

# COMPSCI 3M13 - Principles of Programming Languages

## Topic 6 - Untyped Lambda Calculus 2 : Extensions and Formalities

NCC Moore

McMaster University

Fall 2021

Adapted from “Types and Programming Languages” by Benjamin  
C. Pierce

## More UAE Terms in Lambda Calculus

## Enriching the Calculus

## Turing Completeness and Lambda Calculus

## Formalities

## Extras

# More UAE Terms in $\lambda$ Calculus

Over the course of this topic, we will continue to expand our knowledge of  $\lambda$ -Calculus.

- ▶ We will polish off the last remaining terms of UAE.
- ▶ We will discuss the **enrichment** of the calculus, so that it is less of a pain to work with.
- ▶ We will examine the ways in which  $\lambda$ -Calculus can be used as a general computational engine.
- ▶ We will discuss the formalities of the language.

# isZero

In order to test an expression to see if it is  $c_0$  or not, we must find arguments for the Church numerals which yield `true` if no successors have been applied, and `false` otherwise.

▶ Here's a reminder of what Church numerals look like:

$$c_0 = \lambda s. \lambda z. z \quad (1)$$

$$c_1 = \lambda s. \lambda z. s \ z \quad (2)$$

$$c_2 = \lambda s. \lambda z. s \ (s \ z) \quad (3)$$

$$c_3 = \lambda s. \lambda z. s \ (s \ (s \ z)) \quad (4)$$

$$\vdots$$

# Coke Zero

To design a function that returns `tru` or `fls` when applied to Church numerals, we need to find some **inputs** to a church numeral which yield the correct **output**.

- ▶  $c_0$  simply echos the second argument, so making the second argument `tru` will yield `iszro  $c_0$  = tru`
- ▶ We might observe that each numeral that would return `fls` applies  $z$  to one or more  $s$ .
- ▶ So, if we make  $s$  something that always returns `fls`, no matter what is applied to it, we have success!
  - ▶  $\lambda x. fls$  is the simplest function that fits the above description.

## Zero to Hero

So, our `iszero` function needs to take a church numeral, and apply the above functions to it.

$$\text{iszero} = \lambda m.m : (\lambda x.\text{fls}) \text{tru} \quad (5)$$

$$\begin{aligned} & \frac{\text{iszero } c_0}{(\lambda m.m (\lambda x.\text{fls}) \text{tru}) c_0} \\ \rightarrow & c_0 (\lambda x.\text{fls}) \text{tru} \\ \rightarrow & (\lambda s.\lambda z.z) (\lambda x.\text{fls}) \text{tru} \\ \rightarrow & (\lambda z.z) \text{tru} \\ \rightarrow & \text{tru} \\ \rightarrow & \text{ } \end{aligned}$$

# The Mask of Zero

$$\begin{aligned}
 & \quad \quad \quad \underline{\text{iszero } c_2} \\
 & (\lambda m.m (\lambda x.\text{fls}) \text{tru}) c_2 \\
 \rightarrow & \quad c_2 (\lambda x.\text{fls}) \text{tru} \\
 \rightarrow & \quad (\lambda s.\lambda z. s (s z)) (\lambda x.\text{fls}) \text{tru} \\
 \rightarrow & \quad (\lambda z. (\lambda x.\text{fls}) ((\lambda x.\text{fls}) z)) \text{tru} \\
 \rightarrow & \quad (\lambda x.\text{fls}) ((\lambda x.\text{fls}) \text{tru}) \\
 \rightarrow & \quad \text{fls} \\
 \rightarrow & \quad \text{↯}
 \end{aligned}$$

# Predecessor(!)

Testing to see if something is zero is relatively straightforward, but predecessor requires some cleverness.

- ▶ In UAE, we defined `pred` as an annihilation operation over successors.
- ▶ In  $\lambda$ -Calculus, we essentially need to *reconstruct our numeral*, while keeping a *history of the previous value*.

$$\text{prd} = \lambda m. \text{fst } (m \text{ ss } \text{zz}) \quad (6)$$

Where

$$\text{ss} = \lambda p. \text{pair } (\text{snd } p) (\text{plus } c_1 (\text{snd } p)) \quad (7)$$

$$\text{zz} = \text{pair } c_0 \ c_0 \quad (8)$$



# Predecessor of $c_2$

$$\begin{aligned}
 & \text{pred } c_2 \\
 & (\lambda m.fst (m \ ss \ zz)) \ c_2 \\
 \rightarrow & \ fst (c_2 \ ss \ zz) \\
 \rightarrow & (\lambda p.p \ tru) (c_2 \ ss \ zz) \\
 \rightarrow & c_2 \ ss \ zz \ tru \\
 \rightarrow & (\lambda s.\lambda z.s (s \ z)) \ ss \ zz \ tru \\
 \rightarrow & (\lambda z.ss (ss \ z)) \ zz \ tru \\
 \rightarrow & ss (ss \ zz) \ tru \\
 \rightarrow & (\lambda p.pair (\snd \ p) (\text{plus } c_1 (\snd \ p))) (ss \ zz) \ tru \\
 \rightarrow & pair (\snd (ss \ zz)) (\text{plus } c_1 (\snd (ss \ zz))) \ tru \\
 \rightarrow & (\lambda f.\lambda s.\lambda b.b \ f \ s) (\snd (ss \ zz)) (\text{plus } c_1 (\snd (ss \ zz))) \ tru \\
 \rightarrow \rightarrow \rightarrow & tru (\snd (ss \ zz)) (\text{plus } c_1 (\snd (ss \ zz))) \\
 \rightarrow & (\lambda t.\lambda f.t) (\snd (ss \ zz)) (\text{plus } c_1 (\snd (ss \ zz))) \\
 \rightarrow \rightarrow & \snd (ss \ zz)
 \end{aligned}$$

# Predecessor of $c_2$

$$\begin{aligned}
 & \text{snd } (ss \text{ } zz) \\
 \rightarrow & (\lambda p. p \text{ fls}) (ss \text{ } zz) \\
 \rightarrow & ss \text{ } zz \text{ fls} \\
 \rightarrow & (\lambda p. \text{pair } (\text{snd } p) (\text{plus } c_1 (\text{snd } p))) \text{ } zz \text{ fls} \\
 \rightarrow & \text{pair } (\text{snd } zz) (\text{plus } c_1 (\text{snd } zz)) \text{ fls} \\
 \rightarrow & (\lambda f. \lambda s. \lambda b. b \text{ } f \text{ } s) (\text{snd } zz) (\text{plus } c_1 (\text{snd } zz)) \text{ fls} \\
 \rightarrow \rightarrow \rightarrow & \text{fls } (\text{snd } zz) (\text{plus } c_1 (\text{snd } zz)) \\
 \rightarrow \rightarrow & \text{plus } c_1 (\text{snd } zz) \\
 \rightarrow & (\lambda m. \lambda n. \lambda s. \lambda z. m \text{ } s \text{ } (n \text{ } s \text{ } z)) \text{ } c_1 (\text{snd } zz) \\
 \rightarrow \rightarrow & \lambda s. \lambda z. c_1 \text{ } s \text{ } ((\text{snd } zz) \text{ } s \text{ } z) \\
 \rightarrow &
 \end{aligned}$$

## Latent Evaluations

Because we are using the call by value strategy, we actually can't evaluate any further than this.

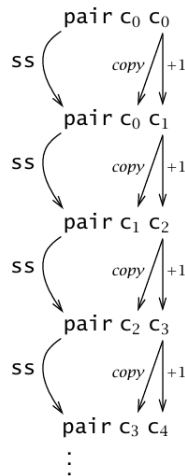
- ▶ Remember that call by value doesn't allow us to evaluate the subexpressions of a term with a lambda.
- ▶ This leads to an interesting property of  $\lambda$ -Calculus: **Latent Evaluations**.
- ▶ Undersupplied  $\lambda$  expressions won't fully execute, in effect, storing evaluations for later retrieval.
- ▶ These evaluations will only be calculated if the expression is supplied with a sufficient number of arguments.
- ▶ We often say these forms are **functionally equivalent** to the normal forms that would arise from full  $\beta$ -reduction, and use them interchangeably.

# Continuing using normal order...

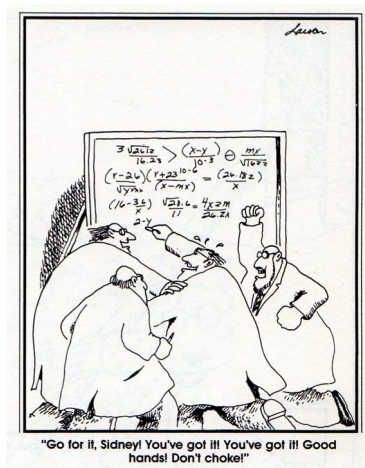
$$\begin{aligned}
 & \lambda s. \lambda z. c_1 s ((\text{snd } zz) s z) \\
 \rightarrow & \lambda s. \lambda z. (\lambda s. \lambda z. s z) s ((\text{snd } zz) s z) \\
 \rightarrow \rightarrow & \lambda s. \lambda z. s (\text{snd } zz s z) \\
 \rightarrow & \lambda s. \lambda z. s ((\lambda p. p \text{ fls}) zz s z) \\
 \rightarrow & \lambda s. \lambda z. s (zz \text{ fls } s z) \\
 \rightarrow & \lambda s. \lambda z. s (\text{pair } c_0 c_0 \text{ fls } s z) \\
 \rightarrow & \lambda s. \lambda z. s ((\lambda f. \lambda s. \lambda b. b f s) c_0 c_0 \text{ fls } s z) \\
 \rightarrow \rightarrow \rightarrow & \lambda s. \lambda z. s (\text{fls } c_0 c_0 s z) \\
 \rightarrow \rightarrow & \lambda s. \lambda z. s (c_0 s z) \\
 \rightarrow & \lambda s. \lambda z. s (c_0 s z) \\
 \rightarrow & \lambda s. \lambda z. s ((\lambda s. \lambda z. z) s z) \\
 \rightarrow \rightarrow & \lambda s. \lambda z. s z \\
 \rightarrow & c_1 \\
 \rightarrow &
 \end{aligned}$$

## For higher numerals...

- ▶ The way this algorithm works is to start at  $c_0$ , and build our way up to the number we're trying to take the predecessor of.
- ▶ One half of the pair keeps track of the last numeral we were on, so when we reach the numeral we're trying to take the predecessor of...
- ▶ We just need to skim off the first element.



# Enriching the $\lambda$ -Calculus



## That Was Pretty Painful!

As the previous, 3 full slide derivation has demonstrated,  $\lambda$ -Calculus can be pretty painful to do any actual work in!

- ▶ Wouldn't it be convenient if there was some sort of advanced electrical computational engine that could perform such calculations...
- ▶ Hundreds even, in the blink of an eye!

Hopefully we are now satisfied that all of our usual values and operators have some expression under  $\lambda$ -Calculus. For convenience of calculation, let's move away from the pure system, and add some additional semantic content.

## Adding UAE

We can convert easily between terms in UAE and our previously defined  $\lambda$  expression equivalents:

$$\text{realbool} = \lambda b.b \text{ true false} \quad (9)$$

$$\text{churchbool} = \lambda b.\text{if } b \text{ then tru else fls} \quad (10)$$

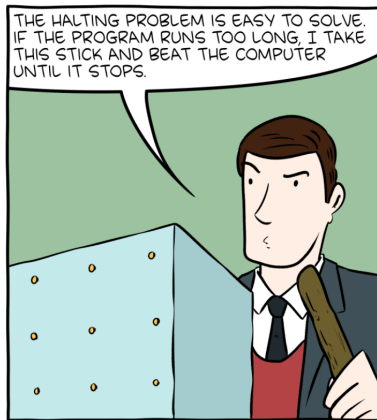
$$\text{realnat} = \lambda c_n.c_n (\lambda x.\text{succ } x) 0 \quad (11)$$

$$\text{churchnat} = \lambda n.(\lambda \text{succ}.\lambda 0.n) s z \quad (12)$$

From here, it's just a matter of using the right operations on the right values.



# Turing Completeness and Lambda Calculus



What if Alan Turing had been an engineer?

# Turing Completeness/Equivalence

To review, a system of computation is considered **Turing Complete** or Turing Equivalent if it can perform all the actions of a Turing machine. The minimal set of things you need to be able to do is:

- ▶ Support conditional branching.
  - ▶ This implies support for arbitrary go-to operations.
- ▶ An infinite amount of tape (or memory).

Technically, no physical computer is Turing Complete, because of physical constraints on memory.

# Church-Turing Thesis

The Church-Turing Thesis states:

- ▶ A function on the natural numbers can be calculated by an effective method if and only if it is computable by a Turing machine.
- ▶ A secondary effect of this is that a program is computable via a Turing Machine iff it is computable using a  $\lambda$  expression.

UAE is not Turing Complete because the Theorem of Evaluation holds.

- ▶ This theorem means a UAE expression's evaluation chain must be finite for a finite term.
- ▶ We can construct a finite Turing Machine which runs infinitely.

## Curious Constructions

Theorem of Evaluation does not hold for  $\lambda$ -Calculus!

- ▶ This doesn't mean it isn't determinate, because it is (depending on your evaluation strategy).

This only means that finite  $\lambda$  expressions can have infinite evaluation chains, such as the  $\Omega$ -Function:

$$\Omega = (\lambda x.x x)(\lambda x.x x) \quad (13)$$

When you  $\beta$ -reduce  $\Omega$ , you get  $\Omega$  right back again!

$$(\lambda x.x x)(\lambda x.x x) \rightarrow (\lambda x.x x)(\lambda x.x x) \quad (14)$$

Because these functions do not converge on a normal form, they are known as **divergent**.

# $R(e(c(u(r(s(i(o(n())))))))))))$

The  $\Omega$ -Function is an interesting function, but it isn't very practical.

- ▶ Its cousin, the Y-Combinator, encodes general recursion in the  $\lambda$ -Calculus.

$$Y = \lambda f. (\lambda x. f (x x)) (\lambda x. f (x x)) \quad (15)$$

- ▶ Unfortunately, it only works under call by name. The following **fixed-point combinator** solves the problem of general recursion for the call by value evaluation strategy.

$$\text{fix} = \lambda f. (\lambda x. f (\lambda y. x x y)) (\lambda x. f (\lambda y. x x y)) \quad (16)$$

## Factorial Again

Recall the factorial function:

$$n! = \begin{cases} 1 & n = 0 \\ n \times (n-1)! & n > 0 \end{cases} \quad (17)$$

We can now encode it as follows.

$$g = \lambda \text{fct} . \lambda n . \text{if } n == 0 \text{ then } 1 \text{ else } n \times (\text{fct } (n-1)) \quad (18)$$

$$\text{factorial} = \text{fix } g \quad (19)$$

To save time and energy, we are encoding this using the enriched calculus.

# Example Time!

factorial 3

fix  $g$  3

→  $(\lambda f.(\lambda x.f (\lambda y.x \times y)) (\lambda x.f (\lambda y.x \times y))) g$  3

→  $(\lambda x.g (\lambda y.x \times y)) (\lambda x.g (\lambda y.x \times y))$  3

setting  $h = \lambda x.g (\lambda y.x \times y)$

yields  $(\lambda x.g (\lambda y.x \times y)) h$  3

→  $g (\lambda y.h h y)$  3

setting  $fct = \lambda y.h h y$

yields  $g fct$  3

→  $(\lambda fct.\lambda n.\text{if } n == 0 \text{ then } 1 \text{ else } n \times (fct (n - 1))) fct$  3

→ →  $\text{if } 3 == 0 \text{ then } 1 \text{ else } 3 \times (fct (3 - 1))$

→ →  $3 \times (fct 2)$

→  $3 \times ((\lambda y.h h y) 2)$

→  $3 \times (h h 2)$

$3 \times (h\ h\ 2)$   
 $\rightarrow 3 \times ((\lambda x.g\ (\lambda y.x\ x\ y))\ h\ 2)$   
 $\rightarrow 3 \times (g\ fct\ 2)$   
 $\rightarrow \rightarrow 3 \times (g\ fct\ 2)$   
 $\rightarrow \rightarrow \rightarrow 3 \times (\text{if } 2 == 0 \text{ then } 1 \text{ else } 2 \times (fct\ (2 - 1)))$   
 $\rightarrow \rightarrow 3 \times 2 \times (fct\ 1)$   
 $\rightarrow \rightarrow \rightarrow 6 \times (h\ h\ 1)$   
 $\rightarrow \rightarrow 6 \times (g\ fct\ 1)$   
 $\rightarrow \rightarrow \rightarrow 6 \times (\text{if } 1 == 0 \text{ then } 1 \text{ else } 1 \times (fct\ (1 - 1)))$   
 $\rightarrow \rightarrow 6 \times 1 \times (fct\ 0)$   
 $\rightarrow \rightarrow \rightarrow 6 \times (h\ h\ 0)$   
 $\rightarrow \rightarrow 6 \times (g\ fct\ 0)$   
 $\rightarrow \rightarrow \rightarrow 6 \times (\text{if } 0 == 0 \text{ then } 1 \text{ else } 0 \times (fct\ (0 - 1)))$   
 $\rightarrow \rightarrow \rightarrow 6$   
 $\nrightarrow$



# Formalities



“There may, indeed, be other applications of the system than its use as a logic.”

- Alonzo Church, 1932 -

## Inductive Syntax of $\lambda$ -Calculus

For the rest of this topic, we will examine the subtleties of a more rigorous definition of the  $\lambda$  calculus, beginning with an inductive definition of its syntax.

Let  $\mathcal{V}$  be a countable set of variable names. The set of terms is the smallest set  $\mathcal{T}$  such that:

1.  $\mathcal{V} \subseteq \mathcal{T}$
2.  $t_1 \in \mathcal{T} \wedge x \in \mathcal{V} \implies \lambda x. t_1 \in \mathcal{T}$
3.  $t_1, t_2 \in \mathcal{T} \implies t_1 t_2 \in \mathcal{T}$

- ▶ Via this definition, we can define size and depth the same way as we did under UAE.

## Free Variables with Every Order

We can define a new function over  $\lambda$ -Calculus, in the style of the `consts` operator of UAE.

The set of *free variables* of a term  $t$ , written  $FV(t)$  is defined as follows:

$$\begin{aligned} FV(x) &= \{x\} \\ FV(\lambda x. t_1) &= FV(t_1) \setminus \{x\} \\ FV(t_1 t_2) &= FV(t_1) \cup FV(t_2) \end{aligned}$$

# Use Substitute!

At the beginning of our discussion of the  $\lambda$ -Calculus, we said it would be necessary to develop a semantic of both the calculus itself, and the substitution operation. Let's start with substitution.

- ▶ We will start with an intuitive definition based on our knowledge of elementary-school algebra, and develop a more robust definition by exposing issues with the naive approach.

Let us define naive substitution as follows:

$$\begin{aligned} [x \mapsto s]x &= s \\ [x \mapsto s]y &= y && \text{if } x \neq y \\ [x \mapsto s]\lambda y. t_1 &= \lambda y. [x \mapsto s]t_1 \\ [x \mapsto s](t_1 \ t_2) &= ([x \mapsto s]t_1) ([x \mapsto s]t_2) \end{aligned}$$

## Substitute is a Move in the Pokemon Games

This works reasonably well in most situations, such as the following:

$$[x \mapsto (\lambda z.z\ w)](\lambda y.x) \rightarrow \lambda y.\lambda z.z\ w \quad (20)$$

But the naive description contains a bug!

- ▶ Consider the following:

$$[x \mapsto y](\lambda x.x) \rightarrow \lambda x.y \quad (21)$$

- ▶ This happens because we pass the substitution through lambdas without checking first to see if the variable we're replacing is bound!

## Maybe I Should Have Substituted a Better Joke...

If we fix the bit where we ignore bound vs. free variables...

$$\begin{aligned} [x \mapsto s]x &= s \\ [x \mapsto s]y &= y && \text{if } y \neq x \\ [x \mapsto s](\lambda y. t_1) &= \begin{cases} \lambda y. t_1 & \text{if } y = x \\ \lambda y. [x \mapsto s]t_1 & \text{if } y \neq x \end{cases} \\ [x \mapsto s](t_1 t_2) &= ([x \mapsto s]t_1) ([x \mapsto s]t_2) \end{aligned}$$

This expression now evaluates the way we expect it to...

$$[x \mapsto y](\lambda x. x) \rightarrow \lambda x. x \quad (22)$$

But the following expression doesn't.

$$[x \mapsto z](\lambda z. x) \rightarrow \lambda z. z \quad (23)$$

- ▶ When we sub in  $z$ , it becomes bound to  $\lambda z$ .
- ▶ This is known as **variable capture**.

# Accept No Substitutes!

In order to avoid having our variables captured, we might add the condition that, in order for a substitution to pass through a  $\lambda$  abstraction, the abstracted variable must not be in the set of free variables contained within the expression we are subbing in.

$$\begin{aligned} [x \mapsto s]x &= s \\ [x \mapsto s]y &= y && \text{if } y \neq x \\ [x \mapsto s](\lambda y. t_1) &= \begin{cases} \lambda y. t_1 & \text{if } y = x \\ \lambda y. [x \mapsto s]t_1 & \text{if } y \neq x \text{ and } y \notin FV(s) \end{cases} \\ [x \mapsto s](t_1 t_2) &= ([x \mapsto s]t_1 ([x \mapsto s]t_2)) \end{aligned}$$

# No Substitutions, Extensions or Refunds!

We're not out of the woods yet, however.

- ▶ Consider the following example:

$$[x \mapsto y z](\lambda y. x y) \quad (24)$$

- ▶ No substitution can be performed, even though it would be reasonable to expect one.
- ▶ By relabelling  $y$  to some other arbitrary label, we can avoid the capture as well. For example:

$$[x \mapsto y z](\lambda y. x y) \rightarrow [x \mapsto y z](\lambda w. x w) \rightarrow (\lambda w. y z w) \quad (25)$$



# Relabel Them Variables!

By convention in  $\lambda$ -Calculus, terms that differ only in the names of bound variables are interchangeable in all contexts.

- By adding the meta-rule that we rename variables whenever a substitution would result in variable capture, we can actually simplify our rules for substitution:

$$\begin{array}{llll} [x \mapsto s]x & = & s & \\ [x \mapsto s]y & = & y & \text{if } y \neq x \\ [x \mapsto s](\lambda y. t_1) & = & \lambda y. [x \mapsto s]t_1 & \text{if } y \neq x \text{ and } y \notin FV(s) \\ [x \mapsto s](t_1 \ t_2) & = & [x \mapsto s]t_1 \ [x \mapsto s]t_2 & \end{array}$$

# Operational Semantics of $\lambda$ -Calculus

Finally, we are ready to discuss the operational semantics of the call by value evaluation strategy of  $\lambda$ -Calculus

$\rightarrow$  (untyped)

Syntax

$t ::=$

$x$

$\lambda x. t$

$t t$

$v ::=$

$\lambda x. t$

terms:

variable

abstraction

application

values:

abstraction value

Evaluation

$t \rightarrow t'$

$$\frac{t_1 \rightarrow t'_1}{t_1 t_2 \rightarrow t'_1 t_2}$$

(E-APP1)

$$\frac{t_2 \rightarrow t'_2}{v_1 t_2 \rightarrow v_1 t'_2}$$

(E-APP2)

$$(\lambda x. t_{12}) v_2 \rightarrow [x \mapsto v_2] t_{12} \quad (\text{E-APPABS})$$

Note that these are the semantics for the **pure**  $\lambda$ -Calculus.

## Things of note

- ▶ The set of values here is somewhat more interesting than in UAE.
  - ▶ Since this strategy doesn't evaluate past  $\lambda$ 's, all  $\lambda$  abstractions are values.
- ▶ In these semantics, we have one application rule (E-AppAbs), and two *congruence* rules, (E-App1) and (E-App2).
- ▶ Note how the placement of values controls the flow of execution.
  - ▶ We may only proceed with (E-App2) if  $t_1$  is a value, implying that (E-App1) is inapplicable.
  - ▶ The reason this strategy is called “call by value” is because the term being substituted in (E-AppAbs) must be a value.

## $\lambda$ -Calculus Self Interpreter

A self-interpreter is a program which implements its own semantics.

- ▶ Some programming languages, including Haskell, have their compilers and interpreters implemented in the language they are implementing.
- ▶ Python's interpreter is written in C...

The following is a self-interpreting  $\lambda$  expression, reliant on the Y-Combinator.

$$Y (\lambda e. \lambda m. m (\lambda x. x) (\lambda m. \lambda n. e \ m \ (e \ n)) (\lambda m. \lambda v. e \ (m \ v))) \quad (26)$$

I don't have an example of it's operation, I just found it while researching this slide deck, thought it was cool, and threw it in at the end.

- ▶ Mogensen, T. (1994). Efficient Self-Interpretation in Lambda Calculus. Journal of Functional Programming.

## Last Slide Comic

