

# Learn Me a Haskell

Jacob Bishop

2023-07-01

## 1 Introduction

WIP

## 2 Starting Out

### 第一个函数

在 `./test/` 文件夹下创建一个 `baby.hs` 的文件，写入：

```
1 doubleMe x = x + x
```

使用 `ghci` 加载该文件（在本项目根目录时使用 `:l tests/baby`）：

```
1 ghci> :l baby
2 [1 of 1] Compiling Main                ( baby.hs, interpreted )
3 Ok, modules loaded: Main.
4 ghci> doubleMe 9
5 18
6 ghci> doubleMe 8.3
7 16.6
```

一个带有 `if` 的函数：

```
1 doubleSmallNumber x =
2   if x > 100
3   then x
4   else x * 2
```

Haskell 中的 `if` 声明是一个表达式，那么 `else` 是强制性的，因为表达式一定要有所返回。因此加上述函数可以改写为：

```
1 doubleSmallNumber' x = (if x > 100 then x else x * 2) + 1
```

这里的 `'` 符号是 Haskell 中的有效字符，且在 Haskell 中并没有特殊的意义，因此可以用于函数名。通常情况下，使用 `'` 代表着一个函数（非懒加载的函数）的严格版本，或是一个有细微变化的函数或者变量。又因为 `'` 是一个有效字符，那么可以创建以下函数：

```
1 conanO'Brien = "It's a-me, Conan O'Brien!"
```

这里又有两点值得注意的地方。首先，函数名不能以大写开头，稍后会进行说明；其次，该函数并没有任何入参。当一个函数没有入参，我们通常称其为一个定义 *definition*，因为一旦定义了它便不能修改其名称，以及其返回。

### list 的介绍

Haskell 中的 `list` 是 **同质的 homogenous** 数据结构。

#### Note

在 `GHCI` 中可以使用 `let` 关键字定义一个名称。换言之，`GHCI` 中的 `let a = 1` 等同于脚本中的 `a = 1`。

通常使用 `++` 操作符将两个数组进行合并：

```
1 ghci> [1,2,3,4] ++ [9,10,11,12]
2 [1,2,3,4,9,10,11,12]
3 ghci> "hello" ++ " " ++ "world"
4 "hello world"
5 ghci> ['w','o'] ++ ['o','t']
6 "woot"
```

可以使用 `:` 操作符将元素直接添加至数组头部：

```
1 ghci> 'A':" SMALL CAT"
2 "A SMALL CAT"
3 ghci> 5:[1,2,3,4,5]
4 [5,1,2,3,4,5]
```

实际上，`[1, 2, 3]` 是 `1:2:3:[]` 的语法糖，其中 `[]` 为一个空数组。如果头部追加 `3`，`[]` 就变成了 `[3]`，再次进行头部追加 `2`，则变为 `[2, 3]`，以此类推。

如果希望通过索引获取数组中的元素，那么可以使用 `!!` 操作符：

```
1 ghci> "Steve Buscemi" !! 6
2 'B'
3 ghci> [9.4,33.2,96.2,11.2,23.25] !! 1
4 33.2
```

超出索引时则会报错。

数组还可以通过操作符 `<`，`<=`，`==`，`>` 以及 `>=` 操作符来进行比较，而比较的方式则是顺序比较。当进行头部比较元素相等时，再进行下一个元素进行比较。

数组的四种基础操作 `head`，`tail`，`last` 以及 `init`：

```
1 ghci> head [5,4,3,2,1]
2 5
3 ghci> tail [5,4,3,2,1]
4 [4,3,2,1]
5 ghci> last [5,4,3,2,1]
6 1
7 ghci> init [5,4,3,2,1]
8 [5,4,3,2]
```

当使用上述四种操作时，需要注意是否应用于空数组，这样的错误在编译期并不能被发现。其它的操作：

1. `length` 获取数组长度；
2. `null` 检查数组是否为空；
3. `reverse` 翻转数组；
4. `take` 获取数组的头几个元素的数组；

5. `drop` 移除数组的头几个元素，并返回剩余元素的数组；
6. `maximum` 获取最大值；
7. `minimum` 获取最小值；
8. `sum` 求和；
9. `product` 求积；
10. `elem` 元素是否存在于数组中。

```
1 ghci> length [5,4,3,2,1]
2 5
3
4 ghci> null [1,2,3]
5 False
6 ghci> null []
7 True
8
9 ghci> reverse [5,4,3,2,1]
10 [1,2,3,4,5]
11
12 ghci> take 3 [5,4,3,2,1]
13 [5,4,3]
14 ghci> take 1 [3,9,3]
15 [3]
16 ghci> take 5 [1,2]
17 [1,2]
18 ghci> take 0 [6,6,6]
19 []
20
21 ghci> drop 3 [8,4,2,1,5,6]
22 [1,5,6]
23 ghci> drop 0 [1,2,3,4]
24 [1,2,3,4]
25 ghci> drop 100 [1,2,3,4]
26 []
27
28 ghci> minimum [8,4,2,1,5,6]
29 1
30 ghci> maximum [1,9,2,3,4]
31 9
32
33 ghci> sum [5,2,1,6,3,2,5,7]
34 31
35 ghci> product [6,2,1,2]
36 24
37 ghci> product [1,2,5,6,7,9,2,0]
```

```

38 0
39
40 ghci> 4 `elem` [3,4,5,6]
41 True
42 ghci> 10 `elem` [3,4,5,6]
43 False

```

## Texas 排列

```

1 ghci> [1..20]
2 [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20]
3 ghci> ['a'..'z']
4 "abcdefghijklmnopqrstuvwxyz"
5 ghci> ['K'..'Z']
6 "KLMNOPQRSTUVWXYZ"

```

带有 step 的排列:

```

1 ghci> [2,4..20]
2 [2,4,6,8,10,12,14,16,18,20]
3 ghci> [3,6..20]
4 [3,6,9,12,15,18]

```

而对于浮点数的排列需要注意精度问题:

```

1 ghci> [0.1, 0.3 .. 1]
2 [0.1,0.3,0.5,0.7,0.8999999999999999,1.0999999999999999]

```

以下是若干用于生产无限长度数组的函数:

**cycle** 循环周期:

```

1 ghci> take 10 (cycle [1,2,3])
2 [1,2,3,1,2,3,1,2,3,1]
3 ghci> take 12 (cycle "LOL ")
4 "LOL LOL LOL "

```

**repeat** 重复:

```

1 ghci> take 10 (repeat 5)
2 [5,5,5,5,5,5,5,5,5,5]

```

另外就是 **replicate** 函数可以重复单个元素:

```

1 ghci> replicate 3 10
2 [10,10,10]

```

## 列表表达式

数学里的 集合表达式 *set comprehensions* 例如  $S = 2 \cdot x | x \in \mathbb{N}, x \leq 10$ ; Haskell 中的列表表达式, 例如 1 至 10 数组中每个元素乘以 2:

```
1 ghci> [x*2 | x <- [1..10]]
2 [2,4,6,8,10,12,14,16,18,20]
```

为列表表达式添加条件（或称谓语 predicate）：

```
1 ghci> [x*2 | x <- [1..10], x*2 >= 12]
2 [12,14,16,18,20]
3 ghci> [ x | x <- [50..100], x `mod` 7 == 3]
4 [52,59,66,73,80,87,94]
```

将列表表达式置于一个函数中便于复用：

```
1 ghci> boomBangs xs = [ if x < 10 then "BOOM!" else "BANG!" | x <- xs, odd x]
2 ghci> boomBangs [7..13]
3 ["BOOM!", "BOOM!", "BANG!", "BANG!"]
```

多个谓词也是可以的：

```
1 ghci> [ x | x <- [10..20], x /= 13, x /= 15, x /= 19]
2 [10,11,12,14,16,17,18,20]
```

除此之外，还可以处理若干数组：

```
1 ghci> [ x*y | x <- [2,5,10], y <- [8,10,11]]
2 [16,20,22,40,50,55,80,100,110]
```

当然也可以加上谓词：

```
1 ghci> [ x*y | x <- [2,5,10], y <- [8,10,11], x*y > 50]
2 [55,80,100,110]
```

那么对于字符串也可以使用列表表达式：

```
1 ghci> let nouns = ["hobo", "frog", "pope"]
2 ghci> let adjectives = ["lazy", "grouchy", "scheming"]
3 ghci> [adjective ++ " " ++ noun | adjective <- adjectives, noun <- nouns]
4 ["lazy hobo", "lazy frog", "lazy pope", "grouchy hobo", "grouchy frog",
5 "grouchy pope", "scheming hobo", "scheming frog", "scheming pope"]
```

现在让我们编写一个自己的 `length`，命名 `length'`（这里的 `_` 意为无需使用的变量）：

```
1 length' xs = sum [1 | _ <- xs]
```

由于字符串是数组，因此我们可以使用列表表达式处理并生产字符串。以下是一个移除所有字符但保留大写字符的函数：

```
1 removeNonUppercase st = [ c | c <- st, c `elem` ['A'..'Z']]
2 removeUppercase st = [ c | c <- st, c `notElem` ['A'..'Z']]
```

## 元组

在某种程度上，元组类似于数组 – 存储若干值至单个变量上。然而有一些基础的差异：数组长度可以无限，元组长度固定；数组中元素类型是同质的，而元组则可以是异质的 *heterogenous*。

对于对元组（当且仅当包含两个元素）有以下操作：

**fst** 获取对元组的第一个元素：

```
1 ghci> fst (8,11)
2 8
3 ghci> fst ("Wow", False)
4 "Wow"
```

**snd** 获取对元组的第二个元素：

```
1 ghci> snd (8,11)
2 11
3 ghci> snd ("Wow", False)
4 False
```

另外一个有意思的函数则是 **zip**，它可以将两个数组按对拼接成对元组的数组

```
1 ghci> zip [1,2,3,4,5] [5,5,5,5,5]
2 [(1,5),(2,5),(3,5),(4,5),(5,5)]
3 ghci> zip [1..5] ["one", "two", "three", "four", "five"]
4 [(1,"one"),(2,"two"),(3,"three"),(4,"four"),(5,"five")]
```

当两个数组的长度不一时，**zip** 则按最短的那个进行对齐，长的数组剩余部分则被丢弃，这是因为 Haskell 是懒加载的缘故。

```
1 ghci> zip [5,3,2,6,2,7,2,5,4,6,6] ["im", "a", "turtle"]
2 [(5,"im"),(3,"a"),(2,"turtle")]
3 ghci> zip [1..] ["apple", "orange", "cherry", "mango"]
4 [(1,"apple"),(2,"orange"),(3,"cherry"),(4,"mango")]
```



## 3 Types and Typeclasses

### 相信类型

通过 `:t` 命令可以得知类型：

```
1 ghci> :t 'a'
2 'a' :: Char
3 ghci> :t True
4 True :: Bool
5 ghci> :t "HELLO!"
6 "HELLO!" :: [Char]
7 ghci> :t (True, 'a')
8 (True, 'a') :: (Bool, Char)
9 ghci> :t 4 == 5
10 4 == 5 :: Bool
```

函数同样拥有类型：

```
1 ghci> :t doubleSmallNumber
2 doubleSmallNumber :: (Ord a, Num a) => a -> a
```

而对于有多个入参的函数而言：

```
1 ghci> addThree x y z = x + y + z
2 ghci> :t addThree
3 addThree :: Num a => a -> a -> a -> a
```

参数由 `->` 符分开，并且入参与返回的类型并无差异，之后我们讨论到为什么是由 `->` 分割而不是 `Int, Int, Int -> Int` 或者其他样式的类型。

接下来是一些常规的类型：

**Int**：对于 32 位的机器而言最大值大概是 2147483647 而最小值则是 -2147483647。

**Integer**：同样也是整数，只不过范围会大很多，而 **Int** 则更高效。

**Float**：单精度。

**Double**：双精度。

**Bool**：布尔值。

**Char**：字符。

### 类型变量

那么 `head` 函数的类型是什么呢？

```
1 ghci> :t head
2 head :: [a] -> a
```

这里的 `a` 则是一个 **类型变量 type variable**，这意味着 `a` 可以是任意类型。这非常像其他语言的泛型，但唯有在 Haskell 中它更为强大，因为它允许我们可以轻易的编写通用的函

数，且不使用任何特定的行为的类型。带有类型变量的函数也被称为 **多态函数 polymorphic functions**。

## Typeclasses 101

`typeclass` 类似于一个接口用于定义一些行为。如果一个类型是 `typeclass` 的一部分，这就意味着它支持并且实现了 `typeclass` 中所描述的行为。

那么 `==` 函数的类型签名是什么呢？

```
1 ghci> :t (==)
2 (==) :: Eq a => a -> a -> Bool
```

### Note

`==` 操作符是一个函数，`+`，`*`，`-`，`/` 以及其它的操作符也都是。如果一个函数只包含特殊字符，那么默认情况下它被认做是一个中缀函数。如果想要检查它的类型，将其传递给另一个函数或作为前缀函数调用它，那么则需要用括号将其包围。

这里有趣的是 `=>` 符号，在该符号之前的被称为一个 **类约束 class constraint**。那么上述的类型声明可以被这么理解：等式函数接受任意两个相同类型的值，并返回一个 `Bool`，而这两个值必须是 `Eq` 类的成员（即类约束）。

`Eq` `typeclass` 提供了一个用于测试是否相等的接口。

而 `elem` 函数则拥有 `(Eq a) => a -> [a] -> Bool` 这样的类型，因为其在数组中使用了 `==` 用于检查元素是否为期望的值。

一些基础的 `typeclasses`：

`Eq` 如上所述。

`Ord` 覆盖了所有标准的比较函数例如 `>`，`<`，`>=` 以及 `<=`。`compare` 函数接受两个同类型的 `Ord` 成员，并返回一个 `ordering`。`Ordering` 是一个可作为 `GT`，`LT` 或是 `EQ` 的类型，分别意为 大于，小于以及等于。

`Show` 可以表示为字符串。

`Read` 有点类似于 `Show` 相反的 `typeclass`，`read` 函数接受一个字符串并返回一个 `Read` 的成员。

```
1 ghci> read "True" || False
2 True
3 ghci> read "8.2" + 3.8
4 12.0
5 ghci> read "5" - 2
6 3
7 ghci> read "[1,2,3,4]" ++ [3]
8 [1,2,3,4,3]
```

但是如果尝试一下 `read "4"` 呢?

```
1 ghci> read "4"
2 <interactive>:1:0:
3   Ambiguous type variable `a' in the constraint:
4     `Read a' arising from a use of `read' at <interactive>:1:0-7
5   Probable fix: add a type signature that fixes these type variable(s)
```

这里 GHCi 告知它并不知道想要返回什么, 通过检查 `read` 的类型签名:

```
1 ghci> :t read
2 read :: Read a => String -> a
```

也就是说其返回的是 `Read` 所约束的类型, 那么如果在之后没有使用到它, 则没有办法知晓其类型。我们可以显式的使用类型注解, 即在表达式后面加上 `::` 与指定的一个类型:

```
1 ghci> read "5" :: Int
2 5
3 ghci> read "5" :: Float
4 5.0
5 ghci> (read "5" :: Float) * 4
6 20.0
7 ghci> read "[1,2,3,4]" :: [Int]
8 [1,2,3,4]
9 ghci> read "(3, 'a')" :: (Int, Char)
10 (3, 'a')
```

大多数表达式可以被编译器推导出其类型, 但是有时编译器并不知道返回值的类型, 例如 `read "5"` 时该为 `Int` 还是 `Float`。那么为了知道其类型, Haskell 会解析 `read "5"`。然而 Haskell 是一个静态类型语言, 因此它需要在代码编译前知道所有类型。

`Enum` 成员是序列化的有序类型 – 它们可被枚举。`Enum` typeclass 的主要优势是可以使用在列表区间; 它们也同样定义了 `successors` 与 `predecessors`, 即可使用 `succ` 以及 `pred` 函数。在这个类中的类型有: `()`, `Bool`, `Char`, `Ordering`, `Int`, `Integer`, `Float` 以及 `Double`。

`Bounded` 成员拥有上下界:

```
1 ghci> minBound :: Int
2 -2147483648
3 ghci> maxBound :: Char
4 '\1114111'
5 ghci> maxBound :: Bool
6 True
7 ghci> minBound :: Bool
8 False
```

元组的成员如果为 `Bounded` 的一部分, 那么元组也是:

```
1 ghci> maxBound :: (Bool, Int, Char)
2 (True, 2147483647, '\1114111')
```

`Num` 是一个数值类:

```
1 ghci> :t 20
2 20 :: (Num t) => t
3 ghci> 20 :: Int
4 20
5 ghci> 20 :: Integer
6 20
7 ghci> 20 :: Float
8 20.0
9 ghci> 20 :: Double
10 20.0
```

这些类型都在 `Num` typeclass 内。如果检查 `*` 的类型，将会看到：

```
1 ghci> :t (*)
2 (*) :: (Num a) => a -> a -> a
```

`Integral` 同样也是数值 typeclass。

`Floating` 包含 `Float` 与 `Double`。

`fromIntegral` 是一个处理数值的常用函数，而其类型为 `fromIntegral :: (Num b, Integral a) => a -> b`。

## 4 Syntax in Functions

### 模式匹配

一个简单案例：

```
1 sayMe :: (Integral a) => a -> String
2 sayMe 1 = "One!"
3 sayMe 2 = "Two!"
4 sayMe 3 = "Three!"
5 sayMe 4 = "Four!"
6 sayMe 5 = "Five!"
7 sayMe x = "Not between 1 and 5"
```

一个递归案例：

```
1 factorial :: (Integral a) => a ->
2 factorial 0 = 1
3 factorial n = n * factorial (n - 1)
```

模式匹配也可以失败：

```
1 charName :: Char -> String
2 charName 'a' = "Albert"
3 charName 'b' = "Broseph"
4 charName 'c' = "Cecil"
```

当输入并不是期望时：

```
1 ghci> charName 'a'
2 "Albert"
3 ghci> charName 'b'
4 "Broseph"
5 ghci> charName 'h'
6 "*** Exception: tut.hs:(53,0)-(55,21): Non-exhaustive patterns in function charName"
```

即出现了非穷尽的匹配，因此我们总是需要捕获所有模式。

模式匹配也可作用于元组：

```
1 addVectors :: (Num a) => (a, a) -> (a, a) -> (a, a)
2 addVectors a b = (fst a + fst b, snd a + snd b)
```

模式匹配也可作用于列表表达式：

```
1 ghci> let xs = [(1,3), (4,3), (2,4), (5,3), (5,6), (3,1)]
2 ghci> [a+b | (a,b) <- xs]
3 [4,7,6,8,11,4]
```

**Note**

`x:xs` 模式的使用很常见，特别是递归函数。但是包含 `:` 的模式只匹配长度为 1 或更多的数组。

如果希望获取前三个元素以及数组剩余元素，那么可以使用 `x:y:z:zs`，那么这样仅匹配有三个或以上的元素的数组。

其它案例：

```
1 tell :: (Show a) => [a] -> String
2 tell [] = "The list is empty"
3 tell (x:[]) = "The list has one element: " ++ show x
4 tell (x:y:[]) = "The list has two elements: " ++ show x ++ " and " ++ show y
5 tell (x:y:_) = "This list is long. The first two elements are: " ++ show x ++ " and " ++ show y
```

该函数是安全的，因为它考虑到了空数组，以及若干元素数组的情况。

之前通过列表表达式编写了 `length` 函数，现在可以通过模式匹配再加上递归的方式实现一遍：

```
1 length' :: (Num b) => [a] -> b
2 length' [] = 0
3 length' (_ : xs) = 1 + length' xs
```

接下来是实现 `sum`：

```
1 sum' :: (Num a) => [a] -> a
2 sum' [] = 0
3 sum' (x:xs) = x + sum' xs
```

同样还有一种被称为 *as* 模式的，即在模式前添加名称以及 `@` 符号，例如 `xs@(x:y:ys)`，该模式将匹配 `x:y:ys`，同时用户可以轻易的通过 `xs` 来获取整个数组，而无需重复使用 `x:y:ys` 进行表达：

```
1 capital :: String -> String
2 capital "" = "Empty string, whoops!"
3 capital all@(x : xs) = "The first letter of " ++ all ++ " is " ++ [x]
```

最后，用户在模式匹配中不能使用 `++` 符号。

**守护!**

守护是一种检测值的某些属性是否为真，看上去像是 `if` 语句，但是其可读性更强：

```
1 bmiTell :: (RealFloat a) => a -> String
2 bmiTell bmi
3 | bmi <= 18.5 = "You're underweight, you emo, you!"
```

```

4 | bmi <= 25.0 = "You're supposedly normal. Pffft, I bet you're ugly!"
5 | bmi <= 30.0 = "You're fat! Lose some weight, fatty!"
6 | otherwise = "You're a whale, congratulations!"

```

守护是由管道符并接着一个函数名以及函数参数进行定义的。守护本质上就是一个布尔表达式，如果为 `True`，那么其关联的函数体被执行；如果为 `False`，那么检查则会移至下一个守护，以此类推。

大多数时候最后一个守护是 `otherwise`，其被简单的定义为 `otherwise = True` 并捕获所有情况。

当然我们可以使用任意参数的函数来守护：

```

1 bmiTell' :: (RealFloat a) => a -> a -> String
2 bmiTell' weight height
3   | weight / height ^ 2 <= 18.5 = "You're underweight, you emo, you!"
4   | weight / height ^ 2 <= 25.0 = "You're supposedly normal. Pffft, I bet you're ugly!"
5   | weight / height ^ 2 <= 30.0 = "You're fat! Lose some weight, fatty!"
6   | otherwise = "You're a whale, congratulations!"

```

另外我们可以实现自己的 `max` 与 `compare` 函数：

```

1 max' :: (Ord a) => a -> a -> a
2 max' a b
3   | a > b = a
4   | otherwise = b

1 compare' :: (Ord a) => a -> a -> Ordering
2 a `compare'` b
3   | a > b = GT
4   | a == b = EQ
5   | otherwise = LT

```

### Note

我们不仅可以通过引号来调用函数，也可以使用引号来定义他们，有时这样会更加便于阅读。

## Where!?

上一节的 `bmiTell'` 函数中的 `weight / height ^ 2` 被重复了三遍，可以只计算一次并通过名称来绑定计算结果：

```

1 bmiTell'' :: (RealFloat a) => a -> a -> String
2 bmiTell'' weight height
3   | bmi <= 18.5 = "You're underweight, you emo, you!"
4   | bmi <= 25.0 = "You're supposedly normal. Pffft, I bet you're ugly!"
5   | bmi <= 30.0 = "You're fat! Lose some weight, fatty!"

```

```

6 | otherwise = "You're a whale, congratulations!"
7 where
8   bmi = weight / height ^ 2

```

我们在守护的结尾添加了 `where` 并定义了 `bmi` 这个名称，这里定义的名称对整个守护可见，这样就无需再重复同样代码了。那么我们可以进行更多的定义：

```

1 bmiTell'' :: (RealFloat a) => a -> a -> String
2 bmiTell'' weight height
3   | bmi <= skinny = "You're underweight, you emo, you!"
4   | bmi <= normal = "You're supposedly normal. Pffft, I bet you're ugly!"
5   | bmi <= fat    = "You're fat! Lose some weight, fatty!"
6   | otherwise    = "You're a whale, congratulations!"
7 where
8   bmi = weight / height ^ 2
9   skinny = 18.5
10  normal = 25.0
11  fat    = 30.0

```

当然我们可以通过**模式匹配**来进行变量绑定！上面 `where` 中的代码可以改写为：

```

1 ...
2 where
3   bmi = weight / height ^ 2
4   (skinny, normal, fat) = (18.5, 25.0, 30.0)

```

现在让我们编写另一个相当简单的函数用作获取名字首字母：

```

1 initials :: String -> String -> String
2 initials first_name last_name = [f] ++ ". " ++ [l] ++ "."
3 where
4   (f : _) = first_name
5   (l : _) = last_name

```

我们可以直接将模式匹配应用于函数参数。

另外，正如我们可以在 `where` 块中定义约束，我们也可以定义函数：

```

1 calcBmis :: (RealFloat a) => [(a, a)] -> [a]
2 calcBmis xs = [bmi w h | (w, h) <- xs]
3 where
4   bmi weight height = weight / height ^ 2

```

`where` 绑定也可以是嵌套的，这在编写函数中很常见：定义一些辅助函数在函数 `where` 子句，然后这些函数的辅助函数又在其自身的 `where` 子句中。

## Let 的用法

与 `where` 绑定很相似的是 `let` 绑定。前者是一个语法构造器用于在函数的尾部进行变量绑定，这些变量可供整个函数使用，包括守护；而后者则是在任意处绑定一个变量，其自身为



表达式，不过只在作用域生效，因此不能被守护中访问。与 Haskell 其他任意的构造一样，`let` 绑定也可使用模式匹配：

```
1 cylinder :: (RealFloat a) => a -> a -> a
2 cylinder r h =
3   let sideArea = 2 * pi * r * h
4       topArea = pi * r ^ 2
5   in sideArea + 2 * topArea
```

这里的结构是 `let <bindings> in <expression>`。在 `let` 中定义的名称可以在 `in` 之后的表达式中访问。这里同样要注意缩进。现在看来 `let` 仅仅将绑定提前，与 `where` 的作用无异。

不同点在于 `let` 绑定是表达式自身，而 `where` 仅为语法构造。还记得之前提到过的 `if else` 语句是表达式，可以在任意处构造：

```
1 ghci> [if 5 > 3 then "Woo" else "Boo", if 'a' > 'b' then "Foo" else "Bar"]
2 ["Woo", "Bar"]
3 ghci> 4 * (if 10 > 5 then 10 else 0) + 2
4 42
```

那么 `let` 绑定也可以：

```
1 ghci> 4 * (let a = 9 in a + 1) + 2
2 42
```

同样可以在当前作用域引入函数：

```
1 ghci> [let square x = x * x in (square 5, square 3, square 2)]
2 [(25,9,4)]
```

如果想要绑定若干变量，我们显然不能再列上对齐它们。这就是为什么需要用分号进行分隔：

```
1 ghci> (let a = 100; b = 200; c = 300 in a*b*c, let foo="Hey "; bar = "there!" in foo ++ bar)
2 (6000000,"Hey there!")
```

正如之前提到的，可以将模式匹配应用于 `let` 绑定：

```
1 ghci> (let (a,b,c) = (1,2,3) in a+b+c) * 100
2 600
```

当然也可以将 `let` 绑定置入列表表达式中：

```
1 calcBmis' :: (RealFloat a) => [(a, a)] -> [a]
2 calcBmis' xs = [bmi | (w, h) <- xs, let bmi = w / h ^ 2]
```

将 `let` 置入列表表达式中类似于一个子句，不过它不会对列表进行筛选，而仅仅绑定名称。该名称可被列表表达式的输出函数可见（即在符号 `|` 前的部分），以及所有的子句，以及绑定后的部分。因此我们可以让函数继续进行筛选：

```
1 calcBmis'' :: (RealFloat a) => [(a, a)] -> [a]
2 calcBmis'' xs = [bmi | (w, h) <- xs, let bmi = w / h ^ 2, bmi >= 25.0]
```

我们不能在 `(w, h) <- xs` 中使用 `bmi`，因为它在 `let` 绑定之前。

在列表表达式中使用 `let` 绑定可以省略 `in` 的那部分，这是因为名称的可视范围已经被预定义好了。不过我们还是可以在一个子句中使用 `let in` 绑定，该名称仅可在该子句中可见。在 `GHCi` 中定义函数与常数时，`in` 部分同样也可以省略。如果这么做了，那么该名称可以被整个交互过程中可见。

```
1 ghci> let zoot x y z = x * y + z
2 ghci> zoot 3 9 2
3 29
4 ghci> let boot x y z = x * y + z in boot 3 4 2
5 14
6 ghci> boot
7 <interactive>:1:0: Not in scope: `boot'
```

## Case 表达式

以下两端代码表达的是同样一件事，它们互为可替换的。

```
1 head' :: [a] -> a
2 head' [] = error "No head for empty lists!"
3 head' (x:_) = x

1 head' :: [a] -> a
2 head' xs = case xs of [] -> error "No head for empty lists!"
3               (x:_) -> x
```

正如所见的那样，`case` 表达式的语法特别简单：

```
1 case expression of pattern -> result
2 pattern -> result
3 pattern -> result
4 ...
```

`expression` 与模式匹配。模式匹配的行为正如预期那样：首个匹配上表达式的那个模式将被使用。如果直到最后都没有合适的模式被找到，那么将会抛出运行时错误。

函数参数的模式匹配只能在定义函数时完成，而 `case` 表达式则可以在任意处使用。例如：

```
1 describeList :: [a] -> String
2 describeList xs =
3   "The list is " ++ case xs of
4     [] -> "empty."
5     [x] -> "a singleton list."
6     xs -> "a longer list."
```

`case` 表达式可以对表达式中间的模型内容进行模式匹配。函数定义中的模式匹配是 `case` 表达式的语法糖，因此我们也可以这样定义：

```
1 describeList' :: [a] -> String
2 describeList' xs = "The list is " ++ what xs
3 where
4     what [] = "empty."
5     what [x] = "a singleton list."
6     what xs = "a longer list."
```

## 5 Recursion

### 你好递归!

递归对于 Haskell 而言很重要, 因为不同于其他命令式语言, Haskell 中的计算是通过声明某物, 而不是声明如何获取。这就是为什么 Haskell 中没有 while 循环以及 for 循环, 取而代之的则是递归。

### Maximum

`maximum` 函数接受一组可排序的列表 (例如 `Ord` typeclass 的实例), 并返回它们之间最大的那个。

现在让我们看一下如何递归的实现这个函数。我们首先可以确立一个边界条件, 同时声明该列表的最大值等同于列表中的唯一元素, 接着声明如果头大于尾时长列表的头是最大值, 如果尾更大那么继续上述过程:

```
1 maximum' :: (Ord a) => [a] -> a
2 maximum' [] = error "maximum of empty list"
3 maximum' [x] = x
4 maximum' (x : xs)
5   | x > maxTail = x
6   | otherwise = maxTail
7 where
8   maxTail = maximum' xs
```

如上所示, 模式匹配与递归非常的相配! 大多数命令式语言并没有模式匹配, 因此需要编写一堆 if else 声明来测试边界条件。而 Haskell 中仅需令它们成为模版。这里使用了 `where` 绑定来定义 `maxTail` 作为列表尾的最大值。

$$\begin{aligned} &\text{maximum}' [2, 5, 1] = \\ &\text{max } 2 \left( \begin{aligned} &\text{maximum}' [5, 1] = \\ &\text{max } 5 \left( \text{maximum}' [1] = 1 \right) \end{aligned} \right) \end{aligned}$$

## 更多的递归函数

让我们使用递归再来实现一些函数。首先是 `replicate`，接受一个 `Int` 以及一些元素，返回一个列表拥有若干重复的元素。

```
1 replicate' :: (Num i, Ord i) => i -> a -> [a]
2 replicate' n x
3   | n <= 0 = []
4   | otherwise = x : replicate' (n - 1) x
```

这里使用守护而不是模式是因为需要测试一个布尔值条件。

### Note

`Num` 并不是 `Ord` 的子类，也就是说一个数值的组成并不依赖于排序。这就是为什么在做加法或减法或比较时，需要同时指定 `Num` 与 `Ord` 的类约束。

接下来是实现 `take`：

```
1 take' :: (Num i, Ord i) => i -> [a] -> [a]
2 take' n _
3   | n <= 0 = []
4 take' _ [] = []
5 take' n (x : xs) = x : take' (n - 1) xs
```

注意这里使用了 `_` 来匹配列表，因为我们并不关心列表里面的情况；同时，我们使用了一个守护，但是并没有 `otherwise` 部分，这意味着如果 `n` 大于 0 的情况下，匹配将会失败并跳转到下一个匹配。第二个匹配指明如果尝试从空列表中提取任何元素，返回空列表。第三个模式将一个列表分割成一个头与一个尾，接着将从一个列表中获取 `n` 个元素相等于拥有 `x` 头与一个尾视作一个列表获取 `n-1` 元素。

```
1 take' :: (Num i, Ord i) => i -> [a] -> [a]
2 take' n _
3   | n <= 0 = []
4 take' _ [] = []
5 take' n (x : xs) = x : take' (n - 1) xs
```

接下来是 `reverse` 函数：

```
1 reverse' :: [a] -> [a]
2 reverse' [] = []
3 reverse' (x : xs) = reverse' xs ++ [x]
```

由于 Haskell 支持无限列表，`reverse` 并没有一个真正的边界检查，但是如果不这么做，那么则会一直计算下去或者生产出一个无限的数据结构，类似于无限列表。无限列表的好处是我们可以任意处进行截断。然后是 `repeat` 函数，其返回一个无限列表：

```
1 repeat' :: a -> [a]
2 repeat' x = x:repeat' x
```

接下来是 `zip` 函数：

```
1 zip' :: [a] -> [b] -> [(a, b)]
2 zip' _ [] = []
3 zip' [] _ = []
4 zip' (x : xs) (y : ys) = (x, y) : zip' xs ys
```

最后一个是 `elem` 函数：

```
1 elem' :: (Eq a) => a -> [a] -> Bool
2 elem' a [] = False
3 elem' a (x : xs)
4   | a == x = True
5   | otherwise = a `elem'` xs
```

## 快排!

这里是主要算法：排序列表是这样的一个列表，它包含所有小于（或等于）前面的列表头的值（这些值都是排序过的），然后是中间的列表头然后是所有大于列表头的值（它们也是排序过的）。注意定义中两次提到了排序，那么我们将进行两次递归！同样也注意我们使用的是动词 *is* 在算法中进行定义，而不是做这个，做那个，再做另一个... 这就是函数式编程的魅力：

```

1 quicksort :: (Ord a) => [a] -> [a]
2 quicksort [] = []
3 quicksort (x:xs) =
4   let
5     smallerSorted = quicksort [a | a <- xs, a <= x]
6     biggerSorted = quicksort [a | a <- xs, a > x]
7   in
8     smallerSorted ++ [x] ++ biggerSorted

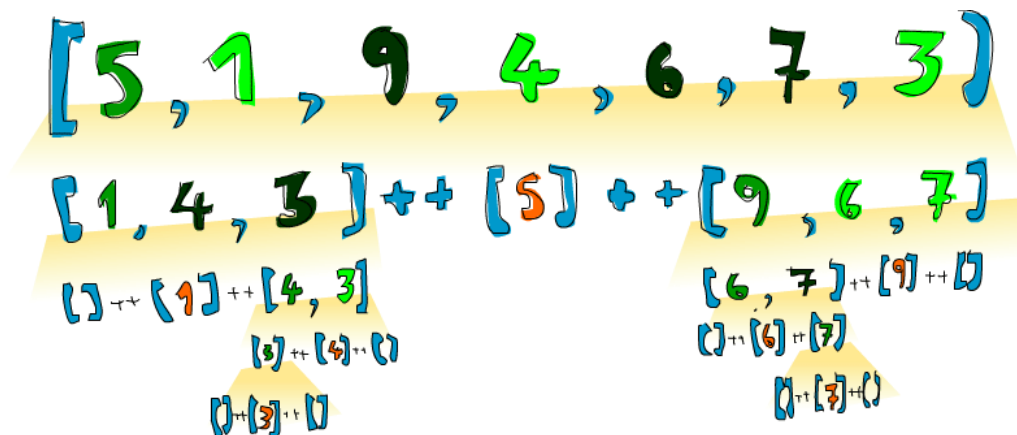
```

测试：

```

1 ghci> quicksort [10,2,5,3,1,6,7,4,2,3,4,8,9]
2 [1,2,2,3,3,4,4,5,6,7,8,9,10]
3 ghci> quicksort "the quick brown fox jumps over the lazy dog"
4 " abcdeeeefghhijklmnoooopqrrsttuuvvwxyz"

```



## 6 Higher Order Functions

### 柯里化函数

在 Haskell 中每个函数实质上仅接受一个参数。那么迄今为止定义的那么多函数是怎么接受多个参数的呢？这就是柯里化函数 **curried functions**。

```
1 ghci> max 4 5
2 5
3 ghci> (max 4) 5
4 5
```

两个参数间用空格间隔就是简单的**函数应用 function application**。空格类似于一个操作符，其拥有最高的优先级。例如 `max`，其签名为 `max :: (Ord a) => a -> a -> a`，可以被重写为 `max :: (Ord a) => a -> (a -> a)`，可以这么理解：`max` 接受一个 `a` 并返回（即 `->`）一个函数，该函数接受一个 `a` 并返回一个 `a`。这就是为什么返回值类型以及函数的参数都是由箭头符进行分隔的。

那么这样做有什么便利？简单来说如果调用一个仅几个参数的函数，我们得到的是一个**部分应用 partially applied**的函数，即一个函数接受的参数与留下未填的参数一样多。

来观测一个简单的函数：

```
1 multThree :: (Num a) => a -> a -> a -> a
2 multThree x y z = x * y * z
```

当使用 `multThree 3 5 9` 或者 `((multThree 3) 5) 9` 时到底发生了什么？首先，`3` 应用至 `multThree`，因为它们由空格进行了分隔（最高优先级）。这就创建了一个接受一个参数的函数，并返回了一个函数。接下来 `5` 被应用至该函数，以此类推。记住我们的函数类型同样也可以重写成 `multThree :: (Num a) => a -> (a -> (a -> a))`。接下来观察：

```
1 ghci> let multTwoWithNine = multThree 9
2 ghci> multTwoWithNine 2 3
3 54
4 ghci> let multWithEighteen = multTwoWithNine 2
5 ghci> multWithEighteen 10
6 180
```

调用函数时输入不足的参数，实际上实在创造新的函数。那么如果希望创建一个函数接受一个值并将其与 `100` 进行比较呢？

```
1 compareWithHundred :: (Num a, Ord a) => a -> Ordering
2 compareWithHundred x = compare 100 x
```

如果带着 `99` 调用它，返回一个 `GT`。注意 `x` 同时位于等式的右侧。那么调用 `compare 100` 返回的是什么呢？它返回一个接受一个数值参数并将其与 `100` 进行比较的函数。现在将其重写：

```
1 compareWithHundred :: (Num a, Ord a) => a -> Ordering
2 compareWithHundred = compare 100
```



类型声明仍然相同，因为 `compare 100` 返回一个函数。`compare` 的类型是 `(Ord a) -> a -> (a -> Ordering)`，带着 `100` 调用它返回一个 `(Num a, Ord a) => a -> Ordering`。这里额外的类约束溜走了，这是因为 `100` 同样也是 `Num` 类的一部分。

中缀函数同样可以通过使用分割被部分应用。要分割中缀函数，只需将其用圆括号括起来，并只在一侧提供参数：

```
1 divideByTen :: (Floating a) => a -> a
2 divideByTen = (/10)
```

调用 `divideByTen 200` 等同于 `200 / 10`，等同于 `(/10) 200`。

那么如果在 `GHCI` 中尝试 `multThree 3 4` 而不是通过 `let` 将其与名称绑定，或是将其传递至另一个函数呢？

```
1 ghci> multThree 3 4
2 <interactive>:1:0:
3   No instance for (Show (t -> t))
4     arising from a use of `print' at <interactive>:1:0-12
5   Possible fix: add an instance declaration for (Show (t -> t))
6   In the expression: print it
7   In a 'do' expression: print it
```

`GHCI` 会提示我们表达式生成了一个类型为 `a -> a` 的函数，但是并不知道该如何将其打印至屏幕。函数并不是 `Show` `typeclass` 的实例，因此我们并不会得到一个函数的展示。

## 来点高阶函数

函数可以接受函数作为其参数，也可以返回函数。

```
1 applyTwice :: (a -> a) -> a -> a
2 applyTwice f x = f (f x)
```

首先注意的是类型声明。之前我们是不需要圆括号的，因为 `->` 是自然地右结合。然而在这里却是强制性的，它们表明了第一个参数是一个接受某物并返回某物的函数，第二个参数同上所述。我们可以用柯里化函数的方式来进行解读，不过为了避免头疼，我们仅需要说该函数接受两个参数并返回一个值。这里第一个参数是一个函数（即类型 `a -> a`），而第二个参数则是 `a`。

函数体非常的简单，仅需要使用参数 `f` 作为一个函数，通过一个空格将 `x` 应用至其，接着再应用一次 `f`。

```
1 ghci> applyTwice (+3) 10
2 16
3 ghci> applyTwice (++) "HAHA" "HEY"
4 "HEY HAHA HAHA"
5 ghci> applyTwice ("HAHA " ++) "HEY"
6 "HAHA HAHA HEY"
```

```

7 ghci> applyTwice (multThree 2 2) 9
8 144
9 ghci> applyTwice (3:) [1]
10 [3,3,1]

```

可以看到单个高阶函数可以被用以多种用途。而在命令式编程中，通常使用的是 `for` 循环、`while` 循环、将某物设置为一个变量、检查其状态等等，为了达到某些行为，还需要用接口将其封装，类似于函数；而函数式编程则使用高阶函数来抽象出相同的模式。

现在让我们实现一个名为 `flip` 的标准库已经存在的函数，其接受一个函数并返回一个类似于原来函数的函数，仅前两个参数被翻转。简单的实现：

```

1 flip' :: (a -> b -> c) -> (b -> a -> c)
2 flip' f = g
3   where
4     g x y = f y x

```

观察类型声明，`flip'` 接受一个函数，该函数接受一个 `a` 与 `b`，并返回一个函数，该返回的函数接受一个 `b` 与 `a`。然而默认情况下函数是柯里化的，第二个圆括号是没有必要的，因为 `->` 默认是右结合的。`(a -> b -> c) -> (b -> a -> c)` 等同于 `(a -> b -> c) -> (b -> (a -> c))`，等同于 `(a -> b -> c) -> b -> a -> c`。我们可以用更简单方式来定义该函数：

```

1 flip'' :: (a -> b -> c) -> b -> a -> c
2 flip'' f y x = f x y

```

这里我们利用了函数都是柯里化的便利。当不带参数 `y` 与 `x` 时调用 `flip'' f` 时，它将返回一个 `f`，该函数接受两个参数，只不过它们的位置是翻转的。

```

1 ghci> flip' zip [1,2,3,4,5] "hello"
2 [('h',1),('e',2),('l',3),('l',4),('o',5)]
3 ghci> zipWith (flip' div) [2,2..] [10,8,6,4,2]
4 [5,4,3,2,1]

```

## Maps & filters

`map` 接受一个函数以及一个列表，将该函数应用至列表中的每一个元素中，生产一个新的列表。让我们来看一下类型签名：

```

1 map :: (a -> b) -> [a] -> [b]
2 map _ [] = []
3 map f (x : xs) = f x : map f xs

```

测试：

```

1 ghci> map (+3) [1,3,5,1,6]
2 [4,6,8,4,9]
3 ghci> map (-1) [1,3,5,1,6]

```

```

4
5 <interactive>:2:1: error:
6   • Could not deduce (Num a0)
7     arising from a type ambiguity check for
8     the inferred type for 'it'
9     from the context: (Num a, Num (a -> b))
10    bound by the inferred type for 'it' :
11        forall {a} {b}. (Num a, Num (a -> b)) => [b]
12    at <interactive>:2:1-20
13    The type variable 'a0' is ambiguous
14    These potential instances exist:
15        instance Num Integer -- Defined in 'GHC.Num'
16        instance Num Double  -- Defined in 'GHC.Float'
17        instance Num Float   -- Defined in 'GHC.Float'
18        ...plus two others
19        ...plus one instance involving out-of-scope types
20        (use -fprint-potential-instances to see them all)
21   • In the ambiguity check for the inferred type for 'it'
22     To defer the ambiguity check to use sites, enable AllowAmbiguousTypes
23     When checking the inferred type
24     it :: forall {a} {b}. (Num a, Num (a -> b)) => [b]
25 ghci> map (subtract 1) [1,3,5,1,6]
26 [0,2,4,0,5]
27 ghci> map (++ "!") ["BIFF", "BANG", "POW"]
28 ["BIFF!", "BANG!", "POW!"]
29 ghci> map (replicate 3) [3..6]
30 [[3,3,3], [4,4,4], [5,5,5], [6,6,6]]
31 ghci> map (map (^2)) [[1,2], [3,4,5,6], [7,8]]
32 [[1,4], [9,16,25,36], [49,64]]
33 ghci> map fst [(1,2), (3,5), (6,3), (2,6), (2,5)]
34 [1,3,6,2,2]

```

原书上一章有提到过 `-1` 这样的情况，报错的原因是 Haskell 将 `-1` 识别为负数而不是减法，需要显式调用 `subtract` 才能识别为 `partial` 函数并应用至列表中的各个元素上。

`filter` 接受一个子句（该子句是一个函数，用于告知某物是否为真），以及一个列表，并返回满足该子句的元素列表。类型签名如下：

```

1 filter :: (a -> Bool) -> [a] -> [a]
2 filter _ [] = []
3 filter p (x : xs)
4   | p x = x : filter p xs
5   | otherwise = filter p xs

```

测试：

```

1 ghci> filter (>3) [1,5,3,2,1,6,4,3,2,1]
2 [5,6,4]
3 ghci> filter (==3) [1,2,3,4,5]
4 [3]

```

```

5 ghci> filter even [1..10]
6 [2,4,6,8,10]
7 ghci> let notNull x = not (null x) in filter notNull [[1,2,3],[],[3,4,5],[2,2],[],[],[]]
8 [[1,2,3],[3,4,5],[2,2]]
9 ghci> filter (`elem` ['a'..'z']) "u LaUgH aT mE BeCaUSe I aM diFfeRent"
10 "uagameasadifeent"
11 ghci> filter (`elem` ['A'..'Z']) "i lauGh At You BecAuse u r aLL the Same"
12 "GAYBALLS"

```

将上一章的 `quicksort` 中的列表表达式替换为 `filter` :

```

1 quicksort :: (Ord a) => [a] -> [a]
2 quicksort [] = []
3 quicksort (x : xs) =
4   let smallerSorted = quicksort (filter (<= x) xs)
5       biggerSorted = quicksort (filter (> x) xs)
6   in smallerSorted ++ [x] ++ biggerSorted

```

现在尝试一下找到 100,100 以下最大能被 3829 的值:

```

1 largestDivisible :: (Integral a) => a
2 largestDivisible = head (filter p [100000, 99999 ..])
3 where
4   p x = x `mod` 3829 == 0

```

接下来尝试一下找到所有奇数平方在 10,000 以下的和, 不过首先要介绍一下 `takeWhile` 函数。该函数接受一个子句以及一个列表, 接着从列表头向后遍历, 在子句返回真时返回元素, 一旦子句返回假则结束遍历。可以在 GHCI 上用一行来完成任任务:

```

1 ghci> sum (takeWhile (<10000) (filter odd (map (^2) [1..])))
2 166650

```

当然也可以用列表表达式:

```

1 ghci> sum (takeWhile (<10000) [n^2 | n <- [1..], odd (n^2)])
2 166650

```

接下来一个问题是处理考拉兹猜想 Collatz sequences, 其数学表达为:

$$f(n) = \begin{cases} n/2 & \text{if } n \equiv 0 \pmod{2} \\ 3n + 1 & \text{if } n \equiv 1 \pmod{2} \end{cases}$$

我们希望知道的是: 对于从 1 至 100 的所有数开始, 有多少链的长度是大于 15 的? 首先编写一个函数用于生成链:

```

1 chain :: (Integral a) => a -> [a]
2 chain 1 = [1]
3 chain n
4   | even n = n : chain (n `div` 2)
5   | odd n  = n : chain (n * 3 + 1)

```

因为链的最后一位肯定是 1，也就是边界，那么这就是一个简单的递归函数了。测试：

```
1 ghci> chain 10
2 [10,5,16,8,4,2,1]
3 ghci> chain 1
4 [1]
5 ghci> chain 30
6 [30,15,46,23,70,35,106,53,160,80,40,20,10,5,16,8,4,2,1]
```

看起来能正常工作，接下来就是获取长度：

```
1 numLongChains :: Int
2 numLongChains = length (filter isLong (map chain [1 .. 100]))
3 where
4   isLong xs = length xs > 15
```

我们将 `chain` 函数映射至 `[1..100]` 来获取一个链的列表，接着根据检查长度是否超过 15 的子句来过滤它们、一旦完成过滤，我们就可以看到结果的列表中还剩多少链。

### Note

该函数的类型是 `numLongChains :: Int`，因为历史原因 `length` 返回一个 `Int` 而不是一个 `Num a`。如果我们需要返回一个更通用的 `Num a`，可以对返回的长度使用 `fromIntegral`。

使用 `map`，我们还可以这样做 `map (*) [0..]`，如果不是因为别的原因要解释柯里化以及偏函数是实值，那么可以将其传递至其它函数，或者置入列表中（仅仅不能将其变为字符串）。迄今为止，我们只映射了单参数的函数至列表，例如 `map (*2) [0..]` 类型是 `(Num a) => [a]`，我们同样可以这么做 `map (*) [0..]`。这里将会对列表中的每个数值应用函数 `*`，即类型为 `(Num a) => a -> a -> a`。将一个参数应用于需要两个参数的函数将会返回需要一个参数的函数。如果将 `*` 映射至列表 `[0..]`，得到的则是一个接受单个参数的函数的列表，即 `(Num a) => [a -> a]`。也就是说 `map (*) [0..]` 生产一个这样的列表 `[(0*), (1*), (2*), (3*), (4*), (5*)..]`。

```
1 ghci> let listOfFuns = map (*) [0..]
2 ghci> (listOfFuns !! 4) 5
3 20
```

## Lambdas

构建 `lambda` 的方式是写一个 `\`，接着是由空格分隔的参数，再然后是 `->` 符，最后是函数体。

对于 `numLongChains` 函数而言，不再需要一个 `where` 子句：

```
1 numLongChains' :: Int
2 numLongChains' = length (filter (\xs -> length xs > 15) (map chain [1 .. 100]))
```

lambdas 也是表达式，上述代码中的表达式 `\xs -> length xs > 15` 返回的就是一个函数，其用于告知列表的长度是否超过 15。

对于不熟悉柯里化以及偏函数应用的人而言，通常会在不必要的地方使用 lambdas。例如表达式 `map (+3) [1,6,3,2]` 以及 `map (x-> x + 3) [1,6,3,2]` 是等价的，因为 `(+3)` 以及 `(x-> x + 3)` 皆为接受一个值并加上 3 的函数。

与普通函数一样，lambdas 可以接受任意数量的参数：

```
1 ghci> zipWith (\a b -> (a * 30 + 3) / b) [5,4,3,2,1] [1,2,3,4,5]
2 [153.0,61.5,31.0,15.75,6.6]
```

与普通函数一样，也可以在 lambdas 中进行模式匹配。唯一不同点在于不能在一个参数内定义若干模式，例如对同样一个参数做 `[]` 以及 `(x:xs)` 模式。如果一个模式匹配再 lambda 中失败了，一个运行时错误则会出现，所以需要特别注意在 lambdas 中进行模式匹配！

```
1 ghci> map (\(a,b) -> a + b) [(1,2),(3,5),(6,3),(2,6),(2,5)]
2 [3,8,9,8,7]
```

lambdas 通常都被圆括号包围，除非我们想让它一直延伸到最右。有趣的来了：函数默认情况下是柯里化的，以下两者相等：

```
1 addThree :: (Num a) => a -> a -> a -> a
2 addThree x y z = x + y + z
```

```
1 addThree :: (Num a) => a -> a -> a -> a
2 addThree = \x -> \y -> \z -> x + y + z
```

## fold

一个 fold 接受一个二元函数，一个起始值（可以称其为 accumulator）以及一个需要被 fold 的列表。二元函数接受两个参数，第一个是 accumulator，第二个则是列表中第一个（或最后一个）元素，然后再生产一个新的 accumulator。接着二元函数再次被调用，带着新的 accumulator 以及新的列表中第一个（或最后一个）元素，以此类推。一旦遍历完整个列表，仅剩 accumulator 剩余，即最终答案。

首先让我们看一下 `foldl` 函数，也被称为左折叠 left fold。

改写 `sum` 的实现：

```
1 sum' :: (Num a) => [a] -> a
2 sum' xs = foldl (\acc x -> acc + x) 0 xs
```

如果考虑到函数是柯里化的，那么可以简化实现：

```
1 sum' :: (Num a) => [a] -> a
2 sum' = foldl (+) 0
```

这是因为 lambda 函数  $\backslash acc\ x \rightarrow acc + x$  等同于  $(+)$ ，所以可以省略掉 `xs` 这个参数，因为调用 `foldl1 (+) 0` 将返回一个接受列表的函数。

接下来通过左折叠来实现 `elem` 这个函数：

```
1 elem' :: (Eq a) => a -> [a] -> Bool
2 elem' y ys = foldl1 (\acc x -> if x == y then True else acc) False ys
```

让我们看一下这里究竟做了些什么。这里的起始值与 accumulator 都是布尔值。在处理 `fold` 时，accumulator 与返回值的类型总是要一致的。这里的起始值为 `False`，即假设寻找的值并不在列表中。接着就是检查当前元素是否为需要找到的那个，如果是则将 accumulator 设为 `True`，不是则不变。

`foldr` 类似于左折叠，只不过 accumulator 是从列表右侧开始。

以下是使用右折叠来实现 `map`：

```
1 map' :: (a -> b) -> [a] -> [b]
2 map' f xs = foldr (\x acc -> f x : acc) [] xs
```

如果将  $(+3)$  映射至 `[1,2,3]`，则是从列表右侧开始，获取最后一个元素，即 `3`，再应用函数得出 `6`，接着将其放入 accumulator 头部，即将 `[]` 变为 `6:[]`，这样 `[6]` 现在变成了新的 accumulator（这里的 `:` 就是将元素添加至头部）。

当然了，我们也可以利用左折叠来实现：`map' f xs = foldl1 (\acc x -> acc ++ [f x]) [] xs`，只不过 `++` 函数比 `:` 而言更加昂贵，因此我们通常从一个列表构建一个新的列表时，会使用右折叠。

折叠可以用于实现任意一个想要一次性遍历列表中所有元素的函数，并基于此进行返回。任何时候想要遍历一个列表来返回一些东西时，那么这个时候就很可能需要一个 `fold`。这也是为什么在函数式编程中，连同 `maps` 与 `filters`，`fold` 是最有用的函数类型。

`foldl1` 与 `foldr1` 类似于 `foldl` 与 `foldr`，区别在于不需要提供一个显式的初始值。它们假设首个（或最后的）元素为起始值，接着从临近的元素开始进行 `fold`。由于依赖列表至少有一个元素，空列表的情况下它们会导致运行时错误；而 `foldl` 与 `foldr` 则不会。

现在通过 `fold` 来实现标准库的函数，看看 `fold` 的强大之处：

```
1 maximum' :: (Ord a) => [a] -> a
2 maximum' = foldr1 (\x acc -> if x > acc then x else acc)
3
4 reverse' :: [a] -> [a]
5 reverse' = foldl (\acc x -> x : acc) []
6
7 product' :: (Num a) => [a] -> a
8 product' = foldr1 (*)
9
10 filter' :: (a -> Bool) -> [a] -> [a]
11 filter' p = foldr (\x acc -> if p x then x : acc else acc) []
12
```

```

13 head' :: [a] -> a
14 head' = foldr1 (\x _ -> x)
15
16 last' :: [a] -> a
17 last' = foldl1 (\_ x -> x)

```

`head` 更好的实现当然是模式匹配，这里只是使用 `fold` 进行展示。这里的 `reverse'` 定义很聪明，从左遍历列表，每次将得到的元素插入至 accumulator 的头部。`acc x -> x : acc` 看起来像是：函数，只不过翻转了参数，这也是为什么可以将 `reverse` 函数改写为 `foldl (flip (:)) []`。

`scanl` 与 `scanr` 很像 `foldl` 与 `foldr`，不同之处在于后者用列表来存储 accumulator 的状态变化；同样 `scanl1` 与 `scanr1` 类似于 `foldl1` 与 `foldr1`。

```

1 ghci> scanl (+) 0 [3,5,2,1]
2 [0,3,8,10,11]
3 ghci> scanr (+) 0 [3,5,2,1]
4 [11,8,3,1,0]
5 ghci> scanl1 (\acc x -> if x > acc then x else acc) [3,4,5,3,7,9,2,1]
6 [3,4,5,5,7,9,9,9]
7 ghci> scanl (flip (:)) [] [3,2,1]
8 [], [3], [2,3], [1,2,3]

```

Scans 可以被认为是一个函数可被 `fold` 的过程监控。让我们回答这样一个问题：需要多少个元素才能让所有自然数的根之和超过 1000？获取所有自然数平方根只需要 `map sqrt [1..]`，那么想要和，可以使用 `fold`，但是又因为我们感兴趣的是求和的这个过程，那么这里就可以使用 `scan`。一旦完成了 `scan`，我们就可以看到有多少和是少于 1000 的了。

```

1 sqrtSums :: Int
2 sqrtSums = length (takeWhile (<1000) (scanl1 (+) (map sqrt [1..]))) + 1

```

```

1 ghci> sqrtSums
2 131
3 ghci> sum (map sqrt [1..131])
4 1005.0942035344083
5 ghci> sum (map sqrt [1..130])
6 993.6486803921487

```

这里使用了 `takeWhile` 而不是 `filter`，因为后者不能作用于无限列表。尽管我们知道列表是递增的，而 `filter` 并不知道，因此这里的 `takeWhile` 在第一个 `sum` 大于 1000 时将截断 `scan` 列表。

## 通过 \$ 符号进行函数应用

接下来让我们看一下 `$` 函数，也被称为函数应用 *function application*。首先来看一下它的定义：



```

1 ($) :: (a -> b) -> a -> b
2 f $ x = f x

```

大多数时候，它是一个便捷的函数使得无需再写很多括号。考虑一下表达式 `sum (map sqrt [1..130])`，因为 `$` 拥有最低的优先级，那么可以将该表达式重写为 `sum $ map sqrt [1..130]`，省了很多键盘敲击！当遇到一个 `$`，在其右侧的表达式会作为参数应用于左边的函数。那么 `sqrt 3 + 4 + 9` 呢？这会将 9, 4 以及 3 的平方根。那么如何得到  $3 + 4 + 9$  的平方根呢，需要 `sqrt (3 + 4 + 9)`，那么如果使用 `$` 可以改为 `sqrt $ 3 + 4 + 9` 因为 `$` 在所有操作符中的优先级最低。这就是为什么可以想象一个 `$` 相当于分别在其右侧以及等式的最右侧编写了一个隐形的圆括号。

那么 `sum (filter (> 10) (map (*2) [2..10]))` 呢？由于 `$` 是右结合的，`f (g (z x))` 就相当于 `f $ g $ z x`。那么用 `$` 重写便是 `sum $ filter (>10) $ map (*2) [2..10]`。

除开省略掉圆括号，`$` 意味着函数应用可以被视为另一个函数，这样的话就可以将其映射 `map` 至一个列表的函数：

```

1 ghci> map ($ 3) [(4+), (10*), (^2), sqrt]
2 [7.0,30.0,9.0,1.7320508075688772]

```

## 组合函数

数学里的组合函数定义为  $(f \circ g)(x) = f(g(x))$ ，意为组合两个函数来产生一个新的函数，当一个参数 `x` 输入时，相当于带着 `x` 输入至 `g` 再将结果输入至 `f`。

Haskell 中的组合函数基本上等同于数学定义中的一样。通过 `.` 函数来组合函数，其定义为：

```

1 (.) :: (b -> c) -> (a -> b) -> a -> c
2 f . g = \x -> f (g x)

```

注意类型声明。`f` 的参数类型与 `g` 的返回类型一致。

函数组合的用途之一是将函数动态的传递给其他函数。当然了，这点 `lambdas` 也可以做到，但是很多情况下，组合函数更加简洁精炼。假设我们有一个列表的数值，并希望将它们全部转为复数。其中一种办法就是将它们全部取绝对值后再添加负号：

```

1 ghci> map (\x -> negate (abs x)) [5,-3,-6,7,-3,2,-19,24]
2 [-5,-3,-6,-7,-3,-2,-19,-24]

```

注意到了 `lambda` 以及组合函数的样式，我们可以将上述重写为：

```

1 ghci> map (negate . abs) [5,-3,-6,7,-3,2,-19,24]
2 [-5,-3,-6,-7,-3,-2,-19,-24]

```

完美！组合函数是右结合的，所以我们可以一次性进行多次组合。表达式 `f (g (z x))` 等同于 `(f . g . z) x`。了解到这些以后，我们可以将：

```
1 ghci> map (\xs -> negate (sum (tail xs))) [[1..5],[3..6],[1..7]]
2 [-14,-15,-27]
```

转换为：

```
1 ghci> map (negate . sum . tail) [[1..5],[3..6],[1..7]]
2 [-14,-15,-27]
```

那么对于接受多个参数的函数而言呢？如果对它们进行组合函数，就需要偏应用它们使得每个函数仅接受一个参数。`sum (replicate 5 (max 6.7 8.9))` 可以被重写为 `(sum . replicate 5 . max 6.7) 8.9` 或者是 `sum . replicate 5 . max 6.7 $ 8.9`。那么这里实际上发生的是：一个函数接受了 `max 6.7` 所接受的参数，并将 `replicate 5` 应用至其，接着一个函数接受计算的结果并对其求和，最后则是带着 `8.9` 调用该函数。不过正常来讲，人类的读取应为：将 `8.9` 应用至 `max 6.7`，接着应用 `replicate 5`，最后则是求和。如果想用组合函数重写一个又很多圆括号的表达式，可以先把最内存函数的最后一个参数放在 `$` 后面，然后在不带最后一个参数的情况下，组合其他的函数调用，即在函数之间加上 `.`。如果有一个表达式 `replicate 100 (product (map (*3) (zipWith max [1,2,3,4,5] [4,5,6,7,8])))`，那么可以重写为 `replicate 100 . product . map (*3) . zipWith max [1,2,3,4,5] [4,5,6,7,8]`。如果表达式的结尾有三个圆括号，转换为组合函数时，则会有三个组合操作符。

组合函数的另一个常见的用法是以所谓的点自由样式（也称无点样式）来定义函数。拿之前我们写的一个例子：

```
1 sum' :: (Num a) => [a] -> a
2 sum' xs = foldl (+) 0 xs
```

`xs` 在两边都暴露出来。由于有柯里化，我们可以在两边省略 `xs`，因为调用 `foldl (+) 0` 会创建一个接受一个列表的函数。编写函数 `sum' = foldl (+) 0` 就是被称为无点样式。那么以下函数如何转换成无点样式呢？

```
1 fn x = ceiling (negate (tan (cos (max 50 x))))
```

我们没法将 `x` 从等式两边的右侧移除，函数体中的 `x` 后面有圆括号。`cos (max 50)` 显然没有意义。那么将 `fn` 表达为组合函数即可：

```
1 fn = ceiling . negate . tan . cos . max 50
```

漂亮！无点样式可以令用户去思考函数的组合方式，而非数据的传递方式，这样更加的简明。你可以将一组简单的函数组合在一起，使之成为一个负责的函数。不过如果函数过于复杂，再使用无点样式往往会达到反效果。因此构造较长的函数组合链并不好，更好的解决方案则是用 `let` 语句将中间的运算结果绑定一个名字，或者说把问题分解成几个小问题再组合到一起。

本章的 `maps` 与 `filters` 中，我们写了：

```
1 oddSquareSum :: Integer
2 oddSquareSum = sum (takeWhile (<10000) (filter odd (map (^2) [1..])))
```

那么更为函数式的写法则是

```
1 oddSquareSum :: Integer
2 oddSquareSum = sum . takeWhile (<10000) . filter odd . map (^2) $ [1..]
```

如果需要可读性，则可以这样：

```
1 oddSquareSum :: Integer
2 oddSquareSum =
3     let oddSquares = filter odd $ map (^2) [1..]
4         belowLimit = takeWhile (<10000) oddSquares
5     in sum belowLimit
```

## 7 Modules

### 加载模块

Haskell 模块是一系列关联的函数，类型以及 typeclasses 的集合。一个 Haskell 的程序是模块的集合，其中主模块加载若干模块并使用这些模块。

Haskell 的标准库也被分成了若干模块，每个模块所包含的函数与类型服务于某些共同的目的。迄今为止我们所有接触到的函数，类型，typeclasses 都位于 `Prelude` 模块，也是默认加载的模块。

Haskell 加载模块的语法就是 `import <module name>`。现在尝试一下加载 `Data.List` 这个模块：

```
1 import Data.List
2
3 numUniques :: (Eq a) => [a] -> Int
4 numUniques = length . nub
```

当 `import Data.List` 后，所有在 `Data.List` 中的函数在公共命名空间中变为可用，也就是说可以在脚本中任意处调用这些函数。`nub` 是一个定义在 `Data.List` 中的函数，接受一个列表并去除其中的重复元素。将 `length` 与 `num` 组合起来 `length . nub` 产生的函数等同于 `xs -> length (nub xs)`。

在使用 `GHCI` 的时候同样也可以加载模块至全局命名空间内，只需：

```
1 ghci> :m + Data.List
```

如果想要加载多个模块，仅需：

```
1 ghci> :m + Data.List Data.Map Data.Set
```

如果加载的脚本（通过 `:l <xxx script>`）中已经加载过模块了，那么便不再需要通过 `:m +` 进行加载。

如果只是想要从模块中加载几个函数，那么可以这样做：

```
1 import Data.List (nub, sort)
```

如果想加载模块，却不包括某些函数，可以这样：

```
1 import Data.List hiding (nub)
```

为了避免重名，可以使用 `qualified`，譬如这样：

```
1 import qualified Data.Map
```

这样的话如果想要调用 `Data.Map` 的 `filter` 函数时，就必须写 `Data.Map.filter`，这样的话 `filter` 仍然会引用普通的 `filter`。不过每次都要写 `Data.Map` 就很麻烦，因此可以这么写：

```
1 import qualified Data.Map as M
```

现在再要调用 `Data.Map` 的 `filter` 时，仅需 `M.filter`。

可以在这里找到标准库中有哪些模块。

通过 `Hoogle` 可以查看函数的位置，这是一个非常棒的 Haskell 搜索引擎，可以通过名称，模块名称甚至是类型签名来进行搜索。

## Data.List

`Data.List` 模块显然都是关于列表的，它提供了些很有用的函数用于列表处理。我们已经学习到了一些（例如 `map` 与 `filter`）这是因为 `Prelude` 已经从 `Data.List` 中加载了不少函数。无需再通过 `qualified import` 导入 `Data.List`，因为它并不会与任何的 `Prelude` 名称重名（除开那些已经被 `Prelude` 从 `Data.List` 中偷走的）。现在让我们看一下其中一些还没有用到过的函数：

`intersperse` 接受一个元素以及一个列表，将该元素插入至列表中每个元素之间：

```
1 ghci> intersperse '.' "MONKEY"
2 "M.O.N.K.E.Y"
3 ghci> intersperse 0 [1..6]
4 [1,0,2,0,3,0,4,0,5,0,6]
```

`intercalate` 接受一个列表以及一个列表的列表，将前者插入至后者之间在打平结果：

```
1 ghci> intercalate " " ["hey", "there", "guys"]
2 "hey there guys"
3 ghci> intercalate [0,0,0] [[1,2,3],[4,5,6],[7,8,9]]
4 [1,2,3,0,0,0,4,5,6,0,0,0,7,8,9]
```

`transpose` 接受一个列表的列表，如果将其视为一个二维的矩阵，那么就是列变为行，行变位列：

```
1 ghci> transpose [[1,2,3],[4,5,6],[7,8,9]]
2 [[1,4,7],[2,5,8],[3,6,9]]
3 ghci> transpose ["hey", "there", "guys"]
4 ["htg","ehu","yey","rs","e"]
```

假设我们有一些多项式  $3x^2 + 5x + 9$ ,  $10x^3 + 9$  以及  $8x^3 + 5x^2 + x - 1$ ，我们希望对它们求和，那么可以使用列表 `[0,3,5,9]`，`[10,0,0,9]` 以及 `[8,5,1,-1]` 来代表它们的系数（分别为  $x^3, x^2, x^1, x^0$  的系数），那么可以这样：

```
1 ghci> map sum $ transpose [[0,3,5,9],[10,0,0,9],[8,5,1,-1]]
2 [18,8,6,17]
```

上述做法就是先转置整个列表 `transpose`，这就代表着每个同幂的系数在一个列表中，再将 `sum` 映射至这个转置后的列表中的每个元素（列表）。

`foldl'` 与 `foldl1'` 相对于它们的懒加载的原版是更严格的版本, 当使用原版的懒加载 `fold` 用于一个很大的列表, 很有可能就会得到一个堆栈溢出的错误。罪魁祸首就是因为 `fold` 的懒加载特性, 累加器的值并不会在 `folding` 时真正的去计算, 相反是在需要结果时才去进行计算 (这也被称为一个型实转换程序 `thunk`)。这发生在每个中间累加器上, 且所有的这些 `thunk` 都会导致堆栈溢出。`strict` 版本由于不是懒加载的, 是真实的计算中间值而不是一直在栈上堆叠。因此当使用懒加载的 `folds` 时遇到了堆栈溢出, 可以尝试一下用它们的严格版本。

`concat` 打平一个列表的列表:

```
1 ghci> concat ["foo", "bar", "car"]
2 "foobarcar"
3 ghci> concat [[3,4,5],[2,3,4],[2,1,1]]
4 [3,4,5,2,3,4,2,1,1]
```

`concatMap` 首先映射一个函数至列表接着将该列表进行 `concat` :

```
1 ghci> concatMap (replicate 4) [1..3]
2 [1,1,1,1,2,2,2,2,3,3,3,3]
```

`and` 接受一个布尔值的列表, 当所有值皆为 `True` 时返回 `True` :

```
1 ghci> and $ map (>4) [5,6,7,8]
2 True
3 ghci> and $ map (==4) [4,4,3,4,4]
4 False
```

`or` 与 `and` 类似, 只不过是任意元素为 `True` 时返回 `True` :

```
1 ghci> or $ map (==4) [2,3,4,5,6,7]
2 True
3 ghci> or $ map (>4) [1,2,3]
4 False
```

`any` 与 `all` 接受一个子句, 然后检查列表中的所有元素, 通常而言我们会使用这两个函数而不是像上面那样先 `map` 接着 `and` 或是 `or`。

```
1 ghci> any (==4) [2,3,5,6,1,4]
2 True
3 ghci> all (>4) [6,9,10]
4 True
5 ghci> all (`elem` ['A'..'Z']) "HEYGUYSwhatsup"
6 False
7 ghci> any (`elem` ['A'..'Z']) "HEYGUYSwhatsup"
8 True
```

`iterate` 接受一个函数以及一个起始值, 将函数应用在这个起始值得到结果后, 再将函数应用至该结果, 以此类推, 返回一个无限列表。

```
1 ghci> take 10 $ iterate (*2) 1
2 [1,2,4,8,16,32,64,128,256,512]
3 ghci> take 3 $ iterate (++ "haha") "haha"
4 ["haha", "hahahaha", "hahahahahaha"]
```

`splitAt` 接受一个数值与一个列表，将列表从数值作为的索引出切分为两部分，返回一个包含了切分后两个列表的二元元组：

```
1 ghci> splitAt 3 "heyman"
2 ("hey","man")
3 ghci> splitAt 100 "heyman"
4 ("heyman","")
5 ghci> splitAt (-3) "heyman"
6 ("","heyman")
7 ghci> let (a,b) = splitAt 3 "foobar" in b ++ a
8 "barfoo"
```

`takeWhile` 是一个非常有用的小函数，它从一个列表中从头开始获取元素，直到元素不再满足子句的条件，返回之前所有满足元素的列表：

```
1 ghci> takeWhile (>3) [6,5,4,3,2,1,2,3,4,5,4,3,2,1]
2 [6,5,4]
3 ghci> takeWhile (/=' ') "This is a sentence"
4 "This"
```

假设我们想要所有三次方都小于 10,000 的值之和，将  $(^3)$  映射至  $[1..]$ ，然后应用一个过滤函数再进行求和这是做不到的，因为过滤一个无限长的列表将永远都不会结束。虽然我们不知道列表里的元素是递增的，但是 Haskell 并不知道，因此需要这么做：

```
1 ghci> sum $ takeWhile (<10000) $ map (^3) [1..]
2 53361
```

将  $(^3)$  应用至一个无限列表，然后一旦元素超过 10,000 那么列表将被截断，这样才可以求和。

`dropWhile` 也类似，它扔掉所有子句中条件判断为真的元素，一旦子句返回 `False`，则返回剩余的列表。

```
1 ghci> dropWhile (/=' ') "This is a sentence"
2 " is a sentence"
3 ghci> dropWhile (<3) [1,2,2,2,3,4,5,4,3,2,1]
4 [3,4,5,4,3,2,1]
```

`span` 有点像 `takeWhile`，不过它返回的是一对列表。第一个列表包含了所有 `takeWhile` 所返回的元素，第二个列表则是剩余的部分：

```
1 ghci> let (fw, rest) = span (/=' ') "This is a sentence" in "First word:" ++ fw ++ ", the
   rest:" ++ rest
2 "First word: This, the rest: is a sentence"
```

`break` 则是列表中第一个元素令子句条件为真时，切分并返回一对列表。`break p` 等同于 `span (not . p)`。

```
1 ghci> break (==4) [1,2,3,4,5,6,7]
2 ([1,2,3],[4,5,6,7])
```

```
3 ghci> span (/=4) [1,2,3,4,5,6,7]
4 ([1,2,3],[4,5,6,7])
```

使用 `break` 时，第一个满足条件的元素将会放在第二个列表的头部。

`sort` 则是对一个列表排序。元素的类型必须属于 `Ord` typeclass:

```
1 ghci> sort [8,5,3,2,1,6,4,2]
2 [1,2,2,3,4,5,6,8]
3 ghci> sort "This will be sorted soon"
4 "Tbdeehiillnooorssstw"
```

`group` 接受一个列表，并将相邻的相同的元素组成子列表:

```
1 ghci> group [1,1,1,1,2,2,2,2,3,3,2,2,2,5,6,7]
2 [[1,1,1,1],[2,2,2,2],[3,3],[2,2,2],[5],[6],[7]]
```

如果在 `group` 一个列表之前进行排序，那么我们就可以知道每个元素在列表中出现了多少次:

```
1 ghci> map (\l@(x:xs) -> (x,length l)) . group . sort $ [1,1,1,1,2,2,2,2,3,3,2,2,2,5,6,7]
2 [(1,4),(2,7),(3,2),(5,1),(6,1),(7,1)]
```

`inits` 与 `tails` 类似于 `init` 与 `tail`，不同之处在于前者们会递归的应用至一个列表直到空:

```
1 ghci> inits "w00t"
2 ["","w","w0","w00","w00t"]
3 ghci> tails "w00t"
4 ["w00t","00t","0t","t",""]
5 ghci> let w = "w00t" in zip (inits w) (tails w)
6 [(("","w00t"),("w","00t"),("w0","0t"),("w00","t"),("w00t",""))]
```

尝试一下使用 `fold` 来实现查找子列表:

```
1 search :: (Eq a) => [a] -> [a] -> Bool
2 -- search needle haystack =
3 --   let nlen = length needle
4 --   in foldl (\acc x -> if take nlen x == needle then True else acc) False (tails haystack)
5 search needle haystack =
6   let nlen = length needle
7   in foldl (\acc x -> (take nlen x == needle) || acc) False (tails haystack)
```

首先调用 `tails` 来处理需要查找的列表，接着查找每个 `tail` 直到该 `tail` 的头是期望找到的。

上述代码的行为实际上就是 `isInfixOf` 函数，该函数在一个列表中查找子列表，在发现了该子列表时返回 `True` :

```
1 ghci> "cat" `isInfixOf` "im a cat burglar"
2 True
3 ghci> "Cat" `isInfixOf` "im a cat burglar"
```



```

4 False
5 ghci> "cats" `isInfixOf` "im a cat burglar"
6 False

```

`isPrefixOf` 与 `isSuffixOf` 则是从前往后进行查找与从后往前进行查找：

```

1 ghci> "hey" `isPrefixOf` "hey there!"
2 True
3 ghci> "hey" `isPrefixOf` "oh hey there!"
4 False
5 ghci> "there!" `isSuffixOf` "oh hey there!"
6 True
7 ghci> "there!" `isSuffixOf` "oh hey there"
8 False

```

`elem` 与 `notElem` 则是查找元素是否在列表中。

`partition` 接受一个列表以及一个子句，返回一对列表，第一个列表包含了所有符合子句条件的元素，其余的在第二个列表：

```

1 ghci> partition (`elem` ['A'..'Z']) "BOBSidneyMORGANeddy"
2 ("BOBMORGAN","sidneyeddy")
3 ghci> partition (>3) [1,3,5,6,3,2,1,0,3,7]
4 ([5,6,7],[1,3,3,2,1,0,3])

```

理解 `partition` 有别于 `span` 与 `break` 是非常重要的：

```

1 ghci> span (`elem` ['A'..'Z']) "BOBSidneyMORGANeddy"
2 ("BOB","sidneyMORGANeddy")

```

因为 `span` 与 `break` 两者是一旦元素满足子句条件便停止，而 `partition` 则是根据子句判断遍历整个列表。

`find` 接受一个子句以及一个列表，返回首个满足子句条件的元素。不过它返回的元素是被包含在 `Maybe` 值之中。我们将会在下一节深入讲解代数数据类型，不过现在我们需要知道的是：一个 `Maybe` 值可以是 `Just something` 或是 `Nothing`。这就像是一个列表既可以是空列表也可以是带有元素的列表，而 `Maybe` 即可以是无元素又可以是一个元素。

```

1 ghci> find (>4) [1,2,3,4,5,6]
2 Just 5
3 ghci> find (>9) [1,2,3,4,5,6]
4 Nothing
5 ghci> :t find
6 find :: (a -> Bool) -> [a] -> Maybe a

```

这里需要注意的是 `find` 的类型，它返回的是 `Maybe a`。

`elemIndex` 有点像 `elem`，不过它返回的不是布尔值，而是一个由 `Maybe` 包含的索引，如果元素不在列表中，则返回 `Nothing`：

```

1 ghci> :t elemIndex
2 elemIndex :: Eq a => a -> [a] -> Maybe Int

```

```

3 ghci> 4 `elemIndex` [1,2,3,4,5,6]
4 Just 3
5 ghci> 10 `elemIndex` [1,2,3,4,5,6]
6 Nothing

```

`elemIndices` 类似于 `elemIndex`，不过它返回的是索引的列表，因为是列表，所以即使找不到元素也可以返回一个空列表，这样就不需要一个 `Maybe` 类型了：

```

1 ghci> ' ' `elemIndices` "Where are the spaces?"
2 [5,9,13]

```

`finxIndex` 像 `find`，不过返回的是索引：

```

1 ghci> findIndex (==4) [5,3,2,1,6,4]
2 Just 5
3 ghci> findIndex (==7) [5,3,2,1,6,4]
4 Nothing
5 ghci> findIndices (`elem` ['A'..'Z']) "Where Are The Caps?"
6 [0,6,10,14]

```

我们已经尝试过了 `zip` 与 `zipWith`，那么对于多个列表就可以使用 `zip3`，`zip4` 等等，以及 `zipWith3`，`zipWith4` 等等，这里最高可以是 `zip` 七个列表。不过有更好的办法可以 `zip` 无穷多个列表，只不过我们现有的知识暂时还没有办法到那儿。

```

1 ghci> zipWith3 (\x y z -> x + y + z) [1,2,3] [4,5,2,2] [2,2,3]
2 [7,9,8]
3 ghci> zip4 [2,3,3] [2,2,2] [5,5,3] [2,2,2]
4 [(2,2,5,2),(3,2,5,2),(3,2,3,2)]

```

跟普通的 `zip` 一样，计算到最短的列表结束为止。

当处理文件或者是其他地方而来的输入，`lines` 则是一个非常有用的函数：

```

1 ghci> lines "first line\nsecond line\nthird line"
2 ["first line","second line","third line"]

```

`'\n'` 是 unix 的换行符。

`unlines` 则是与 `lines` 相反，将字符串列表变回一个由 `'\n'` 分隔的大字符串：

```

1 ghci> unlines ["first line", "second line", "third line"]
2 "first line\nsecond line\nthird line\n"

```

`words` 与 `unwords` 则是分隔与组装一行字符串：

```

1 ghci> words "hey these are the words in this sentence"
2 ["hey","these","are","the","words","in","this","sentence"]
3 ghci> words "hey these         are     the words in this\nsentence"
4 ["hey","these","are","the","words","in","this","sentence"]
5 ghci> unwords ["hey","there","mate"]
6 "hey there mate"

```

`nub` 我们已经见识过了，移除列表中的重复元素：

```

1 ghci> nub [1,2,3,4,3,2,1,2,3,4,3,2,1]
2 [1,2,3,4]
3 ghci> nub "Lots of words and stuff"
4 "Lots fwrданu"

```

**delete** 接受一个元素以及一个列表，删除列表中第一个出现的元素。

```

1 ghci> delete 'h' "hey there ghang!"
2 "ey there ghang!"
3 ghci> delete 'h' . delete 'h' $ "hey there ghang!"
4 "ey tere ghang!"
5 ghci> delete 'h' . delete 'h' . delete 'h' $ "hey there ghang!"
6 "ey tere gang!"

```

**\** 是一个列表比较函数，类似于集合比较，根据右侧的列表中的元素，移除左边列表中匹配的元素。

```

1 ghci> [1..10] \ [2,5,9]
2 [1,3,4,6,7,8,10]
3 ghci> "Im a big baby" \ "big"
4 "Im a baby"

```

处理 `[1..10] \ [2,5,9]` 类似于 `delete 2 . delete 5 . delete 9 $ [1..10]`。

**union** 类似于集合的并集，它返回的是两个列表的集合，做法是将第二个列表中没有在第一个列表中出现的元素，添加至第一个列表的尾部。注意第二个列表中重复的元素会被移除。

```

1 ghci> "hey man" `union` "man what's up"
2 "hey manwt'sup"
3 ghci> [1..7] `union` [5..10]
4 [1,2,3,4,5,6,7,8,9,10]

```

**intersect** 类似于集合的交集，它返回两个列表同时存在的元素：

```

1 ghci> [1..7] `intersect` [5..10]
2 [5,6,7]

```

**insert** 接受一个元素以及一个可以被排序的列表，将该元素插入到最后一个仍然小于或等于下一个元素的位置，**insert** 将会从列表的头开始，直到找到一个元素大于等于它，接着插入到找到的这个元素之前并结束。

```

1 ghci> insert 4 [3,5,1,2,8,2]
2 [3,4,5,1,2,8,2]
3 ghci> insert 4 [1,3,4,4,1]
4 [1,3,4,4,4,1]

```

如果我们对一个排序后的列表使用 **insert**，那么返回的列表仍然是排序好的：

```

1 ghci> insert 4 [1,2,3,5,6,7]
2 [1,2,3,4,5,6,7]
3 ghci> insert 'g' $ ['a'..'f'] ++ ['h'..'z']
4 "abcdefghijklmnopqrstuvwxy"

```

```

5 ghci> insert 3 [1,2,4,3,2,1]
6 [1,2,3,4,3,2,1]

```

因为历史的原因 `length` , `take` , `drop` , `splitAt` , `!!` 以及 `replicate` 接受的都是 `Int` 类型, 即使它们可以更加的泛用接受 `Integral` 或 `Num` typeclasses (取决于函数本身), 但是修改它们则会影响大量已经存在的代码。这就是为什么在 `Data.List` 中引入了 `genericLength` , `genericTake` , `genericDrop` , `genericSplitAt` , `genericIndex` 以及 `genericReplicate` 作为泛化的版本。例如 `length` 的类型签名是 `length :: [a] -> Int` , 那么如果想要一个列表的均值 `let xs = [1..6] in sum xs / length xs` , 这么做会得到一个类型错误, 因为我们不能对 `Int` 使用 `/` 。而 `genericLength` 的类型签名是 `genericLength :: (Num a) => [b] -> a` 。因为一个 `Num` 可以是一个浮点数, 那么 `let xs = [1..6] in sum xs / genericLength xs` 这样做就没有问题了。

对于 `nub` , `delete` , `union` , `intersect` 以及 `group` 而言, 它们的泛化版本则是 `nubBy` , `deleteBy` , `unionBy` , `intersectBy` 以及 `groupBy` 。它们的不同之处在于前一批使用的函数是 `==` 用于比较, 而后一批则是使用输入的比较函数进行比较。 `group` 等同于 `groupBy (==)` 。

例如, 假设我们有一个描述函数美妙值的列表。我们希望基于正负数, 将其它它们分别分段形成子列表。如果用普通的 `group` 那就只能将相同的相邻元素分组在一起, 而使用 `groupBy` 则可以通过 `By` 函数进行同样类型的判断, 当认为是同样类型的即返回 `True` :

```

1 ghci> let values = [-4.3, -2.4, -1.2, 0.4, 2.3, 5.9, 10.5, 29.1, 5.3, -2.4, -14.5, 2.9, 2.3]
2 ghci> groupBy (\x y -> (x > 0) == (y > 0)) values
3 [[-4.3,-2.4,-1.2],[0.4,2.3,5.9,10.5,29.1,5.3],[-2.4,-14.5],[2.9,2.3]]

```

一个更简单的方法是从 `Data.Function` 中加载 `on` 函数, 其定义:

```

1 ghci> :t on
2 on :: (b -> b -> c) -> (a -> b) -> a -> a -> c

```

那么使用 `(==)` ``on` (> 0)` 返回的函数类似于 `\x y -> (x > 0) == (y > 0)` 。 `on` 在 `By` 函数中运用的很广:

```

1 ghci> groupBy ((==) `on` (> 0)) values
2 [[-4.3,-2.4,-1.2],[0.4,2.3,5.9,10.5,29.1,5.3],[-2.4,-14.5],[2.9,2.3]]

```

确实非常有可读性! 我们可以大声的念出来: 根据元素是否大于零进行相等分组。

`sort` , `insert` , `maximum` 以及 `minimum` 也同样拥有更泛用的版本: `sortBy` , `insertBy` , `maximumBy` 以及 `minimumBy` 。 `sortBy` 的类型签名是 `sortBy :: (a -> a -> Ordering) -> [a] -> [a]` 。如果还记得之前的 `Ordering` 类型可以是 `LT` `EQ` 或 `GT` 。 `sort` 等同于 `sortBy compare` , 因为 `compare` 只接受两个类型为 `Ord` typeclass, 并返回它们的排序关系。

列表可以被比较, 一旦可以比较, 它们则是根据字典顺序进行比较。那么如果我们有一个列表的列表, 并想要不根据内部列表的内容, 而是根据长度来进行排序呢? 这就可以使用 `sortBy` 函数:

```

1 ghci> let xs = [[5,4,5,4,4],[1,2,3],[3,5,4,3],[],[2],[2,2]]
2 ghci> sortBy (compare `on` length) xs
3 [[],[2],[2,2],[1,2,3],[3,5,4,3],[5,4,5,4,4]]

```

这里的 `on` 等同于 `\x y -> length x `compare` length y`。也就是说处理 *By* 函数时，需要等式时可以使用 `(==) `on` something`，需要排序时可以使用 `compare `on` something`。

## Data.Char

`Data.Char` 正如其名，是一个用于处理字符的模块，同样对过滤或映射字符串有很大的帮助，因为字符串本身就是字符的列表。

`Data.Char` 提供了很多关于字符范围的函数。也就是函数接受一个字符，并告诉我们哪些假设为真或为假：

- `isControl` 判断一个字符是否是控制字符
- `isSpace` 判断一个字符是否是空格字符，包括空格，tab，换行符等
- `isLower` 判断一个字符是否为小写
- `isUpper` 判断一个字符是否为大写
- `isAlpha` 判断一个字符是否为字母
- `isAlphaNum` 判断一个字符是否为字母或数字
- `isPrint` 判断一个字符是否可打印
- `isDigit` 判断一个字符是否为数字
- `isOctDigit` 判断一个字符是否为八进制数字
- `isHexDigit` 判断一个字符是否为十六进制数字
- `isLetter` 判断一个字符是否为字母
- `isMark` 判断一个字符是否为 unicode 注意字符（法语）
- `isNumber` 判断一个字符是否为数字
- `isPunctuation` 判断一个字符是否为标点符号
- `isSymbol` 判断一个字符是否为货币符号
- `isSeparator` 判断一个字符是否为 unicode 空格或分隔符
- `isAscii` 判断一个字符是否在 unicode 字母表的前 128 位
- `isLatin1` 判断一个字符是否为 unicode 字母表的前 256 位
- `isAsciiUpper` 判断一个字符是否为大写的 Ascii
- `isAsciiLower` 判断一个字符是否为小写的 Ascii

以上所有判断函数的类型声明都是 `Char -> Bool`，大多数时候我们会用它们来过滤字符串或者其它。例如假设我们做一个程序用来获取用户名，而用户名只能由字母与数字构成，我们可以使用 `Data.List` 里的 `all` 函数与 `Data.Char` 里的子句用于判断用户名是否正确：

```

1 ghci> all isAlphaNum "bobby283"
2 True
3 ghci> all isAlphaNum "eddy the fish!"
4 False

```

同样可以通过 `isSpace` 来模拟 `Data.List` 中的 `words` 函数:

```

1 ghci> words "hey guys its me"
2 ["hey","guys","its","me"]
3 ghci> groupBy ((==) `on` isSpace) "hey guys its me"
4 ["hey"," ","guys"," ","its"," ","me"]

```

呃这看起来像是 `words` 不过我们留下来空格, 那么再用一次 `filter`

```

1 ghci> filter (not . any isSpace) . groupBy ((==) `on` isSpace) $ "hey guys its me"
2 ["hey","guys","its","me"]

```

`Data.Char` 同样提供了一个类似 `Ordering` 的数据类型。`Ordering` 类型可以是 `LT`, `EQ` 或 `GT`, 类似于枚举。而 `GeneralCategory` 同样也是枚举, 它提供了字符所处的范围 `generalCategory :: Char -> GeneralCategory`, 由 31 个种类, 这里只列举部分:

```

1 ghci> generalCategory ' '
2 Space
3 ghci> generalCategory 'A'
4 UppercaseLetter
5 ghci> generalCategory 'a'
6 LowercaseLetter
7 ghci> generalCategory '.'
8 OtherPunctuation
9 ghci> generalCategory '9'
10 DecimalNumber
11 ghci> map generalCategory " \t\nA9?|"
12 [Space,Control,Control,UppercaseLetter,DecimalNumber,OtherPunctuation,MathSymbol]

```

由于 `GeneralCategory` 类型属于 `Eq` typeclass, 也就可以这样进行测试 `generalCategory c == Space`。

`toUpper` 转换一个字符成为大写。空格、数字等不变。

`toLower` 转换一个字符成为小写。

`toTitle` 转换一个字符成为 title-case, 大多数字符就是大写。

`digitToInt` 将一个字符转为 `Int` 值, 这个字符必须在 `'1'..'9'`, `'a'..'f'` 或是 `'A'..'F'` 的范围之内。

```

1 ghci> map digitToInt "34538"
2 [3,4,5,3,8]
3 ghci> map digitToInt "FF85AB"
4 [15,15,8,5,10,11]

```

`intToDigit` 则与 `digitToInt` 相反。

```

1 ghci> intToDigit 15
2 'f'
3 ghci> intToDigit 5
4 '5'

```

`ord` 与 `chr` 函数将字符转换成相应的数值，反之亦然：

```

1 ghci> ord 'a'
2 97
3 ghci> chr 97
4 'a'
5 ghci> map ord "abcdefgh"
6 [97,98,99,100,101,102,103,104]

```

下面是 `encode` 与 `decode`：

```

1 encode :: Int -> String -> String
2 encode shift msg
3     let ords = map ord msg
4         shifted = map (+ shift) ords
5     in map chr shifted

```

```

1 ghci> encode 3 "Heeeeey"
2 "Khhhhh|"
3 ghci> encode 4 "Heeeeey"
4 "Liiiii}"
5 ghci> encode 1 "abcd"
6 "bcde"
7 ghci> encode 5 "Marry Christmas! Ho ho ho!"
8 "Rfw~-%Hmwnxyrfx&%Mt%mt%mt%"

```

```

1 decode :: Int -> String -> String
2 decode shift msg = encode (negate shift) msg

```

```

1 ghci> encode 3 "Im a little teapot"
2 "Lp#d#olwwoh#whdsrw"
3 ghci> decode 3 "Lp#d#olwwoh#whdsrw"
4 "Im a little teapot"
5 ghci> decode 5 . encode 5 $ "This is a sentence"
6 "This is a sentence"

```

## Data.Map

关联列表（通常也被称为字典）是一种用于存储键值对却不保证顺序的列表。

由于 `Data.Map` 中的函数会与 `Prelude` 中的 `Data.List` 冲突，因此需要 `qualified import`：

```

1 import qualified Data.Map as Map

```

现在让我们看一下 `Data.Map` 中有什么好东西！

`fromList` 函数接受一个关联列表，并返回一个 `map`：

```

1 ghci> Map.fromList [("betty","555-2938"),("bonnie","452-2928"),("lucille","205-2928")]
2 fromList [("betty","555-2938"),("bonnie","452-2928"),("lucille","205-2928")]
3 ghci> Map.fromList [(1,2),(3,4),(3,2),(5,5)]
4 fromList [(1,2),(3,2),(5,5)]

```

如果有重复的键出现，那么之前的值会被丢弃，这里是 `fromList` 的类型签名：

```

1 Map.fromList :: (Ord k) => [(k, v)] -> Map.Map k v

```

意为接受一个键值类型 `k` 与 `v` 对的列表，并返回一个 `map`。注意这里的键必须属于 `Eq` 以及 `Ord` typeclass，后者是因为需要将键值安排在树结构中。

在需要键值关联的类型时总是使用 `Data.Map`，除非键不属于 `Ord` typeclass。

`empty` 没有参数，返回一个空 `map`：

```

1 ghci> Map.empty
2 fromList []

```

`insert` 接受一键，一值，并返回 `map`：

```

1 ghci> Map.empty
2 fromList []
3 ghci> Map.insert 3 100 Map.empty
4 fromList [(3,100)]
5 ghci> Map.insert 5 600 (Map.insert 4 200 (Map.insert 3 100 Map.empty))
6 fromList [(3,100),(4,200),(5,600)]
7 ghci> Map.insert 5 600 . Map.insert 4 200 . Map.insert 3 100 $ Map.empty
8 fromList [(3,100),(4,200),(5,600)]

```

我们可以通过空 `map`，`insert` 以及 `foldr` 来实现自己的 `fromList`：

```

1 fromList' :: (Ord k) => [(k, v)] -> Map.Map k v
2 fromList' = foldr (\(k, v) acc -> Map.insert k v acc) Map.empty

```

`null` 检查 `map` 是否为空：

```

1 ghci> Map.null Map.empty
2 True
3 ghci> Map.null $ Map.fromList [(2,3),(5,5)]
4 False

```

`size` 告知 `map` 大小：

```

1 ghci> Map.size Map.empty
2 0
3 ghci> Map.size $ Map.fromList [(2,4),(3,3),(4,2),(5,4),(6,4)]
4 5

```

`singleton` 创建一个只有一对键值的 `map`：

```

1 ghci> Map.singleton 3 9
2 fromList [(3,9)]
3 ghci> Map.insert 5 9 $ Map.singleton 3 9
4 fromList [(3,9),(5,9)]

```



`lookup` 类似于 `Data.List` 的 `lookup`，不过它作用于 `map`。如果找到了则返回 `Just something`，反之 `Nothing`

`member` 是一个子句，接受一个键以及一个 `map`，查找该键是否在 `map` 中：

```
1 ghci> Map.member 3 $ Map.fromList [(3,6),(4,3),(6,9)]
2 True
3 ghci> Map.member 3 $ Map.fromList [(2,5),(4,5)]
4 False
```

`map` 与 `filter` 于列表的函数相似：

```
1 ghci> Map.map (*100) $ Map.fromList [(1,1),(2,4),(3,9)]
2 fromList [(1,100),(2,400),(3,900)]
3 ghci> Map.filter isUpper $ Map.fromList [(1,'a'),(2,'A'),(3,'b'),(4,'B')]
4 fromList [(2,'A'),(4,'B')]
```

`toList` 与 `fromList` 相反：

```
1 ghci> Map.toList . Map.insert 9 2 $ Map.singleton 4 3
2 [(4,3),(9,2)]
```

`keys` 与 `elems` 分别返回键与值的列表。`keys` 等同于 `map fst . Map.toList`，而 `elems` 等同于 `map snd . Map.toList`。

`fromListWith` 这个函数很酷。它作用类似于 `fromList`，不过不会丢弃重复键而是使用函数来决定去留。假设我们有这样一个列表（注意在 `ghci` 中使用 `:{` 与 `:}` 作为多行输入的开头与结尾）：

```
1 phoneBook =
2   [ ("betty", "555-2938")
3     , ("betty", "342-2492")
4     , ("bonnie", "452-2928")
5     , ("patsey", "493-2928")
6     , ("patsey", "943-2929")
7     , ("patsey", "827-9162")
8     , ("lucille", "205-2928")
9     , ("wendy", "939-8282")
10    , ("penny", "853-2492")
11    , ("penny", "555-2111")
12  ]
```

如果使用 `fromList` 来生成一个 `map`，我们将会丢弃一些号码！因此这里我们这么做：

```
1 phoneBookToMap :: (Ord k) => [(k, String)] -> Map.Map k String
2 phoneBookToMap = Map.fromListWith (\number1 number2 -> number1 ++ ", " ++ number2)

1 ghci> Map.lookup "patsey" $ phoneBookToMap phoneBook
2 "827-9162, 943-2929, 493-2928"
3 ghci> Map.lookup "wendy" $ phoneBookToMap phoneBook
4 "939-8282"
5 ghci> Map.lookup "betty" $ phoneBookToMap phoneBook
6 "342-2492, 555-2938"
```

如果遇到一个重复的键，传入的函数用于将这些值结合成为其他值。我们还可以首先将所有值映射为单例的列表后在使用 `++` 来结合：

```
1 phoneBookToMap' :: (Ord k) => [(k, a)] -> Map.Map k [a]
2 phoneBookToMap' = Map.fromListWith (++) . map (\(k, v) -> (k, [v]))
```

```
1 ghci> Map.lookup "patsy" $ phoneBookToMap' phoneBook
2 Just ["827-9162", "943-2929", "493-2928"]
```

非常的精妙！另一个用例则是如果我们的一个数值关联的列表创建 `map`，当遇到重复键时，我们希望较大的键的值能保留：

```
1 ghci> Map.fromListWith max [(2,3),(2,5),(2,100),(3,29),(3,22),(3,11),(4,22),(4,15)]
2 fromList [(2,100),(3,29),(4,22)]
```

或者将重复键的值相加：

```
1 ghci> Map.fromListWith (+) [(2,3),(2,5),(2,100),(3,29),(3,22),(3,11),(4,22),(4,15)]
2 fromList [(2,108),(3,62),(4,37)]
```

`insertWith` 则是插入一个键值对进入 `map`，如果键重复则使用传递的函数进行判断：

```
1 ghci> Map.insertWith (+) 3 100 $ Map.fromList [(3,4),(5,103),(6,339)]
2 fromList [(3,104),(5,103),(6,339)]
```

根据官网的介绍现在最好是用 `Data.Map.Strict`（2023/7/14）。

## Data.Set

由于 `Data.Set` 中的名称会与 `Prelude` 与 `Data.List` 重复，因此需要 `qualified import`：

```
1 import qualified Data.Set as Set
```

`fromList` 将列表转换为集合：

```
1 ghci> let set1 = Set.fromList text1
2 ghci> let set2 = Set.fromList text2
3 ghci> set1
4 fromList " ?AIRadefhijlmnorstuy"
5 ghci> set2
6 fromList " !Tabcdefghilmnorstuvwxy"
```

`intersection` 取交集：

```
1 ghci> Set.intersection set1 set2
2 fromList " adefhilmnorstuy"
```

`difference` 取在第一个集合却不在第二个集合的元素：

```
1 ghci> Set.difference set1 set2
2 fromList " ?AIRj"
3 ghci> Set.difference set2 set1
4 fromList " !Tbcgyvw"
```

union 取并集:

```
1 ghci> Set.union set1 set2
2 fromList " !.?AIRTabcdefghijklmnorstuvwy"
```

null , size , member , empty , singleton , insert 以及 delete 函数与预期一致:

```
1 ghci> Set.null Set.empty
2 True
3 ghci> Set.null $ Set.fromList [3,4,5,5,4,3]
4 False
5 ghci> Set.size $ Set.fromList [3,4,5,3,4,5]
6 3
7 ghci> Set.singleton 9
8 fromList [9]
9 ghci> Set.insert 4 $ Set.fromList [9,3,8,1]
10 fromList [1,3,4,8,9]
11 ghci> Set.insert 8 $ Set.fromList [5..10]
12 fromList [5,6,7,8,9,10]
13 ghci> Set.delete 4 $ Set.fromList [3,4,5,4,3,4,5]
14 fromList [3,5]
```

也可以检查子集等:

```
1 ghci> Set.fromList [2,3,4] `Set.isSubsetOf` Set.fromList [1,2,3,4,5]
2 True
3 ghci> Set.fromList [1,2,3,4,5] `Set.isSubsetOf` Set.fromList [1,2,3,4,5]
4 True
5 ghci> Set.fromList [1,2,3,4,5] `Set.isProperSubsetOf` Set.fromList [1,2,3,4,5]
6 False
7 ghci> Set.fromList [2,3,4,8] `Set.isSubsetOf` Set.fromList [1,2,3,4,5]
8 False
```

map 与 filter :

```
1 ghci> Set.filter odd $ Set.fromList [3,4,5,6,7,2,3,4]
2 fromList [3,5,7]
3 ghci> Set.map (+1) $ Set.fromList [3,4,5,6,7,2,3,4]
4 fromList [3,4,5,6,7,8]
```

toList 转为列表, 按字典顺序进行排序:

```
1 ghci> let setNub xs = Set.toList $ Set.fromList xs
2 ghci> setNub "HEY WHATS CRACKALACKIN"
3 " ACEHIKLNIRSTWY"
4 ghci> nub "HEY WHATS CRACKALACKIN"
5 "HEY WATSCRKLIN"
```

## 构建自己的模块

首先我们创建一个名为 `Geometry.hs` 的文件。

当我们提到模块 *exports* 函数，即当导入一个模块时，可以使用该模块导出的函数。  
在模块的开头，需要指定模块名称，接着指定想要导出的函数：

```
1 module Geometry
2   ( sphereVolume,
3     sphereArea,
4     cubeVolume,
5     cubeArea,
6     cuboidArea,
7     cuboidVolume,
8   )
9 where
```

接下来就是定义函数：

```
1 module Geometry
2   ( sphereVolume,
3     sphereArea,
4     cubeVolume,
5     cubeArea,
6     cuboidArea,
7     cuboidVolume,
8   )
9 where
10
11 import Control.Monad
12
13 sphereVolume :: Float -> Float
14 -- sphereVolume radius = (4.0 / 3.0) * pi * (radius ^ 3)
15 sphereVolume = (* (4.0 / 3.0 * pi)) . (^ 3)
16
17 sphereArea :: Float -> Float
18 -- sphereArea radius = 4 * pi * (radius ^ 2)
19 sphereArea = (* (4 * pi)) . (^ 3)
20
21 cubeVolume :: Float -> Float
22 -- cubeVolume side = cuboidVolume side side side
23 cubeVolume = join (join cuboidVolume)
24
25 cubeArea :: Float -> Float
26 -- cubeArea side = cuboidArea side side side
27 cubeArea = join (join cuboidArea)
28
29 cuboidVolume :: Float -> Float -> Float -> Float
30 -- cuboidVolume a b c = rectangleArea a b * c
31 cuboidVolume = ((*) .) . rectangleArea
32
33 cuboidArea :: Float -> Float -> Float -> Float
34 cuboidArea a b c = rectangleArea a b * 2 + rectangleArea a c * 2 + rectangleArea c b * 2
35
```

```

36 rectangleArea :: Float -> Float -> Float
37 -- rectangleArea a b = a * b
38 rectangleArea = (*)

```

我们仅需像下面这样就能使用模块：

```

1 import Geometry

```

需要注意的是 `Geometry.hs` 需要在加载模块的程序同一文件夹下。

模块童谣可以被给予等级结构。每个模块可以拥有若干子模块，子模块里还可以有子模块。

首先创建一个名为 `Geometry` 的文件夹，接着是下面的三个文件夹：`Sphere.hs` , `Cuboid.hs` 以及 `Cube.hs`

`Sphere.hs` :

```

1 module Geometry.Sphere
2   ( volume,
3     area,
4   )
5   where
6
7   volume :: Float -> Float
8   volume = (* (4.0 / 3.0 * pi)) . (^ 3)
9
10  area :: Float -> Float
11  area = (* (4 * pi)) . (^ 3)

```

`Cuboid.hs` :

```

1 module Geometry.Cuboid
2   ( volume,
3     area,
4   )
5   where
6
7   volume :: Float -> Float -> Float -> Float
8   volume = ((*) .) . rectangleArea
9
10  area :: Float -> Float -> Float -> Float
11  area a b c = rectangleArea a b * 2 + rectangleArea a c * 2 + rectangleArea c b * 2
12
13  rectangleArea :: Float -> Float -> Float
14  rectangleArea = (*)

```

以及 `Cube.hs` :

```

1 module Geometry.Cube
2   ( volume,
3     area,
4   )
5   where

```

```
6
7 import Control.Monad
8 import Geometry.Cuboid qualified as Cuboid
9
10 volume :: Float -> Float
11 volume = join (join Cuboid.volume)
12
13 area :: Float -> Float
14 area = join (join Cuboid.area)
```

那么在与文件夹同级的目录下就可以这样加载模块中的子模块：

```
1 import Geometry.Sphere
```

为了避免命名冲突可以使用 qualified import：

```
1 import qualified Geometry.Sphere as Sphere
2 import qualified Geometry.Cuboid as Cuboid
3 import qualified Geometry.Cube as Cube
```

## **8 Making Our Own Types and Typeclasses**

WIP

## 9 Input and Output

WIP



## 10 Functionally Solving Problems

WIP

## **11 Functors, Applicative Functors and Monoids**

WIP

## 12 A Fistful of Monads

WIP

## 13 For a Few Monads More

WIP

## 14 Zippers

WIP