

CSCC24: Principles of Programming Languages

Notes

Joshua Concon

University of Toronto Scarborough – Winter 2018

This course is taught by Dr. Albert Yu Cheong Lai. If you find any problems in these notes, feel free to contact me at conconjoshua@gmail.com.

Contents

1	Thursday, January 11, 2017	4
1.1	The Root Cause of this Course	4
1.1.1	Higher-order Functions on Aggregates	4
1.1.2	Combining Forms	5
1.1.3	Example Topic: Evaluation Order	6
1.1.4	Example Topic: Scheme Macros	6
1.1.5	Dynamic and Static Typing	6
1.1.6	Parametric Polymorphism	7
1.1.7	What is "Powerful"? – The Tradeoff	8
2	Thursday, January 18, 2018	9
2.1	Racket	9
2.2	Basic Data Types	9
2.3	Procedures and Functions	9
2.4	Boolean Operations	9
2.5	Number operations	9
2.6	Equality	10
2.7	Definitions	10

2.8	Anonymous Functions	11
2.9	Conditionals	11
2.9.1	and,or as conditionals	11
2.10	Local bindings	12
2.10.1	Recursive local bindings	12
2.11	Recommended Code Layout	12
3	Thursday, January 25, 2018	14
3.1	Pairs and Lists	14
3.2	User-Defined Records	14
3.3	Pattern Matching	14
3.4	Input and Output	15
3.4.1	ports	15
3.5	Sequencing	15
3.6	Mutable Variables	15
3.7	map	16
3.8	filter	16
4	Thursday, February 1, 2018	17
4.1	Scheme (cont'd)	17
4.1.1	foldl	17
4.1.2	foldr	17
4.1.3	Procedure-Call Stack	18
4.1.4	Non-Tail Calls and Tail Calls	18
4.2	Haskell	18
4.2.1	Expressions and Types	18
4.2.2	Definitions	20
4.2.3	Function Applications	20
5	Thursday, February 8, 2018	21
5.1	Haskell (cont'd)	21
5.1.1	Local Definitions For Expressions	21
5.1.2	Local Definitions For Definitions	21
5.1.3	Pattern Matching	21
5.1.4	Guards	22
5.1.5	Local Definitions under Patterns and Guards	22
5.1.6	List Comprehension	22
5.1.7	Algebraic Data Types	23

5.1.8	Parametric Polymorphism	23
5.1.9	Type-Class Polymorphism	23
6	Thursday, February 15, 2018	25
6.1	Haskell (cont'd)	25
6.1.1	Constraint Instances	25
6.1.2	User-Defined Class	25
6.1.3	Auto-Generating Instance Implementations	26
6.1.4	Haskell's Number System	26
6.1.5	Functor	26
6.1.6	Applicatives	27
7	Thursday, March 1, 2018	28
7.1	Haskell (cont'd)	28
7.1.1	Monads	28
7.1.2	Haskell I/O System	30
7.2	Syntax	30
7.2.1	Context-Free Grammar (CFG)	30
7.2.2	Derivation (aka Generation)	31
7.2.3	Backus-Naur Form (BNF)	31
7.2.4	Parse Tree (aka Derivation Tree)	32
7.2.5	Ambiguous Grammar	33
7.2.6	Unambiguous Grammar Example	33
7.2.7	Left Recursive vs Right Recursive	33
7.2.8	Recursive Descent Parsing	34
7.2.9	Recursive Descent Parser Example	34
8	Thursday, March 8, 2018	35
8.1	Parser Library	35

1 Thursday, January 11, 2017

The purpose of this course is to see the trade-offs between various features in programming languages. This course exists because different programming languages have different features, for example, Java has both class-based OOP and auto-garbage collection while C has neither, but C has union types that Java doesn't have. This means rewriting code into a different language isn't necessarily easy. There may be large semantic differences

1.1 The Root Cause of this Course

A guy named "John Backus" gave a lecture for the acceptance for the Turing Award in 1977. He addressed the question, "Can programming be liberated from the von Neumann style?"

Languages then had been only superficial enhancements to the CPU writing 1 word onto memory at a time i.e:

```
1 s := s + a[i]
```

Backus proposed a new direction for programming languages:

- Higher order functions that work on aggregates (a whole list, an array, a dictionary, etc...)
- Combining forms, for example, function composition ($g \circ f$)
- Reasoning by algebra, for example, the associative law for a function
- If you need a state, use coarse-grained state transitions rather than changing only one word at a time. (So passing an old state into a stateless function that does a lot and returns an answer and a new state.)

1.1.1 Higher-order Functions on Aggregates

Note that the notation to apply a function to several parameters is:
(Haskell)

```
1 f x y z
```

(Scheme:)

```
1 (f x y z)
```

So in Haskell:

```
1 fmap f [x0, x1, ...]
```

will compute

```
1 [f x0, f x1, ...]
```

And

```
1 fmap abs [3, -1, 4]
```

Computes

```
1 [3, -1, 4]
```

And

```
1 foldr (+) 0 [3, 1, 4]
```

Computes

```
1 3+(1+(4+0))
```

Note 2 points:

- "on aggregates" means to work on a whole list at once (such as an array or some "container")
- "Higher-order functions" means that some parameters are functions, so different combinations makes the language more customizable.

Java and MATLAB have the former but lack the latter.

1.1.2 Combining Forms

An obvious example is function composition ($g \circ f$).

In Haskell, this is:

```
1 g . f
```

And in Racket (Scheme) this is:

```
1 compose g f
```

For example, the following code computes the 1-norm of your vector.

```
1 foldr (+) 0 . fmap abs
```

There are other combining forms. There is another example in Haskell.

```
1 (f &&& g) x = (f x, g x)
```

The point is that you can combine functions to perform compound tasks, and this type of language is not about shorter code (although it has that side effect), but about working with building blocks.

1.1.3 Example Topic: Evaluation Order

You can define your own logical "and" in Scheme

```
1 (define (my-and b c) (if b c #f))
2 (my-and #f (list-ref '#t #f #t) 10))
```

The second line fails in Scheme, but if typed in the Haskell version, succeeds.

In most languages, parameters are evaluated before passed into the bodies of functions. In Haskell however, parameters are passed as is. Because of this, in Haskell, many short circuiting operators and control constructs are user-definable, and therefore, very customizable.

1.1.4 Example Topic: Scheme Macros

Scheme offers a macro system for user-defined constructs:

```
1 (define-syntax-rule (my-and b c) (if b c #f))
```

Now if we run the following code, it succeeds.

```
1 (my-and #f (list-ref '#t #f #t) 10))
```

The explanation for this is that this is a macro expansion in Scheme, so the parameters are copy-pasted into the macros. This means that there is a downside, for example:

```
1 (define-syntax-rule (double x) (+ x x))
2 (double (* 3 4))
```

The second line spawns two copies of `(* 3 4)` and performs redundant work, while Haskell's version does not. The Upside is that Scheme's macro system offers other flexibilities not shown in this lecture.

1.1.5 Dynamic and Static Typing

In Scheme:

```

1 (if #f 0 (+ 0 "hello"))
2 (if #t 0 (+ 0 "hello"))

```

The first line fails but the second line succeeds. This is because Types are checked dynamically. When running the program, only the code that is actually run is checked.

In Haskell, the following line fails:

```

1 if True then 0 else 0 + "hello"

```

The reason for this is because types are checked statically, without running, over all the code. (If this code is compiled, then at compile time, if interpreted, then at load time, etc.) So the error of adding 0 to "hello".

Food for thought, Java is both compiled and interpreted.

1.1.6 Parametric Polymorphism

In Haskell, we define:

```

1 trio x = [x, x, x]

```

The inferred type is:

```

1 a -> [a]

```

This is analogous to Java's

```

1 <T> LinkedList<T> trio(<T> x)

```

Note: That the following 2 lines are both legal if we have trio defined

```

1 trio 0
2 trio "hello"

```

"Parametric" means: Supposed you have defined d of type $a \mapsto [a]$, Then you would need one test to know what it does. Say we test $dTrue$ and the answer has length 2. Then we can deduce that dx returns $[x, x]$ for all x .

The basic explanation for this is that d cannot vary behaviour by types. Haskell allows type-determined behaviour, but the function type will look like:

```

1 Foo a => a -> [a]

```

1.1.7 What is "Powerful"? – The Tradeoff

"Macro systems, dynamic typing, ... are powerful." This refers to the flexibility for the implementer or the original author.

"Static typing, parametric polymorphism... are powerful." This refers to the predictability for the user or the maintainer.

Programming is a dialectic class struggle between the user and the implementer. Or between the maintainer and the original author.

2 Thursday, January 18, 2018

2.1 Racket

We won't be using just Scheme, we'll be using Racket which is a version of Scheme. Racket is a platform for implementing and using many languages, and Scheme is one of those that come out of the box.

Racket's version of scheme is somewhat different from the standards with regards to function names, and some features. We will cover Racket, but note that these examples and features may fail for standard Scheme.

2.2 Basic Data Types

```
1  #t, #f ;booleans
2  42 ;numbers, can be ints, rational, floats, complex
3  "hello" ; strings
4  #\h ; this is a char of just the letter h
5  'Chrome ;this is a symbol
```

Symbols Symbols are user-defined atomic values. You think of a name, put a single quote in front. Symbols are not strings, you can't perform string operations onto them.

2.3 Procedures and Functions

For example:

```
1  (sin (/ 0.2 2)) ; sine of 1 over 10
```

2.4 Boolean Operations

```
1  (not expr)
2  (and expr expr)
3  (or expr expr)
4  (or) ; gives #f
5  (and) ; gives #t
6  (boolean? expr) ; tests if you have a boolean
```

2.5 Number operations

```

1  +
2  -
3  *
4  /
5  max
6  min
7  -
8  <
9  >
10 <=
11 >=
12 ; these operations can all take multiple operands

```

```

1  number?
2  complex?
3  real?
4  rational?
5  ; these functions test if a number is a certain type

```

2.6 Equality

There are 3 types of equalities.

```

1  eq?

```

This one is Good for booleans and symbols, uses pointer inequality for aggregates like strings and lists, but has complicated rules for numbers.

```

1  eqv?

```

This one has complicated rules for numbers as well, and different from eq? as it treats the floating-points NaN and signed zero differently.

```

1  equal?

```

This one is for structural equality for most aggregates. So comparing contents of an aggregate.

2.7 Definitions

You can define functions and constants, recursion is allowed.

```

1  ; A constant
2  (define my-width/height (/ 4 3))
3  ; a function with 2 parameters
4  (define (my-log base x)
5    (/ (log x) (log base)))

```

2.8 Anonymous Functions

Basically a function without a name, you can either write lambda or use λ in your code.

Example of a function:

```
1 (lambda (base x) (/ (log x) (log base)))
```

Example of the same function being used

```
1 ((lambda (base x) (/ (log x) (log base))) 128 2)
```

So in this case, base passed in as 128 and x as 2

2.9 Conditionals

If-then-else conditions look like the following

```
1 (if test then-expr else-expr)
```

Test can be non-boolean, and this will be treated as true.

Multiple conditions are as such:

```
1 (cond
2   [(> x y) (sin x)]
3   [(< x y) (cos y)]
4   [else 0])
```

This is if $x > y$ then $\sin x$ else if $x < y$ then $\cos y$ else 0. Test results can be non-boolean, which are treated as true. You can obtain the result of a test to return it.

```
1 (cond
2   [(+ 4 2) => (lambda (x) (* x x))]
3   [else 0])
```

This gives 36.

2.9.1 and, or as conditionals

and evaluates all of its operands from left to right and stops as soon as `#f` operand is read, otherwise the last expression is the answer.

or evaluates all of its operands from left to right and stops as soon as a non `#f` operand is read, and that becomes the answer, otherwise the answer is `#f`

2.10 Local bindings

Local definitions for use in just one expression

```
1 (let ([x expr1]
2      [y expr2]))
3 (+ x y (* 2 x y)))
```

This means to compute $x + y + 2xy$ where $x = \text{expr1}$ and $y = \text{expr2}$. These 2 expressions cannot see the others variables, only the global ones outside of their scope.

```
1 (let ([x 3])
2      (let ([x (* 3 3)]) ; (* 3 3)
3          x))
```

This results in 9 and is not recursive. Let* allows later bindings to see earlier bindings

```
1 (let* ([x 5]
2       [y (+ x 1)]) ; (+ 5 1)
3       (+ x y (* 2 x y)))
```

2.10.1 Recursive local bindings

letrec allows more recursive bindings

```
1 (letrec ([fac
2          (lambda (n)
3            (if (= n 0) 1 (* n (fac (- n 1)))))])
4          [even
5            (lambda (n)
6              (or (= n 0) (not (odd (- n 1)))))])
7          [odd
8            (lambda (n)
9              (not (even (- n 1)))))])
10 (even (fac 5)))
```

This returns whether or not factorial 5 is even. So true.

2.11 Recommended Code Layout

- Open parentheses then immediately first word
- Procedure definition: Body starts on new line, indented
- Long expression: Parts start on new lines, indented

- Closing parentheses not on new lines

3 Thursday, January 25, 2018

3.1 Pairs and Lists

A cons cell is a 2-tuple pair and has the following syntax:

```
1 cons(x y)
```

Essentially a pair of pointers. Has special support for lists.

```
1 '() ; an empty list
2 (list x y z) = (cons x (cons y (cons z '())))
3 '(42 "hi" Chrome)
4 ; Chrome here will be the symbol 'Chrome
```

For a cons cell, you can use car to access the first field and cdr to access the second field.

3.2 User-Defined Records

```
1 (struct dim (width height))
```

This creates a new record type with 2 fields

```
1 (dim 4 7) ; this constructs a value of this type
2 dim? ; this tests for this type
3 dim-width
4 dim-height
5 ; these are the field accessors
```

We can also use struct-copy to clone a record while replacing some values:

```
1 (define d1 (dim 4 7))
2 (define d2 (struct-copy dim d1 [width 5]))
3 ; d2 is (dim 5 7)
```

3.3 Pattern Matching

You can test for a literal, cons cell, or a record type, can get their content as well.

```
1 (struct dim (width height))
2
3 (define (foo x)
4   (match x
5     ['() 'nada]
6     [(cons b _) b]
7     [(dim w h) (* w h)]))
```

```
8 (foo '()) ; returns 'nada
9 (foo '(1 2 3)) ; returns 1
10 (foo (dim 4 7)) ; returns 28
```

3.4 Input and Output

We can print with `display`, `printf` and `displayln`

```
1 (display 5)
2 (newline)
3 (displayln 5)
4 (printf "yes" "price" 5)
```

```
1 (read-line) ; this reads a line
2 (read-string 10) ; this reads up to the upper bound.
3 ; if it reaches the end of the file, it returns eof, which you
  ; can use eq? or eof-object? to test
4 ; for stderr, eprintf is like printf but goes to stderr
```

3.4.1 ports

Racket has ports, analogous to Java Reader/Writer – behind it can be file, string, network connection, message queue, user-defined, etc.

3.5 Sequencing

If we want to evaluate multiple expressions in the order we specify

```
1 (begin
2   (display ln "Please enter your name")
3   (read-line)) ; this returns the last expression
4
5 (begin0 expr1 expr2) ; this returns the first expression, but
  the others are still evaluated.
6
7 (when (> x 0) expr1 expr2 ...)
8 ; if true, evaluates the expressions, returns what the last
  one returns, if false, returns #<void>
```

3.6 Mutable Variables

```
1 (define v 5)
2 (define (f x) (+ x v))
3 (f 0) ; this gives 5
4 (set! v 6)
5 (f 0) ; this gives 6
```

Mutable pairs, lists, strings, arrays, etc. are also available. Use mutation judiciously, is not that necessary.

3.7 map

Takes in a function and a list and applies the function to every element in that list

```
1 (map f (list x y z)) = (list (f x) (f y) (f z))
```

3.8 filter

filter takes in a boolean function and a list (A) and returns a list of the items in the list A that satisfy the boolean function

```
1 (filter number? '(9 "4" 0 "1" "6" 5)) = '(9 0 5)
```


4 Thursday, February 1, 2018

4.1 Scheme (cont'd)

4.1.1 foldl

Consider the problem of summing an entire list and multiplying an entire list. Summing requires us to add up all the elements of the list plus 0 for the first item. Multiplying requires us to multiply up all the elements of the list times 1 for the first item. This is the motivation.

So we define foldl as

```

1 (define (foldl binop a lst)
2   (match lst
3     [(?) a]
4     [(cons hd tl) (foldl binop (binop a hd) tl )]))

```

So Intuitively,

```

1 (foldl binop a (list x y z))

```

Looks like

$$(((a + x) + y) + z)$$

where + is where binop is being performed

4.1.2 foldr

Basically in the opposite direction of foldl, so if

```

1 (foldl binop a (list x y z))

```

Looks like

$$(z + (y + (x + a)))$$

where + is where binop is being performed, then

```

1 (foldr binop a (list x y z))

```

Looks like

$$(((a + x) + y) + z)$$

where + is where binop is being performed, then

4.1.3 Procedure-Call Stack

Consider the following:

```
1 (define (f n) (... (f (- n 1)) ...))
2 (displayln (+ (f 4) (f 1) (f 6)))
```

The Control-flow jumps to into f when it's called and later knows where to return to after the recursive calls. This is done because a stack is used to remember where to return to in recursion, called a **Call Stack**. The Benefit of this is that it supports recursion, but it comes at a price of occupying $\theta(1)$ space while the stack is being used.

4.1.4 Non-Tail Calls and Tail Calls

Non-Tail Calls, are if you still have to do your own processing or computation after getting the results from another function, so basically the results are not returned right away.

For example, for the following function:

```
1 (define (my -sum lst)
2   (match lst [?( ) 0]
3     [(cons hd tl) (+ hd (my -sum tl ))]))
```

Takes $\theta(n)$ space if the list length is n .

A Tail call is the exact opposite, there is no computation after getting the results back from a called function and the function returns the value right away. The complexity of this is $O(1)$ under Tail-Calling optimization in Scheme.

Tail-Calling optimization isn't in every language, Java and Python don't have this.

4.2 Haskell

4.2.1 Expressions and Types

Characters chars are denoted with single quotes

Tuples Not the same as cons

`()` is a special type, called the unit type, used as a return value for functions that don't return anything

Lists Lists are implemented as such:

```
1 3 : ( 1 : ( 4 : [] ) ) = [3,1,4]
```

So a list of length one is an item with an empty list and the colon inbetween separates the items

```
1 [1] = 1 : []
```

Note that the following list has a type `[[integer]]`

```
1 [[3,1,4], [10,20], []]
```

So as for now, arbitrary list nesting is not supported, so basically

```
1 [1, [3]]
```

is not supported (yet).

Note that because of static typing, every item must be the same type, we can't mix integers with floats in the same list.

Strings are a list of chars. The downside of this is that it uses a huge amount of memory, as it's stored as a linked list, and each node and pointer takes up a linear amount of space.

Keyword: Just If you type

```
1 Just 'C'
```

The variable is either `Nothing` or `Char`

Nothing Is the empty type, can be any type

"Left 'C'" Can either be of type `Char Bool`, `Char Int`, ... all we know that the left variable is a character.

"Right False" Can either be of type `Char bool`, `Int bool`, ... all we know that the right variable is a boolean

anonymous functions Ex.

```
1 \x -> x >= 'C'
```

Has the type `Char -> Bool` with `char` as the domain and `bool` as the codomain

4.2.2 Definitions

We can define expressions to variables and vice versa, for example, in the following, we are defining "ten" to be "1+2+3+4" and binding "1+2+3+4" to "ten":

```
1 ten = 1 + 2 + 3 + 4
```

There is also pattern binding, with tuples, but that will be shown later

Functions can also be defined:

```
1 square x = x * x
2 nand a b = not (a && b)
```

We can also define type signatures for the definitions as such

```
1 ten, four :: Integer
```

But Haskell is written such that the type signature can be separated from the definition, so you don't have to put them in the same few lines, they just have to be in the same file.

4.2.3 Function Applications

If you insert one parameter to a function that takes 2 parameters, that function will return a function of 1 parameter. This is how Haskell does multiple parameters

5 Thursday, February 8, 2018

5.1 Haskell (cont'd)

5.1.1 Local Definitions For Expressions

```

1  let x = 4 + 5
2      y = 4 - 5
3  in x+y+2*x*y

```

Layout Version above. Braced Version below

5.1.2 Local Definitions For Definitions

```

1  foo u v = x + y + 2*x*y
2      where — this where refers to the statement in line 1
3          x = u + v
4          y = u - v

```

5.1.3 Pattern Matching

Can be done for expressions or function definitions as well as pattern binding.

```

1  — expression case
2  case expr of
3      [] -> 0
4      42 : xs -> foo xs
5      x : xs -> x + foo xs

```

In this example, the expression *expr* would evaluate to 0 if it was an empty list, have foo applied to its tail if it started with 42 as its first index and x plus foo applied to its tail for any other list.

```

1  — function definition case examples
2
3  mySum [] = 0
4  mySum (x : xs) = x + mySum xs
5
6  nand False _ = True
7  nand True False = True
8  nand True True = False

```

```

1  — pattern binding case
2  [a, b, c] = take 3 someList
3  — a = take
4  — b = 3
5  — c = someList

```

5.1.4 Guards

Guards are extra conditions imposed on patterns.

```

1  — expression case
2  case expr of
3    [] -> 0
4    x : xs | x < 0 -> x + foo xs
5             | x > 2 -> x - foo xs
6             | True -> x * foo xs
7
8  — definition case
9  foo [] = 0
10 foo (x : xs) | x < 0 -> x + foo xs
11               | x > 2 -> x - foo xs
12               | True -> x * foo xs

```

Instead of True for the edge case, can also use "otherwise".

5.1.5 Local Definitions under Patterns and Guards

```

1  foo :: Either String Integer -> Integer
2  foo (Left str) | suffix > "albert" = 42
3                  | otherwise = 24
4                  where
5                      suffix = drop 10 str
6
7  foo (Right x) | x > 0 = 2*y
8                  | x < 0 = y
9                  | otherwise = 0
10                  where
11                      y = div 1000 x

```

Note that the first where belongs to the foo where the input is a Left str, the second where belongs to the foo where the input is a Right integer and if $x = 0$, then y will not be calculated.

5.1.6 List Comprehension

```

1  [x + y | x <- [10,20,30], x > 10, y <- [4,5] ]
2  — this results in
3  --[20+4, 20+5, 30+4, 30+5]

```

We can use pattern matching as well

```

1  [x+3 | Just x <= [Just 10, Nothing, Just 30]]
2  — this results in
3  — [10+3, 30+3]

```

There is also a range notation we can use:

```
1 [1...5] — this is the same as [1,2,3,4,5]
```

5.1.7 Algebraic Data Types

```
1 data MyType = Nada | Duplet Double String | Uno Integer
```

Nada, Duplet, Uno are data constructors. They must start with uppercase letters. They form expressions and patterns.

The following is an example function that takes in "MyType":

```
1 plus1 :: MyType -> MyType
2 plus1 Nada = Nada
3 plus1 (Duplet r s) = Duplet (r+1) s
4 plus1 (Uno i) = Uno (i+1)
```

List, unit, tuple, Maybe, and either are algebraic data types from the standard library.

Recursive definitions are ok too, like the following:

```
1 data Stack = Button | Push Int Stack
```

5.1.8 Parametric Polymorphism

consider the following function type contract

```
1 map :: (a -> b) -> [a] -> [b]
```

Here both a and b are type variables. They start with lowercase letters (actual data types are capitalized, like Bool). The user chooses what types to use for a and b, and the implementer cannot choose what type of a and b, and must let their function work for all types of a and b.

Algebraic data types can be parameterized by type variables too, for example:

```
1 data Either a b = Left a | Right b
2 — We can generalize the previous stack example to
3 data Stack a = Bottom | Push a (Stack a)
```

5.1.9 Type-Class Polymorphism

We notice that comparison functions like

```

1  (==)
2  (<)

```

cannot use completely general it its polymorphism, they must take in items as input that are two of the same type of class that can be compared.

So how does Haskell pull these off?

Haskell uses a **type class** that declares overloaded operations. So the example from before:

```

1  (==)
2  (<)

```

must take in two inputs, lets call them a and b , where they are both of type $Eq\ a$, which is a class that lets the two inputs be compared.

However, note that classes are not the same as types. Eq is not a type, $Bool$ is not a subclass.

So if we were to type this out:

```

1  (==) :: Eq a => a -> a -> Bool
2  - "Eq a" is a "class constraint"

```

In this example, the user chooses what type to use for a , but that chosen type must be an instance of Eq .

Note that constraints propagate down the dependency chain:

```

1  (==) :: Eq a => a -> a -> Bool
2
3  eq3 :: Eq a => a -> a -> a -> Bool
4  eq3 x y z = x==y && y==z

```


6 Thursday, February 15, 2018

6.1 Haskell (cont'd)

6.1.1 Constraint Instances

What if you are comparing instances inside of a data structure (let's say, a list for example)?

Well we can do this:

```

1 instance Eq a => Eq [a] where
2   [] == []      = True
3   (x:xs) == (y:ys) = x==y && xs == ys
4   _ == _        = False

```

Constraints propagate down the dependency chain, including other instance implementations

6.1.2 User-Defined Class

```

1 class ADT a where
2   tag :: a -> String
3
4 instance ADT (Either a b) where
5   tag (Left _) = "Left"
6   tag (Right _) = "Right"
7
8 instance ADT MyType where
9   tag Nada = "Nada"
10  tag (Duplet _ _) = "Duplet"
11  tag (Uno _) = "Uno"

```

Classes in Haskell are similar to Java's Interfaces in the sense that you implement the classes' functions for each instance as well.

```

1 class Eq a => Ord a where
2   (<), (<=), (>), (>=) :: a -> a -> Bool
3   compare :: a -> a -> Ordering
4 data Ordering = LT | EQ | GT
5
6 — The "Eq a =>" here means that
7 — Every Ord instance is also an EQ instance
8 — (Superclass, subclass)

```

For implementers of type and instances, these superclasses must be specified, but for users, they can use `Ord` without mentioning `Eq`.

6.1.3 Auto-Generating Instance Implementations

The compiler is willing to write some instance code for you, for select standard classes: (`Eq`, `Ord`, `Enum` (but no fields allowed), `Show`, and a few others.

```

1 data MyType = Nada | Duplet Double String | Uno Integer
2   deriving (Eq, Ord, Show)
3 data Browser = FireFox | Chrome | Edge | Safari
4   deriving (Eq, Ord, Show)

```

6.1.4 Haskell's Number System

You can't use doubles and integers together for number operators, you have to convert one so that they are the same (Can use `fromIntegral` to convert an integer into a double).

6.1.5 Functor

```

1 fmap_List :: (a -> b) -> [a] -> [b]
2 fmap_List = map
3
4 fmap_Maybe :: (a -> b) -> Maybe a -> Maybe b
5 fmap_Maybe f Nothing = Nothing
6 fmap_Maybe f (Just a) = Just (f a)
7
8 fmap_Either :: (a->b) -> Either e a -> Either e b
9 fmap_Either f (Left e) = Left e
10 fmap_Either f (Right a) = Right (f a)

```

The pattern here is that there is a $f : a \mapsto b$ induces a corresponding $Fa \mapsto Fb$ where F is a parameterized type. There is a class for that.

```

1 class Functor f where
2   fmap :: (a->b) -> f a -> f b

```

So this function generalizes the previous examples above.

Every instance of `Functor` should satisfy:

```

1 fmap id xs = xs
2 fmap g (fmap f xs) = fmap (g . f) xs

```

`fmap` also has an infix alias of `< $ >`, for example:

```
1 sin <$> [1,2,3]
```

6.1.6 Applicatives

You now become ambitious. You ask: What if you have a binary operator, and two lists, ...

```
1 listCross :: (a -> b -> c) -> [a] -> [b] -> [c]
2
3 maybeBoth :: (a -> b -> c) -> Maybe a -> Maybe b -> Maybe c
4 maybeBoth op (Just a) (Just b) = Just (op a b)
5 maybeBoth op _ _ = Nothing
```

And what if you have a ternary operator and three lists?

Can you implement

```
1 ap_List :: [a -> b] -> [a] -> [b]
```

such that, for example:

```
1 ap_List [f,g] [1,2,3] = [f 1, f 2, f 3, g 1, g 2, g 3]
```

Answer:

```
1 ap_List = ListCross (\f -> \x -> f x)
```

Equivalently listCross (\$)

Now can implement ternary too:

```
1 listTernary :: (a->b->c->d)->[a]->[b]->[c]->[d]
2 listTernary ternary as bs cs =
3   ((ternary <$> as) 'ap_List' bs) 'ap_List' cs
```

There is a class for this too.

```
1 class Functor f => Applicative f where
2   pure :: a -> f a
3   (<*>) :: f (a -> b) -> f a -> f b
4
5 — example instance
6 instance Applicative Maybe where
7   pure a = Just a
8   Just f <*> Just a = Just (f a)
9   _ <*> _ = Nothing
```

Applicative subsumes Functor, so we can implement fmap as

```
1 fmap f xs = pure f <*> xs
```

7 Thursday, March 1, 2018

7.1 Haskell (cont'd)

7.1.1 Monads

So far we have been thinking of List, Maybe, Either,... as data structures (ex. containers). Now we must think of them as programs:

- `foo :: Maybe Int` means: a program that may return a number successfully, or may abort
- `foo :: [Int]` means: a non-deterministic program that returns different numbers in different parallel universes.
- `foo :: Either String Int` means: like Maybe, but if it aborts, it uses Left to tell you an error message

Now we re-read Functor and Applicative from this angle

```

1 fmap abs foo
2 — this now means to return the absolute value
3 — of what foo returns
4
5 (+) <$> foo <*> bar
6 — this now means to return the sum of what
7 — foo and bar returns

```

But bar does not know what foo returns, or vice versa.

You now become ambitious. Can you combine two programs such that the return value(s) of the 1st is fed to the 2nd so the 2nd can behave independently? Like such:

```

1 bind :: F a -> (a -> Fb) -> F b
2 — so we can have
3 — prog1st 'bind' prog2nd?

```

We can think of prog2nd as a callback to prog1st.

Other examples would look like the following:

```

1 — bind for Maybe
2
3 bind_Maybe :: Maybe a -> (a -> Maybe b) -> Maybe b

```

```

4 bind_Maybe Nothing _ = Nothing
5 bind_Maybe (Just a) k = k a
6
7 -- bind for List
8
9 bind_List :: [a] -> (a -> [b]) -> [b]
10 bind_List [] _ = []
11 bind_List (a:as) k = k a ++ bind_List as k

```

There is a class for that too.

```

1 class Applicative f => Monad f where
2   return :: a -> f a
3   (>>=) :: f a -> (a -> f b) -> f b
4
5 -- example instance
6
7 instance Monad [] where
8   return a = [a]
9   as >>= k = concat (map k as)

```

Remark: return and pure should be the same thing. Historically, Monad came first, Applicative came later, thus the redundancy. There is a proposed change to make return an alias of pure.

Monad subsumes Applicative: Can implement (`<*>`) as

```

1 fs <*> as = fs >>=
2   \f -> as >>=
3   \a -> return (f a)

```

There are equations for Monad too, such as:

```

1 return a >>= k = k a

```

There is "do-notation" so code looks nicer and the computer emits

```

1 >>= \v ->

```

for you:

```

1 fs <*> as = do
2   f <- fs
3   a <- as
4   return (f a)

```

7.1.2 Haskell I/O System

Parametrized type "IO" for all "I/O" commands. Instance of Monad, Applicative, Functor.

```

1 foo :: IO Char
2 -- this means a program that interacts with the outside world,
   then returns a character (or gets stuck forever, or throws an
   exception).
3
4 putStrLn :: String -> IO ()
5 getLine  :: IO String
6 -- NOT: getLine :: String

```

Do not think about how to extract the string. Use (`>>=`) to feed it to the next program (callback).

```

1 main = getLine >>= \s -> putStrLn ("It 's " ++ s)
2 -- OR
3 main = do
4   s <- getLine
5   putStrLn ("It 's " ++ s)

```

7.2 Syntax

7.2.1 Context-Free Grammar (CFG)

A context-free grammar looks like this bunch of rules:

$$\begin{aligned}
 E &\rightarrow E + E \\
 M &\rightarrow M \times M \\
 A &\rightarrow 0 \\
 A &\rightarrow (E) \\
 E &\rightarrow M \\
 M &\rightarrow A \\
 A &\rightarrow 1
 \end{aligned}$$

Main idea:

- E, M, A are non-terminal symbols aka variables. When you see them, you apply rules to expand
- $+, \times, 1, 0, (,)$ are terminal symbols. They are the characters you want in your language

7.2.2 Derivation (aka Generation)

Derivation is a finite sequence of applying the rules until all non-terminal symbols are gone. Often aim for a specific final string.

$$\begin{aligned} E &\rightarrow M \\ &\rightarrow M \times M \\ &\rightarrow A \times M \\ &\rightarrow 1 \times M \\ &\rightarrow 1 \times A \\ &\rightarrow 1 \times (E) \\ &\rightarrow 1 \times (E + E) \\ &\rightarrow 1 \times (M + E) \\ &\rightarrow 1 \times (A + E) \\ &\rightarrow 1 \times (0 + E) \\ &\rightarrow 1 \times (0 + M) \\ &\rightarrow 1 \times (0 + M \times M) \\ &\rightarrow 1 \times (0 + A \times M) \\ &\rightarrow 1 \times (0 + 1 \times M) \\ &\rightarrow 1 \times (0 + 1 \times A) \\ &\rightarrow 1 \times (0 + 1 \times 1) \end{aligned}$$

Context-free grammars can support: matching parentheses, unlimited nesting.

7.2.3 Backus-Naur Form (BNF)

Backus-Naur Form is a computerized, practical notation for CFG

- Our example grammar in BNF:

7.2.4 Parse Tree (aka Derivation Tree)

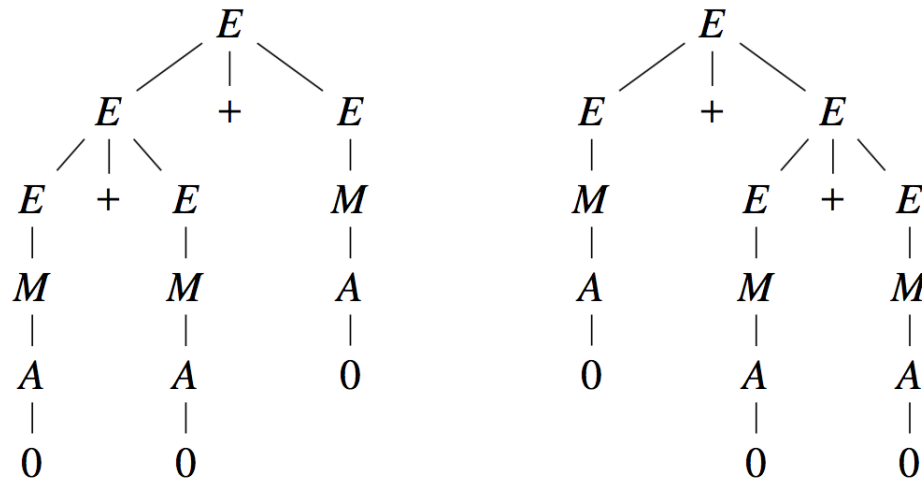
```

graph TD
    E1[E] --- E2[E]
    E1 --- P1[+]
    E1 --- E3[E]
    E2 --- E4[E]
    E2 --- P2[+]
    E2 --- E5[E]
    E4 --- M1[M]
    E4 --- A1[A]
    E4 --- 01[0]
    E5 --- M2[M]
    E5 --- A2[A]
    E5 --- 02[0]
    E3 --- M3[M]
    E3 --- A3[A]
    E3 --- 03[0]
  
```

7 THURSDAY, MARCH 1, 2018

7.2.5 Ambiguous Grammar

Two different trees generate the same $0 + 0 + 0$



If this happens, the grammar is ambiguous.

We try to design unambiguous grammars.

(Bad news: CFG ambiguity is undecidable)

7.2.6 Unambiguous Grammar Example

An unambiguous grammar that generates the same language as our ambiguous grammar example:

```

1 <expr> ::= <expr> "+" <expr> | <mul>
2 <mul> ::= <mul> "*" <mul> | <atom>
3 <atom> ::= "0" | "1" | "(" <expr> ")"
  
```

(Bad news: Equivalence of two CFGs is also undecidable)

7.2.7 Left Recursive vs Right Recursive

```

1 <expr> ::= <expr> "+" <mul>
  
```

This is an example of a left recursive rule. The recursion is at the beginning (left).

```
1 <expr> ::= <mul> "+" <expr>
```

This is an example of a right recursive rule. The recursion is at the end (right).

They affect whether infix operators associate to the left or right.

They also affect some parsing algorithms.

7.2.8 Recursive Descent Parsing

Recursive descent parsing is a simple strategy for writing a parser.

- Write a procedure for each rule
- Non-terminals on Right Hand Side become procedure calls, possible recursive calls. (Thus "recursive descent". Also "top-down".) (Left-recursive grammars need special treatment.)
- Terminal symbols: Consume input and check
- Alternatives require lookahead and/or backtracking

7.2.9 Recursive Descent Parser Example

Example grammar suitable for recursive descent parsing:

```
1 <sub> ::= <atom> "-" <sub> | <atom>
2 <atom> ::= "0" | "1" | "(" <sub> ")"
```

Pseudo-code of recursive descent parser:

```
1 sub:
2   try (atom
3     read; if not "-" then fail
4     sub)
5   if that failed: atom
6
7 atom:
8   read;
9   if "1" or "0": return
10  if "(" : sub
11    read; if not ")" then fail
12  else: fail
```

8 Thursday, March 8, 2018

8.1 Parser Library

```

1  -- | Library of parser definition and operations.
2  module ParserLib where
3
4  import Control.Applicative
5  import Data.Char
6  import Data.Functor
7  import Data.List
8
9  newtype Parser a = PsrOf{
10     -- | Function from input string to:
11     --
12     -- * Nothing, if failure (syntax error);
13     -- * Just (unconsumed input, answer), if success.
14     dePsr :: String -> Maybe (String, a)}
15
16 -- Monadic Parsing in Haskell uses [] instead of Maybe to
17 -- support ambiguous
18 -- grammars and multiple answers.
19
20 -- | Use a parser on an input string.
21 runParser :: Parser a -> String -> Maybe a
22 runParser (PsrOf p) inp = case p inp of
23     Nothing -> Nothing
24     Just (_, a) -> Just a
25     -- OR: fmap (\(-,a) -> a) (p inp)
26
27 -- | Read a character and return. Failure if input is empty.
28 anyChar :: Parser Char
29 anyChar = PsrOf p
30     where
31         p "" = Nothing
32         p (c:cs) = Just (cs, c)
33
34 -- | Read a character and check against the given character.
35 char :: Char -> Parser Char
36 char wanted = PsrOf p
37     where
38         p (c:cs) | c == wanted = Just (cs, c)
39         p _ = Nothing
40 char wanted = satisfy (\c -> c == wanted)    -- (== wanted)

```

```

41 — | Read a character and check against the given predicate.
42 satisfy :: (Char -> Bool) -> Parser Char
43 satisfy pred = PsrOf p
44   where
45     p (c:cs) | pred c = Just (cs, c)
46     p _ = Nothing
47
48 — | Expect the input to be empty.
49 eof :: Parser ()
50 eof = PsrOf p
51   where
52     p "" = Just ("", ())
53     p _ = Nothing
54
55
56 — | Read and check against a given string.
57 string :: String -> Parser String
58 string wanted = PsrOf p
59   where
60     p inp = case stripPrefix wanted inp of
61               Nothing -> Nothing
62               Just suffix -> Just (suffix, wanted)
63     — Refactor this!
64
65 — But you have to compose smaller parsers to build larger
66 — parsers and to return
67 — more interesting answers, e.g., abstract syntax trees.
68 — This is what fmap, pure, <*>, >>= are for. And there are
69 — more...
70 instance Functor Parser where
71   — fmap :: (a -> b) -> Parser a -> Parser b
72   fmap f (PsrOf p) = PsrOf q
73     — (\inp -> fmap (\(rest, a) -> (rest,
74       f a)) (p inp))
75   where
76     q inp = case p inp of
77               Nothing -> Nothing
78               Just (rest, a) -> Just (rest, f a)
79     — fmap (\(rest, a) -> (rest, f a)) (p inp)
80 instance Applicative Parser where
81   — pure :: a -> Parser a
82   pure a = PsrOf (\inp -> Just (inp, a))

```

```

83  -- (<*>) :: Parser (a -> b) -> Parser a -> Parser b
84  -- Consider the 1st parser to be stage 1, 2nd parser stage
    2.
85  PsrOf p1 <*> PsrOf p2 = PsrOf q
86  where
87      q inp = case p1 inp of
88          Nothing -> Nothing
89          Just (middle, f) ->
90              case p2 middle of
91                  Nothing -> Nothing
92                  Just (rest, a) -> Just (rest, f a)
93      -- dePsr (fmap f (PsrOf p2)) middle
94
95  instance Alternative Parser where
96      -- empty :: Parser a
97      -- Always fail. The identity for <|> below.
98      empty = PsrOf (\_ -> Nothing)
99      -- (<|>) :: Parser a -> Parser a -> Parser a
100     -- Try the 1st one. If success, done; if failure, do the 2nd
        one
101     PsrOf p1 <|> PsrOf p2 = PsrOf q
102     where
103         q inp = case p1 inp of
104             j@(Just _) -> j
105             Nothing -> p2 inp
106     -- many :: Parser a -> Parser [a]
107     -- 0 or more times, maximum munch, collect the answers into
        a list.
108
109     -- some :: Parser a -> Parser [a]
110     -- 1 or more times, maximum munch, collect the answers into
        a list.
111
112
113  instance Monad Parser where
114
115
116  -- | Zero or more operands separated by an operator.
117  chainl :: Parser a -- ^ operand parser
118         -> Parser (a -> a -> a) -- ^ operator parser
119         -> a -- ^ start value
120         -> Parser a -- ^ evaluated answer
121  chainl = undefined
122
123  -- | One or more operands separated by an operator.

```

```
124 chainl1 :: Parser a          -- ^ operand parser
125     -> Parser (a -> a -> a)  -- ^ operator parser
126     -> Parser a              -- ^ evaluated answer
127 chainl1 = undefined
128
129 -- | Consume zero or more spaces, maximum munch.
130 spaces :: Parser String
131 spaces = undefined
```