Winter 2020 Zi Hao He

CS442 Notes

1 Functional Programming

- to begin: Functional Programming in general, not a particular language
- need a model
  - simple
  - powerful enough to be useful

Church-Turing Thesis: Any algorithm can be simulated on a Turing machine.

Alan Turing

- Turing machine
- not an inspiring model for programming

Alonzo Church

- λ-calculus
- equivalent to Turing Machines
- the basis for all of functional programming

1.1 Untyped Lambda-Calculus

Calculus

- a system or method of calculation
- syntax + rules for manipulating syntax

Lambda-Calculus (1934) - ”a calculus of functions”

Syntax: Abstract Syntax

$$
\begin{aligned}
& variable, abstraction, application \\
< expr > &::=< var > \mid < abs > \mid < app > \\
< var > &::= a|b|c \\
& \text{variable and body} \\
< abs > &::= \lambda < var > . < expr > \\
& \text{rator and rand} \\
< app > &::=< expr >< expr >
\end{aligned}
$$

Use parenthesis to disambiguate parses

Conventions:

1. Abstractions extend as far to the right as possible

   e.g. $\lambda x.y\ z = \lambda x.(y\ z) \neq (\lambda x.y)\ z$

2. Applications associate left-to-right

   e.g. $x\ y\ z = (x\ y)\ z \neq x\ (y\ z)$

Interpretation

- Var: Self explanatory
- Abs: $\lambda x.E =$ "function" taking argument x and returning expr E
- App: $M\ N$ - result of applying "function" M to argument N

Consider:

$\lambda x.x$ - Here the first x is a binding occurrence and second x is a bound occurrence.

$\lambda x.x\ y\ x$ - The y is a free occurrence (fully parenthesized function is $\lambda x.((x\ y)\ x))$

$\lambda x.x\ (\lambda x.x)\ x$ - Same variable name might be bound in difference places as long as the scope is not clashing

$x\ \lambda x.x$ - The x in the front is free while the other one is bound

Informally:

A variable is bound if it has a bound occurrence.

A variable is free if it has a free occurrence

The same variable can be used in both bound and free occurrences

Formally:

Definition: Let E be an expression. The bound variables of E, $B \cup [E]$ are given by

$B \cup [x] = \emptyset$

$B \cup [\lambda x.E] = B \cup [E] \cup \{x\}$

$B \cup [M\ N] = B \cup [M] \cup B \cup [N]$

x is bound in E if $x \in B \cup [E]$

The free variables of E, $F \cup [E]$ are given by

$F \cup [x] = \{x\}$

$F\cup[\lambda x.E] = F\cup[E] \setminus \{x\}$

$F\cup[M\ N] = F\cup[M] \cup F\cup[N]$

x is free in E if $x \in F\cup[E]$

A variable can be both bound and free: $F\cup[x\ \lambda x.x] = B\cup[x\ \lambda x.x] = \{x\}$

But each occurrence is either bound or free, not both

Two occurrences of a variable x "mean the same thing" if

1. they are both free occurrences
2. OR they have the same binding occurrence

$\lambda x.\lambda y.y\ x$ and $\lambda a.\lambda b.b\ a$ these should mean the same thing

$\lambda x.y\ x$ and $\lambda x.w\ x$ should not mean the same thing because y and w do not have to be the same free variable

You can change the names of bound variables, but not free ones.

Formally:

Definition (α-conversion): For all x, y, M, $\lambda x.M =_\alpha \lambda y.M[y/x]$ if $y \notin F\cup[M]$

$M[N/x]$ = "substitute N for x in M"

More generally, if $C[\lambda x.M]$ denotes an expr in which $\lambda x.M$ occurs as a subexpression, then $C[\lambda x.M] =_\alpha C[\lambda y.M[y/x]]$ if $y \notin F\cup[M]$.

e.g.

$\lambda x.x =_\alpha \lambda y.y$

$x\ \lambda x.x =_\alpha x\ \lambda a.a$

$\lambda a.b\ a =_\alpha \lambda c.b\ c \neq_\alpha \lambda b.b\ b \neq_\alpha \lambda a.d\ a$

Computation:

Expr. $(\lambda x.M)\ N$ is called a (β-)redex (reductible expression)

We expect: $(\lambda x.M)\ N \Rightarrow$ evaluate M, with N substituted for x i.e. $M[N/x]$

Definition (β-reduction): For all M, N, x, $(\lambda x.M)\ N \rightarrow_\beta M[N/x]$

More generally - for all contexts C, $C[(\lambda x.M)\ N] \to_\beta C[M[N/x]]$

$\to_\beta$ is a binary relation on terms

$A \to_\beta B$: A beta reduces to B in one step

$A \to_\beta^n B$: A beta reduces to B in n steps

$A \to_\beta^* B$: A beta reduces to B in 0 or more steps

$A \to_\beta^+ B$: A beta reduces to B in 1 or more steps

Evaluating $M[N/x]$

Want: Substitute N for all free occurrences of x in M

Definition (substitution, naive(wrong)):

$x[E/x] = E$

$y[E/x] = y$, $y \neq x$

$(M\ N)[E/x] = M[E/x]\ N[E/x]$

$(\lambda x.P)[E/x] = \lambda x.P$

$(\lambda y.P)[E/x] = \lambda y.P[E/x]$, $y \neq x$

E.g.

$$\begin{aligned}(\lambda x.x\ y)\ z &\to_\beta (x\ y)[z/x] \\ &= x[z/x]\ y[z/x] \\ &= z\ y\end{aligned}$$

$$(\lambda x.x)\ a \to_\beta x[a/x] = a$$

$$\begin{aligned}(\lambda x.\lambda y.x)\ a\ b &\to_\beta (\lambda y.x)[a/x]\ b \\ &= (\lambda y.x[a/x])\ b \\ &= (\lambda y.a)\ b \\ &\to_\beta a[b/y] \\ &= a\end{aligned}$$

$$\begin{aligned}(\lambda x.\lambda y.y)\ a\ b &\to_\beta (\lambda y.y)[a/x]\ b \\ &= (\lambda y.y[a/x])\ b \\ &= (\lambda y.y)\ b \\ &\to_\beta y[b/y] \\ &= b\end{aligned}$$

Consider:

$$\begin{aligned}(\lambda x.\lambda y.x)\ y\ w &\to_\beta (\lambda y.x)[y/x]\ w\\ &= (\lambda y.x[y/x])\ w\\ &= (\lambda y.y)\ w\\ &\to_\beta y[w/y]\\ &= w\end{aligned}$$

We can see from this that the last rule of $(\lambda y.P)[E/x] = \lambda y.P[E/x]$, $y \neq x$ must be wrong.

What happened? Free variable y became bound after substitution

As a result, the binding occurrence of x changed

$\lambda x.\lambda y.x \leftarrow$ the inner x is bound to the λx on the outside.

This is called **Dynamic binding** - meaning of variables uncertain until runtime

We want **static binding** - meanings of variables fixed before runtime

Definition (substitution, fixed):

$x[E/x] = E$

$y[E/x] = y$, $y \neq x$

$(M\ N)[E/x] = M[E/x]\ N[E/x]$

$(\lambda x.P)[E/x] = \lambda x.P$

$(\lambda y.P)[E/x] = \lambda y.P[E/x]$, $y \notin F \cup [E]$

$(\lambda y.P)[E/x] = \lambda z.(P[z/y][E/x])$, $y \in F \cup [E]$, z a "fresh" variable

Now

$$\begin{aligned}(\lambda x.\lambda y.x)\ y\ w &\to_\beta (\lambda y.x)[y/x]\ w\\ &= (\lambda z.x[z/y][y/x])\ w\\ &= (\lambda z.x[y/x])\ w\\ &= (\lambda z.y)\ w\\ &\to_\beta y[w/z]\\ &= y\end{aligned}$$

Computation: β-reduction until you reach a **Normal Form**.

Definition (β-Normal Form): An expr is in β Normal Form if it has no β-redices.

Definition (Weak Normal Form): An expr is in Weak Normal Form if the only β-redices are inside abstractions. e.g. $\lambda z.(\lambda x.x)\ y$.

Usually, "Normal Form" will mean β-Normal Form.

Consider:

$$\begin{aligned}(\lambda x.x)((\lambda y.y)\ z) &\to_\beta x[((\lambda y.y)\ z)/x]\\ &= (\lambda y.y)\ z\\ &\to_\beta y[z/y]\\ &= z\end{aligned}$$

Or

$$\begin{aligned}(\lambda x.x)((\lambda y.y)\ z) &\to_\beta (\lambda x.x)(y[z/y])\\ &= (\lambda x.x)\ z\\ &\to_\beta x[z/x]\\ &= z\end{aligned}$$

- two difference reduction sequences. Does it matter which we take?

Theorem (Church-Rosser): Let E_1, E_2, E_3 be expressions such that $E_1 \to_\beta^* E_2$ and $E_1 \to_\beta^* E_3$. Then there is an expression E_4 such that (up to α-equivalence) $E_2 \to_\beta^* E_4$ and $E_3 \to_\beta^* E_4$.

Immediate Consequence - An expr E can have at most one β-Normal Form (modulo α-equivalence)

Do all expressions have a β-NF? No!

Consider:

$$\begin{aligned}(\lambda x.x\ x)(\lambda x.x\ x) &\to_\beta (x\ x)[(\lambda x.x\ x)/x]\\ &= (\lambda x.x\ x)(\lambda x.x\ x)\end{aligned}$$

It does not have a β-NF because it always reduces to itself.

Which exprs have a β-NF? - undecidable - equivalent to Halting Problem

Consider:

$$(\lambda x.y)((\lambda x.x\ x)(\lambda x.x\ x)) \to_\beta (\lambda x.y)((\lambda x.x\ x)(\lambda x.x\ x)) \to_\beta \cdots$$

Or, the alternate way of reducing this

$$\begin{aligned}(\lambda x.y)((\lambda x.x\ x)(\lambda x.x\ x)) &\to_\beta y[(\lambda x.x\ x)(\lambda x.x\ x)/x]\\ &= y\end{aligned}$$

$\therefore$ Order does matter when ∞-reductions are possible.

Reduction Strategies - "plans" for choosing a redex to reduce.

Applicative Order Reductions (AOR): Choose the leftmost, innermost redex.

- Innermost = not containing any other redex
- This is called "eager evaluation"

Normal Order Reduction (NOR): Choose the leftmost,outermost

- outermost = not contained in any other redex
- "lazy evaluation"

To the example earlier, the one the results in infinite reduction is AOR and the one that go the β-NF form is NOR.

e.g.

$$(\lambda x.x)((\lambda y.y)\ z) \xrightarrow{\text{AOR}}_\beta (\lambda x.x)\ z \xrightarrow{\text{AOR}}_\beta z$$
$$(\lambda x.x)((\lambda y.y)\ z) \xrightarrow{\text{NOR}}_\beta (\lambda y.y)\ z \xrightarrow{\text{NOR}}_\beta z$$

Theorem (Standardization): If an expr as a β-NF, then NOR is guaranteed to reach it.

But most proramming languages use applicative order, including Scheme.

η-reduction: Consider $(\lambda x.y\ x)\ z \rightarrow_\beta (y\ x)[z/x] = y\ z$

So $\lambda x.y\ x$ behaves exactly like z

Definition (η-reduction): $\lambda x.M\ x \rightarrow_\eta M$ if $x \notin F \cup [M]$

More generally, $C[\lambda x.M\ x] \rightarrow_\eta C[M]$ if $x \notin F \cup [M]$

So we have $\lambda x.y\ x \rightarrow_\eta y$

1.2 Programming in the λ-Calculus

Shorthand for convenience: $[[\cdot]]$: (real-world programming language) $\rightarrow$ λ-calculus

For a real-world expr E, $[[E]]$ is our representation of E in the λ-calculus

e.g. $[[id]] = \lambda x.x$

Note: $[[\cdot]]$ is just shorthand. An expression containing $[[\cdot]]$ is not λ-calculus until all $[[\cdot]]$s have been replaced with what they represent.

Booleans: Let [[true]] $= \lambda x.\lambda y.x$ and [[false]] $= \lambda x.\lambda y.y$.

if ¡bool¿ then ¡true-part¿ else ¡false-part¿

if(b, t, f) = if b then t else f

[[if]] $= \lambda b.\lambda t.\lambda f.b\ t\ f$

Note that $\lambda b.\lambda t.\lambda f.b\ t\ f \rightarrow_{\eta}^{2} \lambda b.b$

Alternatively, [[if B then T else F]] $= [[B]][[T]][[F]]$

Does it work?

$$\begin{aligned}
\text{[[if true then P else Q]]} &= [[true]][[P]][[Q]] \\
&= (\lambda x.\lambda y.x)[[P]][[Q]] \\
&\rightarrow_{\beta} [[P]]
\end{aligned}$$

$$\begin{aligned}
\text{[[if false then P else Q]]} &= [[false]][[P]][[Q]] \\
&= (\lambda x.\lambda y.y)[[P]][[Q]] \\
&\rightarrow_{\beta} [[Q]]
\end{aligned}$$

Now, for the definition of **Not**:

$$\begin{aligned}
[[not]] &= \lambda b.\text{[[if b then false else true]]} \\
&= \lambda b.b[[false]][[true]] \\
&= \lambda b.b(\lambda x.\lambda y.y)(\lambda x.\lambda y.x)
\end{aligned}$$

We can test it:

$$\begin{aligned}
\text{[[not true]]} &= (\lambda b.b[[false]][[true]])\ [[true]] \\
&\rightarrow_{\beta} [[true]][[false]][[true]] \\
&= (\lambda x.\lambda y.x)[[false]][[true]] \\
&\rightarrow_{\beta}^{2} [[false]]
\end{aligned}$$

For the definition of **And**:

$$
\begin{aligned}
[[and]] &= \lambda p.\lambda q.[[\text{if p then q else false}]] \\
&= \lambda p.\lambda q.p\ q\ [[false]]
\end{aligned}
$$

For the definition of **Or**:

$$
\begin{aligned}
[[and]] &= \lambda p.\lambda q.[[\text{if p then true else q}]] \\
&= \lambda p.\lambda q.p\ [[true]]\ q
\end{aligned}
$$

Storage: Use lists. Scheme - lists based on pairs.

(cons a b) creates the pair

(cons a (cons b c)) creates the list

$\therefore$ nested pairs create lists

Need to implement:

- cons
- nil - empty list
- null? - is the list empty?
- car - first component of the pair
- cdr - 2nd component of the pair

Modelling pairs: pair - function that takes a **selector** as a parameter

If the selector is true - return the first component

If the selector is true - return the second component

i.e. $[[pair]] = \lambda s.[[\text{if s then h else t}]]$ where s is the selector, h is the head, t is the tail.

So $[[pair]] = \lambda s.s\ h\ t$

Then $[[cons]] = \lambda h.\lambda t.\lambda s.s\ h\ t$

e.g.

$$
\begin{aligned}
[[\text{cons a (cons b nil)}]] &= [[cons]]\ a\ ([[cons]]\ b\ [[nil]]) \\
&= (\lambda h.\lambda t.\lambda s.s\ h\ t)\ a\ ([[cons]]\ b\ [[nil]]) \\
&\to_\beta^2 \lambda s.s\ a\ ([[cons]]\ b\ [[nil]]) \\
&= \lambda s.s\ a((\lambda h.\lambda t.\lambda s.s\ h\ t)\ b\ [[nil]]) \\
&\to_\beta^2 \lambda s.s\ a\ (\lambda s.s\ b\ [[nil]])
\end{aligned}
$$

car: return the head ⇒ pass the selector "true"

$[[car]] = \lambda l.l\ [[true]] = \lambda l.l\ (\lambda x.\lambda y.x)$

Similarly, **cdr:** return the tail ⇒ pass the selector "false"

$[[cdr]] = \lambda l.l\ [[false]] = \lambda l.l\ (\lambda x.\lambda y.y)$

null?

- must return false when given a pair $(\lambda s.s\ h\ t)$

- pass a selector that consumes h and t, returns false

i.e. $[[null?]] = \lambda l.l(\lambda a.\lambda b.[[false]]) = \lambda l.l(\lambda a.\lambda b.\lambda x.\lambda y.y)$

$$\begin{aligned}
[[null?]](\lambda s.s\ h\ t) &= (\lambda l.l(\lambda a.\lambda b.\lambda x.\lambda y.y))(\lambda s.s\ h\ t) \\
&\to_\beta (\lambda s.s\ h\ t)(\lambda a.\lambda b.\lambda x.\lambda y.y) \\
&\to_\beta (\lambda a.\lambda b.\lambda x.\lambda y.y)\ h\ t \\
&\to_\beta^2 \lambda x.\lambda y.y \\
&= [[false]]
\end{aligned}$$

nil: something to make null? return true.

$[[nil]] = \lambda s.[[true]] = \lambda s.\lambda x.\lambda y.x$

$$\begin{aligned}
[[null?]][[nil]] &= (\lambda l.l(\lambda a.\lambda b.\lambda x.\lambda y.y))(\lambda s.\lambda x.\lambda y.x) \\
&\to_\beta (\lambda s.\lambda x.\lambda y.x)(\lambda a.\lambda b.\lambda x.\lambda y.y) \\
&\to_\beta \lambda x.\lambda y.x \\
&= [[true]]
\end{aligned}$$

Numbers

- Consider only non-negative integers

- Easy - encode n as a list of length n

  $[[0]] = [[nil]]$
  $[[1]] = \lambda s.s\ ?\ [[nil]]$ where ? is just whatever
  $[[2]] = \lambda s.s\ ?\ (\lambda s.s\ ?\ [[nil]])$

Primitives: $[[pred]], [[succ]], [[isZero?]]$

$[[isZero?]] = [[null?]]$

$[[pred]] = [[cdr]]$

$[[succ]] = \lambda n.\lambda s.s\ ?\ n (= \lambda n.[[cons]]\ ?\ n)$

Exercise: $[[pred\ (succ\ m)]] =_\beta m$

Clever Solution: Church Numberals

Represent n as the act of applying a function f n times to an argument x.

$[[0]] = \lambda f.\lambda x.x$

$[[1]] = \lambda f.\lambda x.f\ x$

$[[2]] = \lambda f.\lambda x.f\ (f\ x)$

$[[3]] = \lambda f.\lambda x.f\ (f\ (f\ x))$

$[[n]]\ f\ x = f^n(x)$

addition: $m + n$ - apply f n times to x, then apply f m times to the result

$[[+]] = \lambda m.\lambda n.\lambda f.\lambda x.m\ f\ (n\ f\ x)$

$$
\begin{aligned}
[[+\ 2\ 3]] &= (\lambda m.\lambda n.\lambda f.\lambda x.m\ f\ (n\ f\ x))(\lambda f.\lambda x.f\ (f\ x))(\lambda f.\lambda x.f(f(f\ x))) \\
&\to_\beta^2 \lambda f.\lambda x(\lambda f.\lambda x.f(f\ x))\ f\ ((\lambda f.\lambda x.f(f\ x)))\ f\ x) \\
&\to_\beta \lambda f.\lambda x.(\lambda x.f(f\ x))(\lambda f.\lambda x.f(f(f\ x)))\ f\ x) \\
&\to_\beta \lambda f.\lambda x.f(f((\lambda f.\lambda x.f(f(f\ x)))\ f\ x)) \\
&\to_\beta \lambda f.\lambda x.f(f((\lambda x.f(f(f(\ x)))\ x) \\
&\to_\beta \lambda f.\lambda x.f(f(f(f(f\ x)))) \\
&= [[5]]
\end{aligned}
$$

Special case: $[[succ]] = \lambda n.\lambda f.\lambda x.n\ f\ (f\ x)$ (or $f(n\ f\ x)$)

Subtraction: Harder - find a function to apply n times to produce $[[n-1]]$

Consider: Start with $[[cons\ 0\ 0]]$

apply the function $f : [[cons\ a\ b]] \to [[cons\ a+1\ a]]$

n times: $[[cons\ n\ n-1]]$ - then take the cdr.

$[[pred]] = \lambda n.[[cdr]]\ (n\ (\lambda p.[[cons]]([[succ]][[car\ p]])\ [[car\ p]])([[cons]][[0]][[0]]))$

$[[-]] = \lambda m.\lambda n.n\ [[pred]]\ m$

Multiplication: Apply the n-fold repetition of f, m times

$[[*]] = \lambda m.\lambda n.\lambda f.\lambda x.m\ (n\ f)\ x =_\eta \lambda m.\lambda n.\lambda f.m\ (n\ f)$

Since $n\ f = f^n$, then $m\ (n\ f) = (f^n)^m$

Notice that if you change m and n to f and g, and f to x, you get $f(g(x))$ which is just function composition.

Exponentiation: m^n - the n-fold repetition of m itself

$[[\hat{}]] = \lambda m.\lambda n.\lambda f.\lambda x.n\ m\ f\ x =_\eta \lambda m.\lambda n.n\ m$

Recursion: find the length of a list

$$\begin{aligned}[[len]] &= \lambda l.[[\text{if (null? } l\text{) 0 (succ (len (cdr l)))}]] \\ &= \lambda l.([[\text{null?}]]\ l)\ [[0]]\ ([[\text{succ}]]\ ([[\text{len}]]\ ([[\text{cdr}]]\ l)))\end{aligned}$$

WRONG! [[len]] defined in terms of itself.

How do we get a closed-form representation of len? Solve the equation.

Aside: fixed points. A **fixed point** of a function f is a value x such that $x = f(x)$

e.g. $f(x) = x^2 - 6$ has a fixed point of 3 since $f(3) = 3$

$$\begin{aligned}[[len]] &= \lambda l.([[null?]]\ l)[[0]]([[succ]]([[len]]([[cdr]]\ l))) \\ &\leftarrow_\beta (\lambda f.\lambda l.([[null?]]\ l)[[0]]([[succ]](f\ ([[cdr]]\ l))))\ [[len]] \\ &=_\beta F([[len]]) \text{ where F} = (\lambda f.\lambda l.([[null?]]\ l)[[0]]([[succ]](f\ ([[cdr]]\ l))))\end{aligned}$$

[[len]] satisfies $[[len]] = F([[len]])$

$\therefore$ Need a fixed point of F

Consider:

$$\begin{aligned}X &= (\lambda x.f(x\ x))(\lambda x.f(x\ x)) \\ &\rightarrow_\beta f((\lambda x.f(x\ x))(\lambda x.f(x\ x))) \\ &= f(X)\end{aligned}$$

$\therefore$ X is a fixed point of f

Now parameterize by f, get

$$Y = \lambda f.(\lambda x.f(x\ x))(\lambda x.f(x\ x))$$

This is **Curry's Paradoxical Combinator** (or Y combinator)

- returns the fixed point of any function

$$\begin{aligned} Y\ g &= (\lambda f.(\lambda x.f(x\ x))(\lambda x.f(x\ x)))\ g \\ &\to_\beta (\lambda x.g(x\ x))(\lambda x.g(x\ x)) \\ &\to_\beta g((\lambda x.g(x\ x))(\lambda x.g(x\ x))) \\ &\to_\beta g(Y\ g) \end{aligned}$$

$\therefore$ for any g, $Y\ g =_\beta g(Y\ g)$. i.e. Y g is a fixed point of g.

Any combinator (closed expression) C such that for all g, $C\ g =_\beta g(C\ g)$ is called a **fixed-point combinator**.

$\therefore$ $[[len]] = Y\ F = Y(\lambda f.\lambda l.([[null?]]\ l)\ [[0]]\ ([[succ]]\ (f\ ([[cdr]]\ l)))$