Introduction to Functional Programming in Haskell
Outline

Why learn functional programming?

The essence of functional programming
  What is a function?
  Equational reasoning
  First-order vs. higher-order functions
  Lazy evaluation

How to functional program
  Functional programming workflow
  Data types
  Type classes
  Type-directed programming
  Haskell style
  Refactoring (bonus section)

Type inference
Outline

Why learn functional programming?

The essence of functional programming

How to functional program

Type inference
Why learn (pure) functional programming?

1. This course: strong correspondence of core concepts to PL theory
   - abstract syntax can be represented by algebraic data types
   - denotational semantics can be represented by functions

2. It will make you a better (imperative) programmer
   - forces you to think recursively and compositionally
   - forces you to minimize use of state
   ...essential skills for solving big problems

3. It is the future!
   - more scalable and parallelizable
   - functional features have been added to most mainstream languages
   - many cool new libraries built around functional paradigm
Programming is not about data. It’s about transforming data … That’s why I think functional programming is a natural successor to object-oriented programming. In OOP, we’re constantly concerned about the state of our data. In functional programming, our focus is on the transformation of data. And transformation is where the value is added.

—Dave Thomas, The Pragmatic Programmer
Quotes from pragmatists

Without understanding functional programming, you can’t invent MapReduce, the algorithm that makes Google so massively scalable.

—Joel Spolsky, Joel on Software
So that's the big deal about functional languages; and it is one big fricking deal. There is a freight train barreling down the tracks towards us, with multi-core emblazoned on it; and you'd better be ready by the time it gets here.

—Robert C. Martin, a.k.a. “Uncle Bob”
Outline

Why learn functional programming?

The essence of functional programming
   What is a function?
      Equational reasoning
      First-order vs. higher-order functions
      Lazy evaluation

How to functional program

Type inference
What is a (pure) function?

A function is **pure** if:

- it always returns the same output for the same inputs
- it doesn’t do anything else — no “side effects”

In Haskell: whenever we say “function” we mean a **pure function**!
What are and aren’t functions?

Always functions:
- mathematical functions \( f(x) = x^2 + 2x + 3 \)
- encryption and compression algorithms

Usually not functions:
- C, Python, JavaScript, … “functions” (procedures)
- Java, C#, Ruby, … methods

Haskell only allows you to write (pure) functions!
Why procedures/methods aren’t functions

- output depends on environment
- may perform arbitrary side effects
Outline

Why learn functional programming?

The essence of functional programming
  What is a function?
    Equational reasoning
    First-order vs. higher-order functions
    Lazy evaluation

How to functional program

Type inference
Getting into the Haskell mindset

In Haskell, “=” means *is not change to!*
Getting into the Haskell mindset

Quicksort in Haskell

```haskell
qsort :: Ord a => [a] -> [a]
qusort [] = []
qusort (x:xs) = qsort (filter (\leq x) xs) ++ [x] ++ qsort (filter (> x) xs)
```

QuickSort in C

```c
void qsort(int low, int high) {
    int i = low, j = high;
    int pivot = numbers[low + (high-low)/2];
    while (i <= j) {
        while (numbers[i] < pivot) {
            i++;
        }
        while (numbers[j] > pivot) {
            j--;
        }
        if (i <= j) {
            swap(i, j);
            i++;
            j--;
        }
    }
    if (low < j) qsort(low, j);
    if (i < high) qsort(i, high);
}
void swap(int i, int j) {
    int temp = numbers[i];
    numbers[i] = numbers[j];
    numbers[j] = temp;
}
```
Referential transparency

An expression can be replaced by its value without changing the overall program behavior.

\[ \text{length} \ [1,2,3] + 4 \]
\[ \Rightarrow \]
\[ 3 + 4 \]

Corollary: an expression can be replaced by any expression with the same value without changing program behavior.

Supports equational reasoning.
Equational reasoning

Computation is just substitution!

sum :: [Int] -> Int
sum [] = 0
sum (x:xs) = x + sum xs

\[
\begin{align*}
\text{sum [2,3,4]} &\Rightarrow \text{sum (2:(3:(4:[])))} \\
&\Rightarrow 2 + \text{sum (3:(4:[]))} \\
&\Rightarrow 2 + 3 + \text{sum (4:[])} \\
&\Rightarrow 2 + 3 + 4 + \text{sum []} \\
&\Rightarrow 2 + 3 + 4 + 0 \\
&\Rightarrow 9
\end{align*}
\]
Describing computations

**Function definition**: a list of equations that relate inputs to output
- matched top-to-bottom
- applied left-to-right

**Example: reversing a list**

**imperative view**: how do I rearrange the elements in the list? ✗

**functional view**: how is a list related to its reversal? ✓

```haskell
reverse :: [a] -> [a]
reverse [] = []
reverse (x:xs) = reverse xs ++ [x]
```
Exercise

1. Evaluate: \( \text{double (succ (double 3))} \)
   
   A: 12  B: 14  C: 16

2. Prove, up to associativity of (+), using equational reasoning:
   
   \( \text{double (succ x)} = \text{succ (succ (double x))} \)
Exercise

1. Evaluate: \( \text{foo } (\text{foo } 0 0) 0 \)
   
   - A : 0
   - B : 2
   - C : 3

2. Evaluate: \( \text{foo } 0 (\text{foo } 0 0) \)
   
   - A : 0
   - B : 2
   - C : 3

\[
\text{foo} \, : \, \text{Int -> Int -> Int} \\
\text{foo } 0 \, x = x + 1 \\
\text{foo } x \, 0 = x + 2 \\
\text{foo } x \, y = x + y
\]
Outline

Why learn functional programming?

The essence of functional programming
   What is a function?
   Equational reasoning
      First-order vs. higher-order functions
   Lazy evaluation

How to functional program

Type inference
First-order functions

Examples

- \( \text{cos} :: \text{Float} \rightarrow \text{Float} \)
- \( \text{even} :: \text{Int} \rightarrow \text{Bool} \)
- \( \text{length} :: [\text{a}] \rightarrow \text{Int} \)
Higher-order functions

Examples

- `map :: (a -> b) -> [a] -> [b]`
- `filter :: (a -> Bool) -> [a] -> [a]`
- `(.) :: (b -> c) -> (a -> b) -> a -> c`
Higher-order functions as control structures

**map**: loop for doing something to each element in a list

```haskell
map :: (a -> b) -> [a] -> [b]
map f [] = []
map f (x:xs) = f x : map f xs

map f [2,3,4,5] = [f 2, f 3, f 4, f 5]
map even [2,3,4,5] = [even 2, even 3, even 4, even 5]
= [True, False, True, False]
```

**fold**: loop for aggregating elements in a list

```haskell
foldr :: (a->b->b) -> b -> [a] -> b
foldr f y [] = y
foldr f y (x:xs) = f x (foldr f y xs)

foldr f y [2,3,4] = f 2 (f 3 (f 4 y))
foldr (+) 0 [2,3,4]
= (+) 2 ((+) 3 ((+) 4 0))
= 2 + (3 + (4 + 0))
= 9
```
Can create new functions by **composing** existing functions

- *apply the second function, then apply the first*

### Function composition

\[(\cdot) :: (b \rightarrow c) \rightarrow (a \rightarrow b) \rightarrow a \rightarrow c\]

\[f \cdot g = \lambda x \rightarrow f(g\ x)\]

### Types of existing functions

- `not :: Bool \rightarrow Bool`
- `succ :: Int \rightarrow Int`
- `even :: Int \rightarrow Bool`
- `head :: [a] \rightarrow a`
- `tail :: [a] \rightarrow [a]`

### Definitions of new functions

- `plus2 = succ \cdot succ`
- `odd = not \cdot even`
- `second = head \cdot tail`
- `drop2 = tail \cdot tail`
Currying / partial application

In Haskell, functions that take multiple arguments are implicitly higher order

```
plus :: Int -> Int -> Int
increment :: Int -> Int
increment = plus 1
```

Curried
```
plus 2 3
```

Uncurried
```
plus (2,3)
```

a pair of ints
Outline

Why learn functional programming?

The essence of functional programming
  What is a function?
  Equational reasoning
  First-order vs. higher-order functions
  Lazy evaluation

How to functional program

Type inference
Lazy evaluation

In Haskell, expressions are reduced:

- only when needed
- at most once

Supports:

- infinite data structures
- separation of concerns

```haskell
nats :: [Int]
nats = 1 : map (+1) nats

fact :: Int -> Int
fact n = product (take n nats)

min3 :: [Int] -> [Int]
min3 = take 3 . sort
```

What is the running time of this function?
Outline

- Why learn functional programming?
- The essence of functional programming

How to functional program
  - Functional programming workflow
  - Data types
  - Type classes
  - Type-directed programming
  - Haskell style
  - Refactoring (bonus section)

Type inference
“obsessive compulsive refactoring disorder”
FP workflow (detailed)

1A. Data Description → 1B. Data Examples

2. Function Description (Signature/Purpose/Header)

3. Functional Examples

4. Function Template

5. Code

6. Tests

7. Review & Refactor

Norman Ramsey, On Teaching “How to Design Programs”, ICFP’14
Outline

Why learn functional programming?

The essence of functional programming

How to functional program
  Functional programming workflow
    Data types
      Type classes
      Type-directed programming
    Haskell style
  Refactoring (bonus section)

Type inference
Algebraic data types

Data type definition
- introduces new type of value
- enumerates ways to construct values of this type

Definitions consists of …
- a type name
- a list of data constructors with argument types

Definition is inductive
- the arguments may recursively include the type being defined
- the constructors are the only way to build values of this type

Some example data types

```haskell
data Bool = True | False

data Nat  = Zero  | Succ Nat

data Tree = Node Int Tree Tree | Leaf Int
```
Anatomy of a data type definition

```
data Expr = Lit Int  
           | Plus Expr Expr
```

Example: $2 + 3 + 4 \quad \text{Plus (Lit 2) (Plus (Lit 3) (Lit 4))}$
FP data types vs. OO classes

<table>
<thead>
<tr>
<th>Haskell</th>
<th>Java</th>
</tr>
</thead>
<tbody>
<tr>
<td>`data Tree = Node Int Tree Tree</td>
<td><code>abstract class Tree { ... }</code></td>
</tr>
<tr>
<td></td>
<td>`class Node extends Tree {</td>
</tr>
<tr>
<td></td>
<td>int label;</td>
</tr>
<tr>
<td></td>
<td>Tree left, right;</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td></td>
<td>}</td>
</tr>
<tr>
<td></td>
<td><code>class Leaf extends Tree { ... }</code></td>
</tr>
</tbody>
</table>

- separation of type- and value-level
- set of cases closed
- set of operations open
- merger of type- and value-level
- set of cases open
- set of operations closed

Extensibility of cases vs. operations = the “expression problem”
Type parameters

(Like generics in Java)

data List a = Nil
  | Cons a (List a)

Specialized lists

type IntList = List Int

type CharList = List Char

type RaggedMatrix a = List (List a)
Outline

Why learn functional programming?

The essence of functional programming

How to functional program
   Functional programming workflow
   Data types
      Type classes
   Type-directed programming
   Haskell style
   Refactoring (bonus section)

Type inference
What is a type class?

1. an **interface** that is supported by many different types
2. a **set of types** that have a common behavior

```haskell
class Eq a where
  (==) :: a -> a -> Bool

class Show a where
  show :: a -> String

class Num a where
  (+) :: a -> a -> a
  (*) :: a -> a -> a
  negate :: a -> a
  ...
```

- **types whose values can be compared for equality**
- **types whose values can be shown as strings**
- **types whose values can be manipulated like numbers**
Type constraints

List elements can be of any type

\[
\begin{align*}
\text{length} & : \ [a] \rightarrow \text{Int} \\
\text{length} \ [\ ] & = 0 \\
\text{length} \ (_{\ :} x_{\ :} x_{\ :} x_{\ :} x_{\ :}) & = 1 + \text{length} \ x_{\ :} x_{\ :} x_{\ :} x_{\ :}
\end{align*}
\]

List elements must support equality!

\[
\begin{align*}
\text{elem} & : \text{Eq} \ a \Rightarrow a \rightarrow [a] \rightarrow \text{Bool} \\
\text{elem} \ _{\ []} & = \text{False} \\
\text{elem} \ y \ (x_{\ :} x_{\ :} x_{\ :} x_{\ :}) & = x \ == \ y \ || \ \text{elem} \ y \ x_{\ :} x_{\ :} x_{\ :} x_{\ :}
\end{align*}
\]

use method $\Rightarrow$ add type class constraint

class Eq a where
\( (==) :: a \rightarrow a \rightarrow \text{Bool} \)
Outline

Why learn functional programming?

The essence of functional programming

How to functional program
  Functional programming workflow
  Data types
  Type classes
  Type-directed programming
  Haskell style
  Refactoring (bonus section)

Type inference
Tools for defining functions

Recursion and other functions

```haskell
sum :: [Int] -> Int
sum xs = if null xs then 0
         else head xs + sum (tail xs)
```

(1) case analysis

Pattern matching

```haskell
sum :: [Int] -> Int
sum [] = 0
sum (x:xs) = x + sum xs
```

(2) decomposition

Higher-order functions

```haskell
sum :: [Int] -> Int
sum = foldr (+) 0
```

no recursion or variables needed!
What is type-directed programming?

Use the **type** of a function to help write its **body**.
**Type-directed programming**

Basic goal: transform values of **argument types** into **result type**

<table>
<thead>
<tr>
<th>If argument type is ...</th>
<th>If result type is ...</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>atomic type</strong> (e.g. Int, Char)</td>
<td></td>
</tr>
<tr>
<td>- apply functions to it</td>
<td></td>
</tr>
<tr>
<td><strong>algebraic data type</strong></td>
<td></td>
</tr>
<tr>
<td>- use pattern matching</td>
<td></td>
</tr>
<tr>
<td>- case analysis</td>
<td></td>
</tr>
<tr>
<td>- decompose into parts</td>
<td></td>
</tr>
<tr>
<td><strong>function type</strong></td>
<td></td>
</tr>
<tr>
<td>- apply it to something</td>
<td></td>
</tr>
<tr>
<td><strong>atomic type</strong></td>
<td></td>
</tr>
<tr>
<td>- output of another function</td>
<td></td>
</tr>
<tr>
<td><strong>algebraic data type</strong></td>
<td></td>
</tr>
<tr>
<td>- build with data constructor</td>
<td></td>
</tr>
<tr>
<td><strong>function type</strong></td>
<td></td>
</tr>
<tr>
<td>- function composition or partial application</td>
<td></td>
</tr>
<tr>
<td>- build with lambda abstraction</td>
<td></td>
</tr>
</tbody>
</table>
Outline

Why learn functional programming?

The essence of functional programming

How to functional program
  Functional programming workflow
  Data types
  Type classes
  Type-directed programming
    Haskell style
  Refactoring (bonus section)

Type inference
Good Haskell style

Why it matters:
- layout is significant!
- eliminate misconceptions
- we care about *elegance*

Easy stuff:
- **use spaces!** (tabs cause layout errors)
- align patterns and guards

See style guides on course web page
Function application:

- is *just a space*
- associates to the left
- binds most strongly

Use parentheses only to *override* this behavior:

- $f (g \ x)$
- $f (x + y)$
Outline

Why learn functional programming?

The essence of functional programming

How to functional program
   Functional programming workflow
   Data types
   Type classes
   Type-directed programming
   Haskell style
   Refactoring (bonus section)

Type inference
Refactoring in the FP workflow

Motivations:
- separate concerns
- promote reuse
- promote understandability
- gain insights

“obsessive compulsive refactoring disorder”
Refactoring relations

Semantics-preserving laws prove with equational reasoning and/or induction

- Eta reduction:
  \[ \lambda x \rightarrow f \ x \equiv f \]

- Map–map fusion:
  \[ \text{map } f \ . \ \text{map } g \equiv \text{map } (f \ . \ g) \]

- Fold–map fusion:
  \[ \text{foldr } f \ b \ . \ \text{map } g \equiv \text{foldr } (f \ . \ g) \ b \]

“Algebra of computer programs”

John Backus, *Can Programming be Liberated from the von Neumann Style?*, ACM Turing Award Lecture, 1978
Strategy: systematic generalization

**commas :: [String] -> [String]**

*commas [] = []*

*commas [x] = [x]*

*commas (x:xs) = x : "", " : commas xs*

**Introduce parameters for constants**

**seps :: String -> [String] -> [String]**

*seps _ [] = []*

*seps _ [x] = [x]*

*seps s (x:xs) = x : s : seps s xs*

**Broaden the types**

**intersperse :: a -> [a] -> [a]**

*intersperse _ [] = []*

*intersperse _ [x] = [x]*

*intersperse s (x:xs) = x : s : intersperse s xs*
Strategy: abstract repeated templates

**abstract** (v): extract and make reusable (as a function)

```
showResult :: Maybe Float -> String
showResult Nothing = "ERROR"
showResult (Just v) = show v

moveCommand :: Maybe Dir -> Command
moveCommand Nothing = Stay
moveCommand (Just d) = Move d

safeAdd :: Int -> Maybe Int -> Int
safeAdd x Nothing = x
safeAdd x (Just y) = x + y
```
Strategy: abstract repeated templates

Describe repeated structure in function

```haskell
maybe :: b -> (a -> b) -> Maybe a -> b
maybe b _ Nothing = b
maybe _ f (Just a) = f a
```

Reuse in implementations

```haskell
showResult  = maybe "ERROR" show
moveCommand = maybe Stay Move
safeAdd x   = maybe x (x+)
```
Refactoring data types

data Expr = Var Name
  | Add Expr Expr
  | Sub Expr Expr
  | Mul Expr Expr

vars :: Expr -> [Name]
vars (Var x) = [x]
vars (Add l r) = vars l ++ vars r
vars (Sub l r) = vars l ++ vars r
vars (Mul l r) = vars l ++ vars r

eval :: Env -> Expr -> Int
eval m (Var x) = get x m
eval m (Add l r) = eval m l + eval m r
eval m (Sub l r) = eval m l - eval m r
eval m (Mul l r) = eval m l * eval m r
Refactoring data types

Factor out shared structure

```haskell
data Expr = Var Name
          | BinOp Op Expr Expr

data Op = Add | Sub | Mul

vars :: Expr -> [Name]
vars (Var x) = [x]
vars (BinOp _ l r) = vars l ++ vars r

eval :: Env -> Expr -> Int
eval m (Var x) = get x m
eval m (BinOp o l r) = op o (eval m l) (eval m r)
where
  op Add = (+)
  op Sub = (-)
  op Mul = (*)
```
Outline

Why learn functional programming?

The essence of functional programming

How to functional program

Type inference
How to perform type inference

If a literal, data constructor, or named function: write down the type – you’re done!

Otherwise:

1. expand any syntactic sugar
2. pick an application $e_1 e_2$
3. recursively infer their types $e_1 : T_1$ and $e_2 : T_2$
4. $T_1$ should be a function type $T_1 = T_{\text{arg}} \rightarrow T_{\text{res}}$
5. unify $T_{\text{arg}} = ? T_2$, yielding type variable assignment $\sigma$
6. return $e_1 e_2 : \sigma T_{\text{res}}$ (where $T_{\text{res}}$ with type variables substituted)

If any of these steps fails, it is a type error!