flour`.

The `:ignorecase` or `:i` modifier makes the regex insensitive to upper and lower case, so `m/ :i perl /` matches not only `perl`, but also `Perl` or `PERL` (though nobody really writes the programming language in all uppercase letters).



## 9.5 Backtracking control

In the course of matching a regex against a string, the regex engine may reach a point where an alternation has matched a particular branch or a quantifier has greedily matched all it can but the final portion of the regex fails to match. In this case, the regex engine backs up and attempts to match another alternative or matches one fewer character on the quantified portion to see if the overall regex succeeds. This process of failing and trying again is called *backtracking*.

When matching `m/\w+ 'en' /` against the string `oxen`, the `\w+` group first matches the whole string (because of the greediness of `+`), but then the `en` literal at the end can't match anything. `\w+` gives up one character to match `oxe`. `en` still can't match, so the `\w+` group again gives up one character and now matches `ox`. The `en` literal can now match the last two characters of the string, and the overall match succeeds.

While backtracking is often useful and convenient, it can also be slow and confusing. A colon `:` switches off backtracking for the previous quantifier or alternation. So `m/ \w+ : 'en' /` can never match any string, because the `\w+` always eats up all word characters, and never releases them.

The `:ratchet` modifier disables backtracking for a whole regex, which is often desirable in a small regex called often from other regexes. The duplicate word search regex had to anchor the regex to word boundaries, because `\w+` would allow matching only part of a word. Disabling backtracking produces simpler behavior where `\w+` always matches a full word:

```
1 my regex word { :ratchet \w+ [ \\' \w+]? }
2 my regex dup { <word=&word> \W+ $<word> }
3
4 # no match, doesn't match the 'and'
5 # in 'strand' without backtracking
6 'strand and beach' ~~ m/<&dup>/
```

However the effect of `:ratchet` applies only to the regex in which it appears. The outer regex still backtracks, and can also retry the regex `word` at a different starting position.

## Chapter 9 | PATTERN MATCHING

The regex `{ :ratchet ... }` pattern is common that it has its own shortcut: `token { ... }`. The duplicate word searcher is idiomatic when written:

```
1 my token word { \w+ [ \ ' \w+]? }
2 my regex dup { <word> \W+ $<word> }
```

A token that also switches on the `:sigspace` modifier is a rule:

```
1 my rule wordlist { <word> ** \, 'and' <word> }
```

### 9.6 Substitutions

Regexes are not only popular for data validation and extraction, but also data manipulation. The `subst` method matches a regex against a string. If a match is found, it substitutes the portion of the string that matches with its second argument.

```
1 my $spacey = 'with many superfluous spaces';
2 say $spacey.subst(rx/ \s+ /, ' ', :g);
3 # output: with many superfluous spaces
```

The `:g` at the end tells the substitution to work *globally* to replace every match. Without `:g`, it stops after the first match.

Note the use of `rx/ ... /` rather than `m/ ... /` to construct the regex. The former constructs a regex object. The latter not only constructs the regex object, but immediately matches it against the topic variable `$_`. Using `m/ ... /` in the call to `subst` creates a match object and passes it as the first argument, rather than the regex itself.

## 9.7 Other Regex Features

Sometimes you want to call other regexes, but don't want them to capture the matched text. For example, when parsing a programming language you might discard whitespaces and comments. You can achieve that by calling the regex as `<.otherrule>`.

For example, if you use the `:sigspace` modifier, every continuous piece of whitespaces calls the built-in rule `<.ws>`. This use of a rule rather than a character class allows you to define your own version of whitespace characters (see grammars).

Sometimes you just want to take a look ahead, and check if the next characters fulfill some properties without actually consuming them, so that the following parts of the regex can still match them. This is common in substitutions. In normal English text, you always place a whitespace after a comma. If somebody forgets to add that whitespace, a regex can clean up after the lazy writer:

```
1 my $str = 'milk,flour,sugar and eggs';
2 say $str.subst(/',' <?before \w>/, ' ', ':g);
3 # output: milk, flour, sugar and eggs
```

The word character after the comma is not part of the match, because it is in a look-ahead, which `<?before ... >` introduces. The leading question mark indicates an *zero-width assertion*: a rule that never consumes characters from the matched string. You can turn any call to a subrule into an zero width assertion. The built-in token `<alpha>` matches an alphabetic character, so you can rewrite this example as:

```
1 say $str.subst(/',' <?alpha>/, ' ', ':g);
```

An leading exclamation mark negates the meaning; another variant is:

```
1 say $str.subst(/',' <!space>/, ' ', ':g);
```

A look in the opposite direction is also possible, with `<?after>`. In fact many built-in anchors can be written with look-ahead and look-behind assertions, though usually not quite as efficient:

Table 9.3: Emulation of anchors with look-around assertions

Anchor	Meaning	Equivalent Assertion
<code>^</code>	start of string	<code>&lt;!after .&gt;</code>
<code>^^</code>	start of line	<code>&lt;?after ^   \n &gt;</code>
<code>\$</code>	end of string	<code>&lt;!before .&gt;</code>
<code>»</code>	right word boundary	<code>&lt;?after \w&gt; &lt;!before \w&gt;</code>

## 9.8 Match objects

```

1  sub line-and-column(Match $m) {
2      my $line   = ($m.orig.substr(0, $m.from).split("\n")).elems;
3      # RAKUDO workaround for RT #70003, $m.orig.rindex(...) directly fails
4      my $column = $m.from - ('' ~ $m.orig).rindex("\n", $m.from);
5      $line, $column;
6  }
7
8  my $s = "the quick\nbrown fox jumped\nover the the lazy dog";
9
10 my token word { \w+ [ \' \w+]? }
11 my regex dup { <word> \W+ $<word> }
12 if $s ~~ m/ <dup> / {
13     my ($line, $column) = line-and-column($/);
14     say "Found '{$<dup><word>}' twice in a row";
15     say "at line $line, column $column";
16 }
17
18 # output:

```

```
19 # Found 'the' twice in a row
20 # at line 3, column 6
```

Every regex match returns an object of type `Match`. In boolean context, a match object returns `True` for successful matches and `False` for failed ones. Most properties are only interesting after successful matches.

The `orig` method returns the string that was matched against. The `from` and `to` methods return the positions of the start and end points of the match.

In the previous example, the `line-and-column` function determines the line number in which the match occurred by extracting the string up to the match position (`$m.orig.substr(0, $m.from)`), splitting it by newlines, and counting the elements. It calculates the column by searching backwards from the match position and calculating the difference to the match position.

The `index` method searches a string for another substring, and returns the position of the search string.

The `rindex` method does the same, but searches backwards from the end of the string, so it finds the position of the last occurrence of the substring.

Using a match object as an array yields access to the positional captures. Using it as a hash reveals the named captures. In the previous example, `$<dup>` is a shortcut for `$/<dup>` or `$/{'dup'}`. These captures are again `Match` objects, so match objects are really trees of matches.

The `caps` method returns all captures, named and positional, in the order in which their matched text appears in the source string. The return value is a list of `Pair` objects, the keys of which are the names or numbers of the capture and the values the corresponding `Match` objects.

```
1 if 'abc' =~ m/(.) <alpha> (.) / {
2   for $/.caps {
```

## Chapter 9 | PATTERN MATCHING

```
3         say .key, ' => ', .value;
4
5     }
6 }
7
8 # Output:
9 # 0 => a
10 # alpha => b
11 # 1 => c
```

In this case the captures are in the same order as they are in the regex, but quantifiers can change that. Even so, `$/ .caps` follows the ordering of the string, not of the regex. Any parts of the string which match but not as part of captures will not appear in the values that `caps` returns.

To access the non-captured parts too, use `$/ .chunks` instead. It returns both the captured and the non-captured part of the matched string, in the same format as `caps`, but with a tilde `~` as key. If there are no overlapping captures (which could only come from look-around assertions), the concatenation of all the pair values that `chunks` returns is the same as the matched part of the string.

# 10

## Grammars

Grammars organize regexes, just like classes organize methods. The following example demonstrates how to parse JSON, a data exchange format already introduced (multis).

```
1 # file lib/JSON/Tiny/Grammar.pm
2
3 grammar JSON::Tiny::Grammar {
4     rule TOP          { ^[ <object> | <array> ]$ }
5     rule object       { '{' ~ '}' <pairlist>      }
6     rule pairlist     { [ <pair> ** [ \, ] ]?     }
7     rule pair         { <string> ':' <value>      }
8     rule array        { '[' ~ ']' [ <value> ** [ \, ] ]? }
9
10    proto token value { <...> };
11
12    token value:sym<number> {
13        '-'?
```

## Chapter 10 | GRAMMARS

```
14      [ 0 | <[1..9]> <[0..9]>* ]
15      [ \. <[0..9]>+ ]?
16      [ <[eE]> [\+|\-]? <[0..9]>+ ]?
17  }
18
19  token value:sym<true>    { <sym>    };
20  token value:sym<false>   { <sym>    };
21  token value:sym<null>    { <sym>    };
22  token value:sym<object>  { <object> };
23  token value:sym<array>   { <array>  };
24  token value:sym<string>  { <string> };
25
26  token string {
27      \" ~ \" [ <str> | \\ <str_escape> ]*
28  }
29
30  token str {
31      [
32          <!before \t>
33          <!before \n>
34          <!before \>
35          <!before \>
36          .
37      ]+
38      # <-[\"\\\\t\n]>+
39  }
40
41  token str_escape {
42      <[\"\\/bfnrt]> | u <xdigit>**4
43  }
44
45  }
46
47
48  # test it:
49  my $tester = '{
50      "country": "Austria",
51      "cities": [ "Wien", "Salzburg", "Innsbruck" ],
```



```

52     "population": 8353243
53   }';
54
55   if JSON::Tiny::Grammar.parse($tester) {
56       say "It's valid JSON";
57   } else {
58       # TODO: error reporting
59       say "Not quite...";
60   }

```

A grammar contains various named regexes. The call to `JSON::Tiny.parse($string)` starts by calling `TOP`.

Regexes inside a grammar do not need a scope declarator, the default to the same lookup semantics as methods in classes. The call to another regex simplifies to `<regex.name>`, which also creates a capture of the same name.

Regexes inside a grammar do not need a scope declarator, the default to the same lookup semantics as methods in classes. The call to another regex simplifies to `<regex.name>`, which also creates a capture of the same name.

Rule `TOP` anchors the match to the start and end of the string, so that the whole string has to be in valid JSON format for the match to succeed. It then either matches an `<array>` or an `<object>`. Subsequent calls are straightforward, and reflect the structure in which JSON components can appear.

Regexes can be recursive. An array contains value, and in turn a value can be an array. That won't cause any infinite loops as long as at least one regex per recursive call consumes at least one character. If a set of regexes were to call each other recursively without ever progressing in the string, the recursion could go on infinitely, never progressing in the string and never proceeding to other parts of the grammar.

The only new regex syntax used in the `JSON::Tiny` grammar is the *goal matching* syntax `'{' ~'}' [ ... ]`, which resembles `'{' ... '}'`, but gives a better error message upon failure. It sets the term to the right of the tilde character as the goal, and then matches the final term `[ ... ]`. If the goal does not match, Perl will issue an error.

Another novelty is the declaration of a *proto token*:

```

1  proto token value { <...> };
2
3  token value:sym<number> {
4      '-'?
5      [ 0 | <[1..9]> <[0..9]>* ]
6      [ \. <[0..9]>+ ]?
7      [ <[eE]> [\+|\-]? <[0..9]>+ ]?
8  }
9
10 token value:sym<true>    { <sym>    };
11 token value:sym<false>  { <sym>    };

```

The `proto token` syntax marks `value` as a set of alternatives instead of a single regex. Each alternative has a name of the form `token value:sym<thing>`, which can read as *alternative of value with parameter sym set to thing*. The body of such an alternative is a normal regex, where the call `<sym>` matches the value of the parameter, in this example `thing`.

When calling the rule `<value>`, the grammar engine attempts to match every alternative (and can do so in parallel). The longest match wins.

## 10.1 Grammar Inheritance

The similarity of grammars to classes goes deeper than storing regexes in a namespace as a class might store methods—you can inherit from and extend grammars, mix roles into them, and take advantage of polymorphism. In fact, a grammar is a class which by default inherits from `Grammar` instead of `Any`.

Suppose you want to enhance the JSON grammar to allow single-line C++ or JavaScript comments. (These begin with `//` and continue until the end of the line.) The simplest enhancement is to allow such a comment in any place where whitespace is valid.

Most of the grammar uses *rules*, which as you may recall are like tokens with the `:sigspace` modifier enabled. As this uses the `ws` token to find significant whitespace, the simplest approach is to override that token:

```
1 grammar JSON::Tiny::Grammar::WithComments
2     is JSON::Tiny::Grammar {
3
4     token ws {
5         \s* [ '//' \N* \n ]?
6     }
7 }
8
9 my $tester = '{
10     "country": "Austria",
11     "cities": [ "Wien", "Salzburg", "Innsbruck" ],
12     "population": 8353243 // data from 2009-01
13 }';
14
15 if JSON::Tiny::Grammar::WithComments.parse($tester) {
16     say "It's valid (modified) JSON";
17 }
```

The first two lines introduce a grammar that inherits from `JSON::Tiny::Grammar` through the use of the `is` trait. As subclasses inherit methods from superclasses, so any grammar rule not present in the derived grammar will come from its base grammar.

In this minimal JSON grammar, whitespace is never mandatory, so `ws` can match nothing at all. After optional spaces, two slashes `'//'` introduce a comment, after which must follow an arbitrary number of non-newline characters, and then a newline. In prose, it extends to the rest of the line.

Inherited grammars may also add variants to proto tokens:

```
1 grammar JSON::ExtendedNumeric is JSON::Tiny::Grammar {
2     token value:sym<nan> { <sym> }
3     token value:sym<inf> { <[+-]>? <sym> }
4 }
```

In this, grammar a call to `<value>` matches either one of the newly added alternatives, or any of the old alternatives from the parent grammar `JSON::Tiny::Grammar`. Such extensibility is difficult to achieve with ordinary, `|` delimited alternatives.

## 10.2 Extracting data

The `parse` method of a grammar returns a `Match` object, through which you can access all the relevant information of the match. If you were to do this yourself, you'd have to write a function which traverses the match tree recursively to find and to extract the interesting data. An alternative solution exists: *reduction methods*, also called *action methods*.

```

1  class JSON::Tiny::Actions {
2      method TOP($/)      { make $/.values.[0].ast }
3      method object($/)   { make $<pairlist>.ast.hash }
4      method pairlist($/) { make $<pair>».ast }
5      method pair($/)     { make $<string>.ast => $<value>.ast }
6      method array($/)    { make [$<value>»].ast }
7      method string($/)   { make join ' ', $/.caps>>.value>>.ast }
8
9      # TODO: make that
10     # make +$/
11     # once prefix:<+> is sufficiently polymorphic
12     method value:sym<number>($/) { make eval $/ }
13     method value:sym<string>($/) { make $<string>.ast }
14     method value:sym<true> ($/) { make Bool::True }
15     method value:sym<false> ($/) { make Bool::False }
16     method value:sym<null> ($/) { make Any }
17     method value:sym<object>($/) { make $<object>.ast }
18     method value:sym<array> ($/) { make $<array>.ast }
19
20     method str($/)          { make ~$/ }
21
22     method str_escape($/) {
23         if $<xdigit> {
24             make chr(:16($<xdigit>.join));

```

```

25         } else {
26             my %h = '\|' => "\\|",
27                     'n'  => "\\n",
28                     't'  => "\\t",
29                     'f'  => "\\f",
30                     'r'  => "\\r";
31             make %h{$/};
32         }
33     }
34 }
35
36 my $actions = JSON::Tiny::Actions.new();
37 JSON::Tiny::Grammar.parse($str, :$actions);

```

This example passes an actions object to the grammar's parse method. Whenever the grammar engine finishes parsing one rule, it calls a method on the actions object with the same name as the current rule. If no such method exists, the grammar engine calls no method and moves along.

If a method does exist, the grammar engine passes the current match object as a positional argument.

Each match object has a slot called *ast* (short for *abstract syntax tree*) for a payload object. This slot can hold a custom data structure that you create from the action methods. Calling `make $thing` in an action method sets the *ast* attribute of the current match object to `$thing`.

In the case of the JSON parser, the payload can be the data structure that the JSON string represents.

For each matching rule an action method is called, which populates the *ast* slot of the match object, thus transforming the match tree to a more high-level tree - here the actual JSON tree.

Although the rules and action methods live in different namespaces (and in a real-world project probably even in separate files), here they are adjacent to demonstrate their correspondence:

```
1 rule TOP          { ^ [ <object> | <array> ]$ }
2 method TOP($/) { make $/.values.[0].ast }
```

The rule has an alternation with two branches, `object` and `array`. Both have a named capture. `$.values` returns a list of all captures, here just one - either the object or the array capture.

The action method takes the AST attached to the match object of that sub capture, and promotes it as its own AST by calling `make`.

```
1 rule object      { '{' ~ '}' <pairlist> }
2 method object($/) { make $<pairlist>.ast.hash }
```

The reduction method for `object` extracts the AST of the `pairlist` submatch and turns it into a hash by calling its `hash` method.

```
1 rule pairlist    { [ <pair> ** [ \, ] ]? }
2 method pairlist($/) { make $<pair>».ast; }
```

The `pairlist` rule matches multiple comma-separated pairs. The reduction method calls the `.ast` method on each matched pair and installs the result list in its own AST.

```
1 rule pair        { <string> ':' <value> }
2 method pair($/) { make $<string>.ast => $<value>.ast }
```

A pair consists of a string key and a value, so the action method constructs a Perl 6 pair with the `=>` operator.

The other action methods work the same way. They transform the information they extract from the match object into “native” Perl 6 data structures, and call `make` to set it as their own AST.

The action methods that belong to a proto token are parametric in the same way as the alternative:

```
1 token value:sym<null>      { <sym>    };
2 method value:sym<null>($/) { make Any }
3
4 token value:sym<object>    { <object> };
5 method value:sym<object>($/) { make $<object>.ast }
```

When a <value> call matches, the action method with the same parametrization as the matching alternative executes.





# 11

## Built-in types, operators and methods

Many operators work on a particular *type* of data. If the type of the operands differs from the type of the operand, Perl will make copies of the operands and convert those to the needed types. For example, `$a + $b` will convert a copy of both `$a` and `$b` to numbers (unless they are numbers already). This implicit conversion is called *coercion*.

Besides operators, other syntactic elements coerce their elements: `if` and `while` coerce to truth values (bool), `for` views things as lists, and so on.

### 11.1 Numbers

Sometimes coercion is intuitively transparent. Perl 6 has several numeric types which you can intermix freely. For example, subtracting a floating point value from an integer works, as in `123 - 12.1e1`.

## Chapter 11 | BUILT-IN TYPES, OPERATORS AND METHODS

The most important types are:

### Int

Int objects store integer numbers of arbitrary size. If you write a literal that consists only of digits, such as 12, it is an Int.

### Rat

Rat, short for *rational*, stores fractional numbers without loss of precision. Because Rat tracks its numerator and denominator as integers, mathematical operations on Rats with large components can become quite slow. For this reason, rationals with large denominators automatically degrade to Num.

Writing a fractional value with a dot as the decimal separator produces a Rat, for example 3.14.

### Num

Num is the floating point type. It stores sign, mantissa, and exponent, each with a fixed width. Calculations involving Num numbers are usually quite fast, but subject to limited precision.

Numbers in scientific notation such as 6.022e23 are of type Num.

### Complex

Complex is the complex number type. Complex numbers have two parts to them: a real part and an imaginary part. If either part is NaN, then the entire number may possibly be NaN.

Numbers in the form  $a + bi$  are of type Complex.

The following operators are available for all number types:

```

1 Binary operators:
2 Operator      Description
3 **            Exponentiation: $a**$b is $a to the power of $b
4 *             multiplication
5 /             division
6 div           integer division
7 +             addition
8 -             subtraction
9
10 Unary operators:
11 Operator      Description
12 +             conversion to number
13 -             negation

```

Most mathematical functions are available both as methods and functions, so you can write both `(-5).abs` and `abs(-5)`.

```

1 Method        Description
2 abs           absolute value
3 sqrt          square root
4 log           natural logarithm
5 log10         logarithm to base 10
6
7 ceil          rounding up to an integer
8 floor         rounding down to an integer
9 round         rounding to next integer
10 sign         -1 for negative, 0 for 0, 1 for positive values

```

The trigonometric functions `sin`, `cos`, `tan`, `asin`, `acos`, `atan`, `sec`, `cosec`, `cotan`, `asec`, `acosec`, `acotan`, `sinh`, `cosh`, `tanh`, `asinh`, `acosh`, `atanh`, `sech`, `cosech`, `cotanh`, `asech`, `acosech` and `acotanh` are available, and work in units of radians by default. You may specify the unit with an argument of `Degrees`, `Gradians` or `Circles`. For example, `180.sin(Degrees)` is approximately 0.

## 11.2 Strings

Strings stored as `Str` are sequences of characters, independent of character encoding. The `Buf` type is available for storing binary data. The `encode` method converts a `Str` to `Buf`. `decode` goes the other direction.

The following operations are available for strings:

```

1  Binary operators:
2  Operator      Description
3  ~             concatenation: 'a' ~ 'b' is 'ab'
4  x             replication: 'a' x 2 is 'aa'
5
6  Unary operators:
7  Operator      Description
8  ~             conversion to string: ~1 becomes '1'
9
10 Methods:
11 .chomp         remove trailing newline
12 .substr($start, $length)  extract a part of the string. $length defaults
13                                     to the rest of the string
14 .chars         number of characters in the string
15 .uc            upper case
16 .lc            lower case
17 .ucfirst       convert first character to upper case
18 .lcfirst       convert first character to lower case
19 .capitalize    convert the first character of each word to upper case, and
20                all other characters to lower case

```

## 11.3 Bool

A Boolean value is either True or False. Any value can coerce to a boolean in boolean context. The rules for deciding if a value is true or false depends on the type of the value:

### Strings

Empty strings and "" evaluate to False. All other strings evaluate to True.

### Numbers

All numbers except zero evaluate to True.

### Lists and Hashes

Container types such as lists and hashes evaluate to False if they are empty, and to True if they contain at least one value.

Constructs such as `if` automatically evaluate their conditions in boolean context. You can force an explicit boolean context by putting a `?` in front of an expression. The `!` prefix negates the boolean value.

```
1 my $num = 5;
2
3 # implicit boolean context
4 if $num { say "True" }
5
6 # explicit boolean context
7 my $bool = ?$num;
8
9 # negated boolean context
10 my $not_num = !$num;
```



# Index

`+=`, 10  
`.*` method calls, 81  
`.+` method calls, 81  
`.?` method calls, 81  
`<=>`, 26  
`=>`, 10  
`~`; regex meta character, 107

operator; fat arrow, 18

abstract syntax tree, 111  
accessor methods, 68  
action methods, 110  
anonymous subroutines, 31  
Any, 10  
arguments, 29  
arity, 61  
array, 7  
assignment, 7, 18  
attributes, 67  
autovivification, 10

behavior, 66  
bless, 70  
block, 7, 9, 11  
Bool, 119

Buf, 118

calling sets, 81  
callsame, 64  
callwith, 64  
candidates, 54  
Capture, 44  
captures, 44  
class, 66  
classes, 66  
classes; accessors, 68  
classes; attributes, 67  
classes; behavior, 66  
classes; encapsulation, 67  
classes; has, 66  
classes; methods, 68  
cmp, 26  
coercion, 115  
Complex, 116  
composition, 78  
composition; conflicts, 79  
composition; methods, 79  
composition; multi methods, 80  
composition; resolution, 79  
constraint, 43

- constraint; type, 43
- constraints, 55
- constructors, 70
- does, 79, 83
- double-quoted strings, 12
- encapsulation, 67
- enum, 59
- fat arrow, 10
- file handle, 7
- files; handle, 7
- first-class subroutines, 31
- flattening composition, 78
- for, 9
- functions; arity, 61
- functions; protos, 63
- functions; substr, 61
- goal matching, 107
- has, 66
- hash, 8
- identifier, 7
- implicit constraints, 61
- infix, 19
- Int, 116
- interpolation, 12
- invocant, 7, 69
- JSON, 53
- leg, 26
- lexical, 7
- Match, 103
- Match.caps, 103
- Match.chunks, 104
- Match.from, 103
- Match.orig, 103
- Match.to, 103
- Match; access as a hash, 103
- meta operator; [], 20
- meta operator; reduction, 20
- method, 7
- methods, 68
- methods; multidispatch, 63
- multidispatch, 53
- multidispatch; constraints, 55
- multidispatch; narrowness, 57
- multimethods, 63
- multis, 53
- named captures, 103
- nominal type, 56
- Num, 116
- objects; bless, 70
- operator, 17
- operator precedence, 22
- operator; =, 18
- operator; assignment, 18
- operator; infix operators, 19
- operator; max, 19
- operator; postcircumfix, 19
- operator; postfix, 19
- operator; x, 21
- operator;=>, 18
- operators; +=, 10
- operators; ., 69
- operators; m//, 100
- operators; postincrement, 11
- operators; preincrement, 11



- operators; print, 12
- operators; rx//, 100
- operators; say, 12
- operators; trigonometry, 117
- pair, 10, 18
- parameter, 32
- parameter type constraint, 43
- parameters; anonymous, 60
- postcircumfix, 19
- postfix, 19
- postincrement, 11
- precedence, 22
- preincrement, 11
- print, 12
- printf, 21
- proto token, 107
- protos, 63
- Rakudo, 3
- Rat, 116
- rational type, 116
- reduction methods, 110
- regex, 89
- regex; \w, 91
- regex; \* quantifier, 93
- regex; \*\* quantifier, 93
- regex; + quantifier, 93
- regex; . character, 90
- regex; :, 99
- regex; :g, 100
- regex; :i, 98
- regex; :ignorecase modifier, 98
- regex; :ratchet, 99
- regex; :s modifier, 98
- regex; :sigspace modifier, 98

- regex; ? quantifier, 93
- regex; \$, 94
- regex; \$/, 95
- regex; ^, 94
- regex; alternation, 94
- regex; anchors, 94
- regex; avoid captures, 101
- regex; backreference, 97
- regex; backtracking, 99
- regex; captures, 95
- regex; character class addition, 92
- regex; character class subtraction, 92
- regex; character range, 92
- regex; custom character classes, 91
- regex; disable backtracking, 99
- regex; global substitution, 100
- regex; greedy matching, 93
- regex; grouping, 94
- regex; lookahead, 101
- regex; Match object, 103
- regex; metasyntactic characters, 90
- regex; modifiers, 98
- regex; named, 97
- regex; named captures, 103
- regex; negative look-ahead assertion, 101
- regex; non-greedy matching, 93
- regex; quantified capture, 95
- regex; quantifier, 93
- regex; rule, 100
- regex; special characters, 90
- regex; string end anchor, 94
- regex; string start anchor, 94
- regex; subrule, 97
- regex; token, 99
- regex; zero-width assertion, 101
- regular expressions, 89

- repetition operator, 21
- required methods, 81
- return, 41
- return value, 29
- return; implicit, 42
- role, 75
- roles, 75
- roles; requirements, 81
- roles; runtime application, 82
- rule, 100
- say, 12
- scalar, 7
- scoping; subroutines, 30
- sigil, 7
- sigils; &, 67
- signature, 29
- signature unpacking, 47
- signatures; subroutines, 32
- single-quoted strings, 12
- slurpy, 40
- sort; stable, 11
- stable sort, 11
- state, 66
- statement, 7
- Str, 118
- string, 7
- string literal, 7
- strings, 118
- strings; double-quoted, 12
- strings; literal, 7
- strings; single-quoted, 12
- subroutine, 29
- subroutines; anonymous, 31
- subroutines; declaration, 29
- subroutines; first-class, 31
- subroutines; scoping, 30
- subroutines; signature, 32
- subrule, 97
- subset type, 55
- subsignature, 47
- subst, 100
- substitutions, 100
- substr, 61
- term, 19
- token, 99
- topic, 11
- topic variable, 11
- traits; implicit constraints, 61
- traits; is rw, 68
- trigonometric functions, 117
- twigils, 67
- twigils; , 67
- twigils; ., 68
- twigils; accessors, 68
- type, 115
- type capture, 60
- type object, 66
- types; Bool, 119
- types; Buf, 118
- types; capture, 60
- types; Complex, 116
- types; constraints, 55
- types; Int, 116
- types; nominal, 56
- types; Num, 116
- types; Rat, 116
- types; rational, 116
- types; Str, 118
- types; subset, 55
- units, 117

unpacking, 47

v6, 7

value identity, 24

variable; scalar, 7

variables; \$\_, 11

variables; lexical, 7

where, 43

Zen slice, 13