

Lecture 8

Pairs

- Representing pairs:

$$\text{LET } (E_1, E_2) = \lambda f. f E_1 E_2$$
$$\text{LET } \text{fst} = \lambda p. p \text{ true}$$
$$\text{LET } \text{snd} = \lambda p. p \text{ false}$$

- (E_1, E_2) represents an ordered pair

- first component (i.e. E_1) is accessed with `fst`
- second component (i.e. E_2) is accessed with `snd`

- The definitions work, e.g.:

$$\begin{aligned} \text{fst } (E_1, E_2) &= (\lambda p. p \text{ true}) (E_1, E_2) \\ &= (E_1, E_2) \text{ true} \\ &= (\lambda f. f E_1 E_2) \text{ true} \\ &= \text{true } E_1 E_2 \\ &= (\lambda x y. x) E_1 E_2 \\ &= E_1 \end{aligned}$$

Tuples

- *n*-tuples easily defined in terms of pairs:

$$\text{LET } (E_1, E_2, \dots, E_n) = (E_1, (E_2, (\dots (E_{n-1}, E_n) \dots)))$$

- (E_1, \dots, E_n) is an *n*-tuple
 - with components E_1, \dots, E_n and length *n*
 - pairs are 2-tuples

- Extracting components of *n*-tuples:

$$\text{LET } E \downarrow^n 1 = \text{fst } E$$

$$\text{LET } E \downarrow^n 2 = \text{fst}(\text{snd } E)$$

⋮

$$\text{LET } E \downarrow^n i = \text{fst}(\underbrace{\text{snd}(\text{snd}(\dots(\text{snd } E) \dots))}_{i-1 \text{ snds}}) \quad (\text{if } i < n)$$

⋮

$$\text{LET } E \downarrow^n n = \underbrace{\text{snd}(\text{snd}(\dots(\text{snd } E) \dots))}_{n-1 \text{ snds}}$$

Verifying tuple selection works

$$\begin{aligned}(E_1, E_2, \dots, E_n) \downarrow^n 1 &= (E_1, (E_2, (\dots))) \downarrow^n 1 \\ &= \mathbf{fst} (E_1, (E_2, (\dots))) \\ &= E_1\end{aligned}$$

$$\begin{aligned}(E_1, E_2, \dots, E_n) \downarrow^n 2 &= (E_1, (E_2, (\dots))) \downarrow^n 2 \\ &= \mathbf{fst} (\mathbf{snd} (E_1, (E_2, (\dots)))) \\ &= \mathbf{fst} (E_2, (\dots)) \\ &= E_2\end{aligned}$$

- $(E_1, E_2, \dots, E_n) \downarrow^n i = E_i$ all i such that $1 \leq i \leq n$
- Usually write $E \downarrow i$ instead of $E \downarrow^n i$
 - when n clear from context
 - e.g. $(E_1, \dots, E_n) \downarrow i = E_i$ (where $1 \leq i \leq n$)

Representing numbers

- Goal: define λ -expression \underline{n} representing n
- Goal: represent arithmetical operations
 - e.g. need `suc`, `pre`, `add` and `iszero` representing:
 - the successor function ($n \mapsto n + 1$),
 - the predecessor function ($n \mapsto n - 1$),
 - addition,
 - test for zero

- Required properties:

$$\text{suc } \underline{n} = \underline{n+1} \quad (\text{for all numbers } n)$$

$$\text{pre } \underline{n} = \underline{n-1} \quad (\text{for all numbers } n > 0)$$

$$\text{add } \underline{m} \ \underline{n} = \underline{m+n} \quad (\text{for all numbers } m \text{ and } n)$$

$$\text{iszero } \underline{0} = \text{true}$$

$$\text{iszero } (\text{suc } \underline{n}) = \text{false}$$

Preliminary notation

- Define $f^n x$ to mean n applications of f to x
 - For example, $f^5 x = f(f(f(f(f x))))$
 - By convention $f^0 x$ defined to mean x

- More generally:

$$\text{LET } E^0 E' = E'$$

$$\text{LET } E^n E' = \underbrace{E(E(\dots(E E')\dots))}_{n \text{ Es}}$$

- Note that $E^n(E E') = E^{n+1} E' = E(E^n E')$
 - $f^4 x = f(f(f(f x))) = f(f^3 x) = f^3(f x)$

Church's numerals

- Representation below due to Church
- Church's numerals:

$$\text{LET } \underline{0} = \lambda f x. x$$

$$\text{LET } \underline{1} = \lambda f x. f x$$

$$\text{LET } \underline{2} = \lambda f x. f(f x)$$

⋮

$$\text{LET } \underline{n} = \lambda f x. f^n x$$

⋮

- Arithmetical operations

$$\text{LET } \text{suc} = \lambda n f x. n f(f x)$$

$$\text{LET } \text{add} = \lambda m n f x. m f (n f x)$$

$$\text{LET } \text{iszero} = \lambda n. n (\lambda x. \text{false}) \text{true}$$

Properties (exercise)

- $\text{suc } \underline{0} = \underline{1}$
- $\text{suc } \underline{5} = \underline{6}$
- $\text{iszero } \underline{0} = \text{true}$
- $\text{iszero } \underline{5} = \text{false}$
- $\text{add } \underline{0} \underline{1} = \underline{1}$
- $\text{add } \underline{2} \underline{3} = \underline{5}$
- $\text{suc } \underline{n} = \underline{n+1}$
- $\text{iszero } (\text{suc } \underline{n}) = \text{false}$
- $\text{add } \underline{0} \underline{n} = \underline{n}$
- $\text{add } \underline{m} \underline{0} = \underline{m}$
- $\text{add } \underline{m} \underline{n} = \underline{m + n}$

Predecessor function pre

- pre \underline{n} defined using $\lambda f x. f^n x$ (i.e. \underline{n})
 - goal: get a function that applies f only $n-1$ times
 - trick: ‘throw away’ the first application of f in f^n
- First define a function `prefn` on pairs:
 - `prefn f (true, x) = (false, x)`
 - `prefn f (false, x) = (false, f x)`

From this it follows that:

- $(\text{prefn } f)^n (\text{false}, x) = (\text{false}, f^n x)$
 - $(\text{prefn } f)^n (\text{true}, x) = (\text{false}, f^{n-1} x)$ (if $n > 0$)
- n applications of `prefn` to (true, x) result in $n-1$ applications of f to x
 - Definition of `prefn`

LET `prefn = λf p. (false, (fst p → snd p | (f(snd p))))`

Properties of `prefn`

- $\text{prefn } f (b, x) = (\text{false}, (b \rightarrow x \mid f x))$
- $\text{prefn } f (\text{true}, x) = (\text{false}, x)$
- $\text{prefn } f (\text{false}, x) = (\text{false}, f x)$
- $(\text{prefn } f)^n (\text{false}, x) = (\text{false}, f^n x)$
- $(\text{prefn } f)^n (\text{true}, x) = (\text{false}, f^{n-1} x)$

Definition of pre

• LET $\text{pre} = \lambda n f x. \text{snd } (n \text{ (prefn } f) \text{ (true, } x))$

• If $n > 0$

$$\begin{aligned} \text{pre } \underline{n} f x &= \text{snd } (\underline{n} \text{ (prefn } f) \text{ (true, } x)) && \text{(defn of pre)} \\ &= \text{snd } ((\text{prefn } f)^n \text{ (true, } x)) && \text{(defn of } \underline{n}) \\ &= \text{snd}(\text{false, } f^{n-1} x) && \text{(as above)} \\ &= f^{n-1} x \end{aligned}$$

• hence by extensionality

$$\begin{aligned} \text{pre } \underline{n} &= \lambda f x. f^{n-1} x \\ &= \underline{n-1} && \text{(definition of } \underline{n-1}) \end{aligned}$$

• Properties of pre

• $\text{pre } (\text{suc } \underline{n}) = \underline{n}$

• $\text{pre } \underline{0} = \underline{0}$

Another numeral systems

- Numerals with simpler predecessor function

LET $\hat{0} = \lambda x.x$

LET $\hat{1} = (\text{false}, \hat{0})$

LET $\hat{2} = (\text{false}, \hat{1})$

⋮

LET $\hat{n+1} = (\text{false}, \hat{n})$

⋮

- Can define $\widehat{\text{suc}}$, $\widehat{\text{iszero}}$, $\widehat{\text{pre}}$ such that

- $\widehat{\text{suc}} \hat{n} = \hat{n+1}$

- $\widehat{\text{iszero}} \hat{0} = \text{true}$

- $\widehat{\text{iszero}} (\widehat{\text{suc}} \hat{n}) = \text{false}$

- $\widehat{\text{pre}} (\widehat{\text{suc}} \hat{n}) = \hat{n}$

Definition by recursion

- To represent multiplication would like to define `mult` such that:

$$\text{mult } m \ n = \underbrace{\text{add } n \ (\text{add } n \ (\dots (\text{add } n \ \underline{0}) \dots))}_{m \ \text{adds}}$$

- Achieved if `mult` satisfies:

$$\text{mult } m \ n = (\text{iszero } m \rightarrow \underline{0} \mid \text{add } n \ (\text{mult } (\text{pre } m) \ n))$$

- If this held then, for example,

$$\begin{aligned} \text{mult } \underline{2} \ \underline{3} &= (\text{iszero } \underline{2} \rightarrow \underline{0} \mid \text{add } \underline{3} \ (\text{mult } (\text{pre } \underline{2}) \ \underline{3})) && \text{(by the equation)} \\ &= \text{add } \underline{3} \ (\text{mult } \underline{1} \ \underline{3}) && \\ &\quad \text{(by properties of iszero, the conditional and pre)} \\ &= \text{add } \underline{3} \ (\text{iszero } \underline{1} \rightarrow \underline{0} \mid \text{add } \underline{3} \ (\text{mult } (\text{pre } \underline{1}) \ \underline{3})) && \\ &\quad \text{(by the equation)} \\ &= \text{add } \underline{3} \ (\text{add } \underline{3} \ (\text{mult } \underline{0} \ \underline{3})) && \\ &\quad \text{(by properties of iszero, the conditional and pre)} \\ &= \text{add } \underline{3} \ (\text{add } \underline{3} \ (\text{iszero } \underline{0} \rightarrow \underline{0} \mid \text{add } \underline{3} \ (\text{mult } (\text{pre } \underline{0}) \ \underline{3}))) && \\ &\quad \text{(by the equation)} \\ &= \text{add } \underline{3} \ (\text{add } \underline{3} \ \underline{0}) && \\ &\quad \text{(by properties of iszero and the conditional)} \end{aligned}$$

Recursion

- Equation above suggests `mult` be defined by:

`mult = λm n. (iszero m → 0 | add n (mult (pre m) n))`
↑
N.B.

- This cannot be used to define `mult`
 - `mult` must *already be defined* for the λ -expression to the right of the equals to make sense
- There is a technique for constructing λ -expressions that satisfy arbitrary equations
 - applied to the equation above, this gives the desired definition of `mult`.

Fixed points

- Y is such that, for any expression E :

$$Y E = E (Y E)$$

- $Y E$ is unchanged when E is applied to it
 - if $E E' = E'$ then E' is called a *fixed point* of E
- A λ -expression Fix with the property

$$\text{Fix } E = E(\text{Fix } E)$$

(for any E) is called a *fixed-point operator*

- infinitely many different fixed-point operators
- Y is the most famous one

The fixed-point operator Y

- Definition of Y :

$$\text{LET } Y = \lambda f. (\lambda x. f(x x)) (\lambda x. f(x x))$$

- Y is a fixed-point operator:

$$\begin{aligned} Y E &= (\lambda f. (\lambda x. f(x x)) (\lambda x. f(x x))) E && \text{(defn of } Y) \\ &= (\lambda x. E(x x)) (\lambda x. E(x x)) && (\beta\text{-conversion}) \\ &= E ((\lambda x. E(x x)) (\lambda x. E(x x))) && (\beta\text{-conversion}) \\ &= E (Y E) && \text{(the line before last)} \end{aligned}$$

- Hence every E has a fixed point $Y E$

Defining mult

- Define multfn by:

$$\text{LET multfn} = \lambda f m n. (\text{iszero } m \rightarrow \underline{0} \mid \text{add } n (f (\text{pre } m) n))$$

\uparrow \uparrow

- Define mult by:

$$\text{LET mult} = \text{Y multfn}$$

- Then:

$$\begin{aligned} \text{mult } m n &= (\text{Y multfn}) m n && \text{(defn of mult)} \\ &= \text{multfn } (\text{Y multfn}) m n && \text{(property of Y)} \\ &= \text{multfn mult } m n && \text{(defn of mult)} \\ &= (\lambda f m n. (\text{iszero } m \rightarrow \underline{0} \mid \text{add } n (f (\text{pre } m) n))) \text{mult } m n && \text{(defn of multfn)} \\ &= (\text{iszero } m \rightarrow \underline{0} \mid \text{add } n (\text{mult } (\text{pre } m) n)) && \text{(\beta-conversion)} \end{aligned}$$

Recursion in general

- An equation of the form

$$f x_1 \cdots x_n = E$$

is called *recursive* if f occurs free in E

- Y provides a way of solving such equations
- start with an equation:

$$f x_1 \dots x_n = \sim f \sim$$

where $\sim f \sim$ is some λ -expression containing f

- The following defines f so this holds:

$$\text{LET } f = Y (\lambda f x_1 \dots x_n. \sim f \sim)$$

- Then:

$$\begin{aligned} f x_1 \dots x_n &= Y (\lambda f x_1 \dots x_n. \sim f \sim) x_1 \dots x_n && \text{(defn of } f) \\ &= (\lambda f x_1 \dots x_n. \sim f \sim) (Y (\lambda f x_1 \dots x_n. \sim f \sim)) x_1 \dots x_n && \text{(fixed-point property)} \\ &= (\lambda f x_1 \dots x_n. \sim f \sim) f x_1 \dots x_n && \text{(defn of } f) \\ &= \sim f \sim && (\beta\text{-conversion}) \end{aligned}$$

Functions with several arguments

- λ -expressions can only be applied to a single argument
- However, this argument can be a *tuple*
- Thus can write:

$$E(E_1, \dots, E_n)$$

which actually abbreviates:

$$E(E_1, (E_2, (\dots (E_{n-1}, E_n) \dots)))$$

- **Example:** $E(E_1, E_2)$ abbreviates $E(\lambda f. f E_1 E_2)$

Currying

- Can encode multi-argument functions as:
 - (i) $(f\ x_1\ \dots\ x_n)$, or
 - (ii) the application of f to n -tuple (x_1, \dots, x_n)
- In (i), f expects its arguments ‘one at a time’
 - said to be *curried*
 - after a logician called Curry
 - actually invented by Schönfinkel
- and, or and add are curried
- Curried functions can be ‘partially applied’
 - for example, $\text{add } \underline{1}$
 - the result of partially applying add to $\underline{1}$
 - denotes the function $n \mapsto n+1$

curry and uncurry

- Consider:

$$\text{LET curry} = \lambda f x_1 x_2. f (x_1, x_2)$$
$$\text{LET uncurry} = \lambda f p. f (\text{fst } p) (\text{snd } p)$$

- If sum and prod are defined by:

$$\text{sum} = \text{uncurry add}$$
$$\text{prod} = \text{uncurry mult}$$

- sum, prod are ‘uncurried’ versions of add, mult

- For example:

$$\begin{aligned} \text{sum } (\underline{m}, \underline{n}) &= \text{uncurry add } (\underline{m}, \underline{n}) \\ &= (\lambda f p. f (\text{fst } p) (\text{snd } p)) \text{add } (\underline{m}, \underline{n}) \\ &= \text{add } (\text{fst } (\underline{m}, \underline{n})) (\text{snd } (\underline{m}, \underline{n})) \\ &= \text{add } \underline{m} \ \underline{n} \\ &= \underline{m+n} \end{aligned}$$

curry and uncurry are inverses

- Can show:

$$\text{curry} (\text{uncurry } E) = E$$

$$\text{uncurry} (\text{curry } E) = E$$

- Hence:

$$\text{add} = \text{curry sum}$$

$$\text{mult} = \text{curry prod}$$

n -ary currying and uncurrying

- For $n > 0$ define:

$$\text{LET } \text{curry}_n = \lambda f x_1 \cdots x_n. f (x_1, \dots, x_n)$$

$$\text{LET } \text{uncurry}_n = \lambda f p. f (p \downarrow^n 1) \cdots (p \downarrow^n n)$$

- If E represents a function expecting an n -tuple argument

- then $\text{curry}_n E$ represents the curried function which takes its arguments one at a time

- If E represents a curried function of n arguments

- then $\text{uncurry}_n E$ represents the uncurried version which expects a single n -tuple as argument

- Can show:

$$\text{curry}_n (\text{uncurry}_n E) = E$$

$$\text{uncurry}_n (\text{curry}_n E) = E$$

Notation for uncurried functions

- Generalized λ -abstractions:

$$\text{LET } \lambda(V_1, \dots, V_n). E = \text{uncurry}_n (\lambda V_1 \dots V_n. E)$$

- Example: $\lambda(x, y). \text{mult } x y$ abbreviates:

$$\begin{aligned} & \text{uncurry}_2 (\lambda x y. \text{mult } x y) \\ &= (\lambda f p. f (p \downarrow^2 1) (p \downarrow^2 2)) (\lambda x y. \text{mult } x y) \\ &= (\lambda f p. f (\text{fst } p) (\text{snd } p)) (\lambda x y. \text{mult } x y) \\ &= \lambda p. \text{mult } (\text{fst } p) (\text{snd } p) \end{aligned}$$

- Thus:

$$\begin{aligned} & (\lambda(x, y). \text{mult } x y) (E_1, E_2) \\ &= (\lambda p. \text{mult } (\text{fst } p) (\text{snd } p)) (E_1, E_2) \\ &= \text{mult } (\text{fst}(E_1, E_2)) (\text{snd}(E_1, E_2)) \\ &= \text{mult } E_1 E_2 \end{aligned}$$

Generalized β -conversion

- Can derive:

$$(\lambda(V_1, \dots, V_n). E) (E_1, \dots, E_n) = E[E_1, \dots, E_n/V_1, \dots, V_n]$$

- $E[E_1, \dots, E_n/V_1, \dots, V_n]$ is the *simultaneous substitution* of E_1, \dots, E_n for V_1, \dots, V_n , respectively
 - none of these variables occur free in any of E_1, \dots, E_n
- Can be derived from ordinary β -conversion
 - and the definitions of tuples
 - and generalized λ -abstractions
 - A tuple of arguments is passed to each argument position in the body of the generalized abstraction
 - then each individual argument can be extracted from the tuple without affecting the others

More syntactic sugar for abstractions

- Convenient to extend notation $\lambda V_1 V_2 \dots V_n. E$
 - each V_i can either be an identifier
 - or a tuple of identifiers
- $\lambda V_1 V_2 \dots V_n. E$ still $\lambda V_1. (\lambda V_2. (\dots (\lambda V_n. E) \dots))$
 - if V_i is a tuple of identifiers
 - then the expression is a generalized abstraction

- **Example:**

$$\lambda f (x, y). f x y$$

means

$$\lambda f. (\lambda(x, y). f x y)$$

which means

$$\lambda f. \text{uncurry } (\lambda x y. f x y)$$

which equals

$$\lambda f. (\lambda p. f (\text{fst } p) (\text{snd } p))$$

Mutual recursion

- Consider the equations:

$$\begin{aligned}f_1 &= F_1 f_1 \cdots f_n \\f_2 &= F_2 f_1 \cdots f_n \\&\vdots \\f_n &= F_n f_1 \cdots f_n\end{aligned}$$

- Solution is:

$$f_i = Y (\lambda(f_1, \dots, f_n). (F_1 f_1 \cdots f_n, \dots, F_n f_1 \cdots f_n)) \downarrow i \\(1 \leq i \leq n)$$

- Works because if:

$$\vec{f} = Y (\lambda(f_1, \dots, f_n). (F_1 f_1 \cdots f_n, \dots, F_n f_1 \cdots f_n))$$

- Then $f_i = \vec{f} \downarrow i$ and hence:

$$\begin{aligned}\vec{f} &= (\lambda(f_1, \dots, f_n). (F_1 f_1 \cdots f_n, \dots, F_n f_1 \cdots f_n)) \vec{f} \\&= (F_1(\vec{f} \downarrow 1) \cdots (\vec{f} \downarrow n), \dots, F_n(\vec{f} \downarrow 1) \cdots (\vec{f} \downarrow n)) \\&= (F_1 f_1 \cdots f_n, \dots, F_n f_1 \cdots f_n) \quad (\text{as } \vec{f} \downarrow i = f_i)\end{aligned}$$

- Hence: $f_i = F_i f_1 \cdots f_n$

Extending the λ -calculus

- Can represent data-objects and data-structures with λ -expressions
 - often inefficient to do so
 - computers have hardware for arithmetic
 - why not use this, rather than λ -conversion
- Can ‘interface’ computation rules to λ -calculus
- Idea:
 - add a set of *new constants*
 - give rules for reducing applications involving these constants
 - such rules are called a δ -rules

Example δ -rules

- Add numerals and $+$ as new constants
- Possible δ -rule:

$$+ \ m \ n \xrightarrow{\delta} m+n$$

- $E_1 \xrightarrow{\delta} E_2$ means E_2 results by applying a δ -rule to some subexpression of E_1
- Must be careful to keep properties of λ -calculus
 - e.g. the Church-Rosser property

Safe δ -rules

- δ -rules are safe if they have the form:

$$c_1 \ c_2 \ \cdots \ c_n \xrightarrow{\delta} e$$

- where c_1, \dots, c_n are constants
- and e is either a constant or a closed abstraction (such λ -expressions are sometimes called *values*)
- **Example:** add as constants `Suc`, `Pre`, `IsZero`, Δ_0 , Δ_1 , Δ_2 , \dots with the δ -rules:

$$\text{Suc } \Delta_n \xrightarrow{\delta} \Delta_{n+1}$$

$$\text{Pre } \Delta_{n+1} \xrightarrow{\delta} \Delta_n$$

$$\text{IsZero } \Delta_0 \xrightarrow{\delta} \text{true}$$

$$\text{IsZero } \Delta_{n+1} \xrightarrow{\delta} \text{false}$$

- Δ_n represents the number n ,
- `Suc`, `Pre`, `IsZero` are new constants (i.e. not defined λ -expressions like `suc`, `pre`, `iszero`)
- `true` and `false` defined above (both are values)

