Foundations of functional programming
Matthew Parkinson
12 Lectures (Lent 2008)

Overview

Motivation

Understanding:
• simple notion of computation

Encoding:
• Representing complex features in terms of simpler features

Functional programming in the wild:
• Visual Basic and C# have functional programming features.
**Pure λ-calculus**

\[ M ::= x \mid (M M) \mid (\lambda x. M) \]

**Syntax:**
- \( x \) variable
- \( (M M) \) (function) application
- \( (\lambda x. M) \) (lambda) abstraction

World smallest programming language:
- \( \alpha, \beta, \eta \) reductions
- when are two programs equal?
- choice of evaluation strategies

**Pure λ-calculus is universal**

Can encode:
- Booleans
- Integers
- Pairs
- Disjoint sums
- Lists
- Recursion within the λ-calculus.

Can simulate a Turing or Register machine (Computation Theory), so is universal.

---

**Applied λ-calculus**

\[ M ::= x \mid \lambda x. M \mid M M \mid c \]

**Syntax:**
- \( x \) variables
- \( \lambda x. M \) (lambda) abstraction
- \( M M \) (function) application
- \( c \) (constants)

Elements of \( c \) used to represent integers, and also functions such as addition
- \( \delta \) reductions are added to deal with constants

---

**Combinators**

\[ M ::= M M \mid c \quad \text{(omit } x \text{ and } \lambda x. M) \]

We just have \( c \in \{S, K\} \) regains power of λ-calculus.

Translation to/from lambda calculus including almost equivalent reduction rules.
Evaluation mechanisms/facts

Eager evaluation (Call-by-value)
Lazy evaluation (Call-by-need)
Confluence “There’s always a meeting place downstream”
Implementation Techniques

Real implementations

• “Functional Languages”
• Don’t do substitution, use environments instead.
• Haskell, ML, F#, Visual Basic, C#

SECD

Abstract machine for executing the \(\lambda\)-calculus.
4 registers Stack, Environment, Control and Dump.

Continuations

• \(\lambda\)-expressions restricted to always return “()” [continuations] can implement all \(\lambda\)-expressions
• Continuations can also represent many forms of non-standard control flow, including exceptions
• call/cc
State

How can we use state and effects in a purely functional language?

Types

This course is primarily untyped.
We will mention types only where it aids understanding.

Pure λ-calculus

Syntax

Variables: x,y,z,...
Terms: M,N,L,... := λx.M | M N | x
We write M=N to say M and N are syntactically equal.
Free variables and permutation

We define free variables of a $\lambda$-term as

- $\text{FV}(M N) = \text{FV}(M) \cup \text{FV}(N)$
- $\text{FV}(\lambda x. M) = \text{FV}(M) \setminus \{x\}$
- $\text{FV}(x) = \{x\}$

We define variable permutation as

- $x < x' z > = z$
- $x < y \cdot z > = x$ (provided $x \neq y$ and $x \neq z$)
- $(\lambda x. M) < y \cdot z > = \lambda (x < y \cdot z >). (M < y \cdot z >)$
- $(M N) < y \cdot z > = (M < y \cdot z >) (N < y \cdot z >)$

Recap: Equivalence relations

An equivalence relation is a reflexive, symmetric and transitive relation.

$R$ is an equivalence relation if

- Reflexive
  $\forall x. \ x R x$
- Transitive
  $\forall xyz. \ x R y \land y R z \Rightarrow x R z$
- Symmetric
  $\forall xy. \ x R y \Rightarrow y R x$

Congruence and Contexts

A congruence relation is an equivalence relation, that is preserved by placing terms under contexts.

Context (term with a single hole (*)):

$C ::= \lambda x. C \ | \ C M \ | \ M C \ | \*$

Context application $C[M]$ fills hole (*) with $M$.

$R$ is a compatible relation if

- $\forall M N C. \ M R N \Rightarrow C[M] R C[N]$

$R$ is a congruence relation if it is both an equivalence and a compatible relation.

$\alpha$-equivalence

Two terms are $\alpha$-equivalent if they can be made syntactically equal (\(=\)) by renaming bound variables.

$\alpha$-equivalence (\(=_{\alpha}\)) is the least congruence relation satisfying

- $\lambda x. M =_{\alpha} \lambda y. M<x \cdot y>$ where $y \notin \text{FV}(\lambda x. M)$
Intuition of $\alpha$-equivalence

Consider $\lambda x. \lambda y. x \ y \ z \ x$.
We can see this as $\lambda x. \lambda y. x \ y \ z \ x$ and hence the bound names are irrelevant.

We only treat terms up to $\alpha$-equivalence.

$\alpha$-equivalence (alternative defn)

Use $\lambda x. s. M$ as a shorthand, where
- $x s ::= x s, x \ | \ []$
- $\lambda [] M = M$
- $\lambda x s. x. M = \lambda x s. \lambda x. M$

Definition
- $\lambda [] x =_\alpha \lambda [] x$
- $\lambda x s. x_1. x_2 =_\alpha \lambda y s. y_1. y_2$
  if $(x_1 = x_2$ and $y_1 = y_2$)
  or $(x_1 \neq x_2$ and $y_1 = y_2$ and $\lambda x s. x_1 =_\alpha \lambda y s. y_1$)
- $\lambda x s. M_1. M_2 =_\alpha \lambda y s. N_1. N_2$
  iff $\lambda x s. M_1 =_\alpha \lambda y s. N_1$ and $\lambda x s. M_2 =_\alpha \lambda y s. N_2$

Capture avoiding substitution

If $x \notin FV(M)$,
- $M [L/x] = M$
otherwise:
- $(M \ N) [L/x] = (M [L/x] \ N [L/x])$
- $(\lambda y. M) [L/x] = (\lambda z. M < z \cdot y > [L/x])$
  where $z \notin FV(x, L, \lambda y. m)$
- $x [L/x] = L$

Note: In the $(\lambda y. M)$ case, we use a permutation to pick an $\alpha$-equivalent term that does not capture variables in $L$. 

Are these alpha-equivalent?

$\lambda x. x =_\alpha \lambda y. y$
$\lambda x. y. x =_\alpha \lambda y. \lambda x. y$
$\lambda x. y =_\alpha \lambda y. y$
$(\lambda x. x) (\lambda y. y) =_\alpha (\lambda y. y) (\lambda x. x)$
$\lambda x. \lambda y. (x \ z \ y) =_\alpha \lambda z. \lambda y. (z \ z \ y)$
\[ (x \, y)[L/y] = x \, L \]
\[ (\lambda \, x. \, y) [x/w] = \lambda \, x. \, y \]
\[ (\lambda \, x. \, (x \, y)) [L/x] = (\lambda \, x. \, (x \, y)) \]
\[ (\lambda \, x. \, y) [x/y] = (\lambda \, z. \, x) \]
\[ (\lambda \, y. \, (\lambda \, x. \, z)) [x \, w/ \, z] = (\lambda \, y. \, (\lambda \, x. \, (x \, w))) \]

**Extra brackets**

To simplify terms we will drop some brackets:

\[ \lambda \, y. \, M = \lambda \, x. \, (\lambda \, y. \, M) \]
\[ L \, M \, N = (L \, M) \, N \]
\[ \lambda \, x. \, M \, N = \lambda \, x. \, (M \, N) \]

Some examples

\[ (\lambda \, x. \, x \, x) \, (\lambda \, x. \, x \, x) \, y \, z = (((\lambda \, x. \, (x \, x)) \, (\lambda \, x. \, (x \, x))) \, y) \, z \]
\[ \lambda \, x \, y \, z. \, \lambda \, x. \, (\lambda \, y. \, (\lambda \, z. \, ((x \, y) \, z))) \]

**\(\beta\eta\)-reduction**

We define \(\beta\)-reduction as:

\[ (\lambda \, x. \, M) \, N \rightarrow_\beta \, M \, [N/x] \]

This is the workhorse of the \(\lambda\)-calculus.

We define \(\eta\)-reduction as: If \(x \notin \text{FV}(M)\), then

\[ \lambda \, x. \, (M \, x) \rightarrow_\eta \, M \]

This collapses trivial functions.

Consider \((\text{fn} \, x \mapsto \sin \, x)\) is this the same as \(\sin\) in ML?

**\(\beta\eta\) examples**

\[ (\lambda \, x. \, x \, y) \, (\lambda \, z. \, z) \rightarrow_\beta \, \lambda \, z \, z \, y \]
\[ (\lambda \, x. \, y) \, (\lambda \, z. \, z) \rightarrow_\beta \, (\lambda \, z. \, z) \, y \]
\[ \lambda \, x. \, M \, N \, x \rightarrow_\eta \, (M \, N) \]
\[ (\lambda \, x. \, x \, x) \, (\lambda \, x. \, x \, x) \rightarrow_\beta \, (\lambda \, x. \, x \, x) \, (\lambda \, x. \, x \, x) \]
\[ (\lambda \, y \, x. \, x) \, (\lambda \, y. \, y) \rightarrow_\beta \, (\lambda \, y \, x. \, y) \]
Reduction in a context

We actually define $\beta$-reduction as:
$$C[ (\lambda x.M) N ] \to_{\beta} C[ M[N/x] ]$$
and $\eta$-reduction as:
$$C[ (\lambda x.(M x)) ] \to_{\eta} C[ M ]$$ (where $x \notin \text{FV}(M))$
where $C ::= \lambda x.C \mid C M \mid M C \mid \ast$ (from "Context and Congruence" slide)

Note: to control evaluation order we can consider different contexts.

How many reductions?

$$(\lambda x.x) ((\lambda x.x)(\lambda x.x))$$

$$(\lambda x.x) (\lambda x.x)(\lambda x.x)$$

$$((\lambda x y.x) z) w$$

$$(\lambda x y.z) ((\lambda x.x)(\lambda x.x))(\lambda x.x)$$

$$\lambda x y. z$$

Normal-form (NF)

A term is in normal form if it there are no $\beta$ or $\eta$ reductions that apply.

Examples in NF:
- $x; \lambda x.y$; and $\lambda x y. x (\lambda x.y)$

and not in NF:
- $(\lambda x.x) y; (\lambda x. x x)(\lambda x. x x);$ and $(\lambda x. y x)$

$\beta$-normal-form:
- $NF ::= \lambda x. NF \mid NF_2$
- $NF_2 ::= NF_2 NF \mid x$
Normal-forms

A term has a normal form, if it can be reduced to a normal form:
- \((\lambda x.x) \, y\) has a normal form \(y\)
- \((\lambda x. y \, x)\) has a normal form \(y\)
- \((\lambda x. x \, x) \,(\lambda x. x \, x)\) does not have a normal form

Note: \((\lambda x.x) \,(\lambda x.x)\) is sometimes denoted \(\Omega\).

Note: Some terms have normal forms and infinite reduction sequences, e.g. \((\lambda x. y) \, \Omega\).

Weak head normal form

A term is in WHNF if it cannot reduce when we restrict the context to
\[ C ::= C\, M | M\, C | \ast \]
That is, we don’t reduce under a \(\lambda\).
\(\lambda x. \Omega\) is a WHNF, but not a NF.

Multi-step reduction

\[ M \rightarrow^* N \quad \text{iff} \]
- \(M \rightarrow_\beta N\)
- \(M \rightarrow_\eta N\)
- \(M = N\) (reflexive)
- \(\exists L. M \rightarrow^* L \text{ and } L \rightarrow^* N\) (transitive)

The transitive and reflexive closure of \(\beta\) and \(\eta\) reduction.

Equality

We define equality on terms, \(=\), as the least congruence relation, that additional contains
- \(\alpha\)-equivalence (implicitly)
- \(\beta\)-reduction
- \(\eta\)-reduction

Sometimes expressed as \(M = M'\) iff there exists a sequence of forwards and backwards reductions from \(M\) to \(M'\):
- \(M \rightarrow^* N_1 \leftarrow \rightarrow^* M_1 \leftarrow \rightarrow^* N_2 \leftarrow \ldots \rightarrow^* N_k \leftarrow \rightarrow^* M'\)

Exercise: Show these are equivalent.
Equality properties

If \( M \rightarrow^* N \) or \( N \rightarrow^* M \), then \( M = N \).
The converse is not true (Exercise: why?)

If \( L \rightarrow^* M \) and \( L \rightarrow^* N \), then \( M = N \).
If \( M \rightarrow^* L \) and \( N \rightarrow^* L \), then \( M = N \).

Church-Rosser Theorem

Theorem: If \( M = N \), then there exists \( L \) such that \( M \rightarrow^* L \) and \( N \rightarrow^* L \).

Consider \( (\lambda x. ax)((\lambda y. by)c) \):
- \( (\lambda x. ax)((\lambda y. by)c) \rightarrow^\beta a((\lambda y. by)c) \rightarrow^\beta a(bc) \)
- \( (\lambda x. ax)((\lambda y. by)c) \rightarrow^\beta (\lambda x. ax)(bc) \rightarrow^\beta a(bc) \)
Note: Underlined term is reduced.

Consequences

If \( M = N \) and \( N \) is in normal form, then \( M \rightarrow^* N \).
If \( M = N \) and \( M \) and \( N \) are in normal forms, then \( M \equiv N \).
Conversely, if \( M \) and \( N \) are in normal forms and are distinct, then \( M \not\equiv N \). For example, \( \lambda x y.x \not\equiv \lambda x y.y \).

Diamond property

Key to proving Church–Rosser Theorem is demonstrating the diamond property:
- If \( M \rightarrow^* N_1 \) and \( M \rightarrow^* N_2 \), then there exists \( L \) such that \( N_1 \rightarrow^* L \) and \( N_2 \rightarrow^* L \).

Exercise: Show how this property implies the Church–Rosser Theorem.
The diamond property does not hold for the single step reduction:
• If $M \to^\beta N_1$ and $M \to^\beta N_2$, then there exists $L$ such that $N_1 \to^\beta L$ and $N_2 \to^\beta L$.

Consider $(\lambda x.xx) (l\ a)$ where $l = \lambda x.x$. This has two initial reductions:
• $(\lambda x.xx) (l\ a) \to^\beta (\lambda x.xx) a \to^\beta a a$
• $(\lambda x.xx) (l\ a) \to^\beta (l\ a) (l\ a)$

Now, the second has two possible reduction sequences:
• $(l\ a) (l\ a) \to^\beta a (l\ a) \to^\beta a a$
• $(l\ a) (l\ a) \to^\beta (l\ a) a \to^\beta a a$

Following first path, we have reached normal-form, while second is potentially infinite.
Normal order reduction

Perform leftmost, outermost $\beta$-reduction.
(leave $\eta$-reduction until the end)

Reduction context
- $C ::= \lambda x. C \mid C'$
- $C' ::= C' M \mid NF_2 C \mid \cdot$

where $NF_2$ is from $\beta$-normal-form definition.
This definition is guaranteed to reach normal-form if one exists.

Other reduction orders

Call-by-name:
- $C ::= C M \mid \cdot$

Call-by-value:
- $V ::= x \mid \lambda x. M$ (values)
- $C ::= C M \mid \cdot \mid (\lambda x.M) C$
- $C[ (\lambda x.M) V ] \rightarrow C[ M[V/x] ]$

Motivation

We want to use different datatypes in the $\lambda$-calculus.

Two possibilities:
- Add new datatypes to the language
- Encode datatypes into the language

Encoding makes program language simpler, but less efficient.

Encoding Data

Typo in printout

Typo in printout

Typo in printout
Encoding booleans

To encode booleans we require IF, TRUE, and FALSE such that:

\[
\begin{align*}
\text{IF TRUE } M N &= M \\
\text{IF FALSE } M N &= N
\end{align*}
\]

Here, we are using \( = \) as defined earlier.

Encoding booleans

Definitions:

- TRUE = \( \lambda m \ n\). m
- FALSE = \( \lambda m \ n\). n
- IF = \( \lambda b \ m \ n\). b m n

TRUE and FALSE are both in normal–form, so by Church–Rosser, we know TRUE ≠ FALSE.

Note that, IF is not strictly necessary as

- \( \forall P. \ IF \ P = P \) (Exercise: show this).

Logical operators

We can give AND, OR and NOT operators as well:

- AND = \( \lambda x y. \ IF \ x \ y \) FALSE
- OR = \( \lambda x y. \ IF \ x \ TRUE \ y \)
- NOT = \( \lambda x. \ IF \ x \ FALSE \ TRUE \)

Exercise: Show

- If \( L = \) TRUE then IF \( L \ M N = M \).
- If \( L = \) FALSE then IF \( L \ M N = N \).
Encoding pairs

Constructor:
• PAIR = \lambda xyf. fxy

Destructors:
• FST = \lambda p.p TRUE
• SND = \lambda p.p FALSE

Properties: \forall pq.
• FST (PAIR p q) = p
• SND (PAIR p q) = q

Encoding sums

Constructors:
• INL = \lambda x. PAIR TRUE x
• INR = \lambda x. PAIR FALSE x

Destructor:
• CASE = \lambda s f g. IF (FST s) (f (SND s)) (g (SND s))

Properties:
• CASE (INL x) f g = f x
• CASE (INR x) f g = g x

Encoding sums (alternative defn)

Constructors:
• INL = \lambda x f g. f x
• INR = \lambda x f g. g x

Destructors:
• CASE = \lambda s f g. s f g

As with booleans destructor unnecessary.
• \forall p. CASE p = p

Church Numerals

Define:
• 0 = \lambda f x . x
• 1 = \lambda f x. f x
• 2 = \lambda f x. f (f x)
• 3 = \lambda f x. f (f (f x))
• ...
• n = \lambda f x. f(...(f x)...)  

That is, n takes a function and applies it n times to its argument: n f is f^n.
Arithmetic

Definitions
• \( \text{ADD} = \lambda mnfx. \, m \, (n \, f \, x) \)
• \( \text{MULT} = \lambda mnfx. \, m \, (n \, f) \, x = \lambda m n f. \, (m \, n \, f) \)
• \( \text{EXP} = \lambda m n f x. \, m \, n \, m \, f \, x \)

Example:
\( \text{ADD} \, m \, n \rightarrow^* \lambda f x. \, m \, f \, (n \, f \, x) \rightarrow^* \, f^m(n \, x) = f^{m+n} \, x \)

More arithmetic

Definitions
• \( \text{SUC} = \lambda n x. f (n \, f \, x) \)
• \( \text{ISZERO} = \lambda n. \, (\lambda x. \, \text{FALSE}) \, \text{TRUE} \)

Properties
• \( \text{SUC} \, 0 = n+1 \)
• \( \text{ISZERO} \, 0 = \text{TRUE} \)
• \( \text{ISZERO} \, (n+1) = \text{FALSE} \)

We also require decrement/predecessor!

Building decrement

<table>
<thead>
<tr>
<th>n</th>
<th>PFN(n)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(0,0)</td>
</tr>
<tr>
<td>1</td>
<td>suc</td>
</tr>
<tr>
<td>2</td>
<td>(1,0)</td>
</tr>
<tr>
<td>3</td>
<td>suc</td>
</tr>
<tr>
<td>4</td>
<td>(2,1)</td>
</tr>
<tr>
<td>5</td>
<td>suc</td>
</tr>
<tr>
<td>6</td>
<td>(3,2)</td>
</tr>
<tr>
<td>7</td>
<td>suc</td>
</tr>
</tbody>
</table>

Decrement and subtraction

Definitions:
• \( \text{PFN} = \lambda n. \, n \, (\lambda p. \, (\text{SUC}(\text{FST} \, p), \, \text{FST} \, p)) \, (\text{PAIR} \, 0 \, 0) \)
• \( \text{PRE} = \lambda n. \, \text{SND} \, (\text{PFN} \, n) \)
• \( \text{SUB} = \lambda m n. \, \text{n} \, \text{PRE} \, m \)

Exercise: Evaluate
• \( \text{PFN} \, 5 \)
• \( \text{PRE} \, 0 \)
• \( \text{SUB} \, 4 \, 6 \)
Lists

Constructors:
- NIL = PAIR TRUE (λ.z.z)
- CONS = λxy. PAIR FALSE (PAIR x y)

Destructor:
- NULL = FST
- HD = λl. FST (SND l)
- TL = λl. SND (SND l)

Properties:
- NULL NIL = TRUE
- HD (CONS M N) = M

Recursion

How do we actually iterate over a list?

Fixed point combinator (Y)

We use a fixed point combinator Y to allow recursion.
In ML, we write:
letrec f(x) = M in N
this is really
let f = Y (λf.λx. M) in N
and hence
(λf.N) (Y λf. λx. M)
Defining recursive function

Consider defining a factorial function with the following property:
\[ \text{FACT} = \lambda n. (\text{ISZERO } n) \perp (\text{MULT } n \ (\text{FACT} \ (\text{PRE } n))) \]

We can define
\[ \text{PREFACT} = \lambda f n. (\text{ISZERO } n) \perp (\text{MULT } n \ (f \ (\text{PRE } n))) \]

Properties
- Base case: \( \forall F. \text{PREFACT } F \ 0 = 1 \)
- Inductive case: \( \forall F. \) If \( F \) behaves like factorial up to \( n \), then \( \text{PREFACT } F \) behaves like factorial up to \( n+1 \);

General approach

If you need a term, \( M \), such that
- \( M = P \ M \)
Then \( M = Y P \) suffices

Example:
- \( \text{ZEROES} = \text{CONS } \ 0 \ \text{ZEROES} = (\lambda p. \text{CONS } \ 0 \ p) \ \text{ZEROES} \)
- \( \text{ZEROES} = Y (\lambda p. \text{CONS } \ 0 \ p) \)

Fixed points

Discrete Maths: \( x \) is a fixed point of \( f \), iff \( f \ x = x \)

Assume, \( Y \) exists (we will define it shortly) such that
- \( Y \ f = f \ (Y \ f) \)

Hence, by using \( Y \) we can satisfy this property:
\[ \text{FACT} = Y (\text{PREFACT}) \]

Exercise: Show \( \text{FACT} \) satisfies property on previous slide.

Mutual Recursion

Consider trying to find solutions \( M \) and \( N \) to:
- \( M = P \ M \ N \)
- \( N = Q \ M \ N \)

We can do this using pairs:
\[ L = Y (\lambda p. \text{PAIR } (P \ (\text{FST } p)) (\text{SND } p)) \]
\[ M = \text{FST } L \]
\[ N = \text{SND } L \]

Exercise: Show this satisfies equations given above.
Y

Y has no normal form

We assume:
• M has no normal form, iff M x has no normal form. (Exercise: prove this)

Proof of Y has no normal form:
• Y f = f (Y f) (by Y property)
• Assume Y f has a normal form N.
• Hence f (Y f) can reduce to f N, and f N is also a normal form.
• Therefore, by Church Rosser, f N = N, which is a contradiction, so Y f cannot have a normal form.
• Therefore, Y has no normal form.

Head normal form

How can we characterise well-behaved λ-terms?
• Terms with normal forms? (Too strong, FACT does not have normal form)
• Terms with weak head normal form (WHNF)? (Too weak, lots of bad terms have this, for example λ.x.Ω).
• New concept: Head normal form.

HNF

A term is in head normal form, iff it looks like
\[ \lambda x_1,...,x_m. y \ M_1 ... M_k \ (m,k \geq 0) \]

Examples:
• x , \ λxy.x , \ λz.(\λx.a)z ,
\[ \lambda f (\lambda x. f(xx)) (\lambda x. f(xx)) \]

Non-examples:
• \ λy.(\λx.a) y \ → \ λy.a
• \ λf. (\λx. f(xx)) (\λx. f(xx))
Properties

Head normal form can be reached by performing head reduction (leftmost)
- $C' ::= C' \ M$ |
- $C ::= \lambda x.C \mid C'$

Therefore, $\Omega$ has no HNF (Exercise: prove this.)

If $M \ N$ has a HNF, then so does $M$. Therefore, if $M$ has no HNF, then $M \ N_1 \ldots \ N_k$ does not have a HNF. Hence, $M$ is a “totally undefined function”.

ISWIM: Syntax

From the $\lambda$-calculus
- $x$ (variable)
- $\lambda x.M$ (abstraction)
- $M \ N$ (application)

Local declarations
- let $x = M$ in $N$ (simple declaration)
- let $f \ x_1 \ldots \ x_n = M$ in $N$ (function declaration)
- letrec $f \ x_1 \ldots \ x_n = M$ in $N$ (recursive declaration)

and post-hoc declarations
- $N$ where $x = M$
- ...

ISWIM

$\lambda$-calculus as a programming language
(The next 700 programming languages [Landin 1966])

ISWIM: Syntactic sugar

$N$ where $x = M$ = let $x = M$ in $N$
let $x = M$ in $N$ = $(\lambda x.N) \ M$
let $f \ x_1 \ldots \ x_n = M$ in $N$ = let $f = \lambda x_1 \ldots x_n.M$ in $N$
letrec $f \ x_1 \ldots \ x_n = M$ in $N$ = let $f = Y(\lambda f.\lambda x_1 \ldots x_n.M)$ in $N$

Desugaring explains syntax purely in terms of $\lambda$-calculus.
ISWIM: Constants

\[ M ::= x \mid c \mid \lambda x. M \mid M N \]

Constants c include:
• 0 1 2 −1 −2 ... (integers)
• + − x / (arithmetic operators)
• = ≠ < > (relational operators)
• true false (booleans)
• and or not (boolean connectives)

Reduction rules for constants: e.g.
• + 0 0 \rightarrow 0

Call-by-value and IF-THEN-ELSE

ISWIM uses the call-by-value \(\lambda\)-calculus.

Consider: \(\text{IF TRUE } 1\ \Omega\)

\[ \text{IF E THEN M ELSE N} = (\text{IF E } (\lambda x. M) (\lambda x. N)) (\lambda z. z) \]
where \(x \notin FV(M N)\)

Pattern matching

Has
• \((M, N)\) (pair constructor)
• \(\lambda(p_1, p_2). M\) (pattern matching pairs)

Desugaring
• \(\lambda(p_1, p_2). M = \lambda z.(\lambda(p_1, p_2). M)(\text{fst } z)(\text{snd } z)\)
where \(z \notin FV(M)\)

Real \(\lambda\)-evaluator

Don’t use \(\beta\) and substitution
Do use environment of values, and delayed substitution.
Environments and Closures

Consider β-reduction sequence
\[(λxy.x+y)\ 3\ 5\ →\ (λy.3+y)\ 5\ →\ 3\ +\ 5\ →\ 8.\]
Rather than produce \((λy.3+y)\) build a closure:
\[
\text{Clo}(\ y,\ x+y,\ x=3)
\]
The arguments are
- bound variable;
- function body; and
- environment.

SECD Machine

Virtual machine for ISWIM.
The SECD machine has a state consisting of four components S, E, C and D:
- S: The “stack” is a list of values typically operands or function arguments; it also returns result of a function call;
- E: The “environment” has the form \(x_1=a_1;...;x_n=a_n\), expressing that the variables \(x_1,..,x_n\) have values \(a_1,...,a_n\) respectively; and
- C: The “control” is a list of commands, that is λ-terms or special tokens/instructions.

SECD Machine

- D: The “dump” is either empty (−) or is another machine state of the form \((S,E,C,D)\). A typical state looks like
  \((S_1,E_1,C_1,(S_2,E_2,C_2,...,(S_n,E_n,C_n,−)...)\)
  It is essentially a list of triples \((S_1,E_1,C_1),...,,(S_n,E_n,C_n)\)
  and serves as the function call stack.

State transitions: constant
Inefficient as requires construction of closures. Perform some conversions in advance:

- \([ c ] = \text{const } c\)
- \([ x ] = \text{var } x\)
- \([ M \ N ] = [ N ] ; [ M ] ; \text{app}\)
- \([ \lambda x.M ] = \text{Closure}(x,[M])\)
- \([ M + N ] = [M] ; [N] ; \text{add}\)
- ...

More intelligent compilations for “let” and tail recursive functions can also be constructed.
Example

We can see ($(\lambda xy.x+y) \; 3 \; 5$) compiles to
- \texttt{const 5; const 3; Closure(x,C_0); app; app}
where
- $C_0 = \text{Closure}(y,C_1)$
- $C_1 = \text{var } x; \text{ var } y; \text{ add}$

Recursion

The usual fixpoint combinator fails under the SECD machine: it loops forever.

A modified one can be used:
- $\lambda fx. f(\lambda y. x \; x \; y) \; (\lambda y. x \; x \; y)$
This is very inefficient.

Better approach to have closure with pointer to itself.

Recursive functions ($Y(\lambda fx.M)$)

\[
\begin{array}{|c|c|c|c|}
\hline
\text{Clo}(x,M,f=:\text{Clo}(x,M,f=:\text{Clo}(x,M,f=\cdot;E');a;S)} & - & x=a;f=\text{Clo}(x,M,f=:\text{Clo}(x,M,f=:\cdot;E');E'c & E \\
E & C & D & M \\
\hline
\end{array}
\]

Implementation in ML

SECD machine is a small–step machine.

Next we will see a big–step evaluator written in ML.
Implementation in ML

datatype Expr = Name of string |
| Numb of int |
| Plus of Expr * Expr |
| Fn of string * Expr |
| Apply of Expr * Expr

datatype Val =
| IntVal of val |
| FnVal of string * Expr * Env |

and Env = Empty | Defn of string * Val * Env

Implementation in ML

fun lookup (n, Defn (s,v,r)) =
  if s=n then v else lookup(n, r)
| lookup(n, Empty) = raise oddity()

fun eval (Name(s), r) = lookup(s, r)
| eval(Fn(bv,body),r) = FnVal(bv,body,r)
| eval(Apply(e,e'), r) =
  case eval(e, r) of
  | IntVal(i) => raise oddity()
  | FnVal(bv,body,env) =>
    let val arg = eval(e', r) in
    eval(body, Defn(bv, arg, env)
...

Exercises

How could we make it lazy?
Combinator logic

Syntax:

\[ P, Q, R ::= S \mid K \mid P \cdot Q \]

Reductions:

\[ K P Q \rightarrow_w P \]
\[ S P Q R \rightarrow_w (P R) (Q R) \]

Note that the term \( S K \) does not reduce: it requires three arguments. Combinator reductions are called “weak reductions”.

Identity combinator

Consider the reduction of, for any \( P \):

- \( S K K P \rightarrow_w K P (K P) \rightarrow_w P \)

Hence, we define \( I = S K K \), where \( I \) stands for identity.

Church-Rosser

Combinators also satisfy Church–Rosser:

- if \( P = Q \), then exists \( R \) such that \( P \rightarrow_w^* R \) and \( Q \rightarrow_w^* R \)
Encoding the λ-calculus

Use extended syntax with variables:

- \( P ::= S \mid K \mid P P \mid x \)

Define meta-operator on combinators \( \lambda^* \) by

- \( \lambda^*x.x = I \)
- \( \lambda^*x.P = K P \) (where \( x \notin \text{FV}(P) \))
- \( \lambda^*x.P Q = S(\lambda^*x.P)(\lambda^*x.Q) \)

There and back again

λ-calculus to SK:

- \( (\lambda x. M)_{CL} = (\lambda^* x. (M)_{CL}) \)
- \( (x)_{CL} = x \)
- \( (M N)_{CL} = (M)_{CL} (N)_{CL} \)

SK to λ-calculus:

- \( (x)_\lambda = x \)
- \( (K)_\lambda = \lambda xy.x \)
- \( (S)_\lambda = \lambda xyz. x z (y z) \)
- \( (P Q)_\lambda = (P)_\lambda (Q)_\lambda \)

Example translation

\[
(\lambda^* x. \lambda^* y. y \ x)
\]

- \( = \lambda^*x. S(\lambda^*y.y)(\lambda^*y.x) \)
- \( = \lambda^*x. (S\ I)\ (K\ x) \)
- \( = S(\lambda^*x.(S\ I))\ (\lambda^*x.K\ x) \)
- \( = S( (K\ I))\ (S\ (\lambda^*x.K)\ (\lambda^*x.x)) \)
- \( = S( (K\ I))\ (S\ (K\ K)\ I) \)

Properties

Free variables are preserved by translation

- \( \text{FV}(M) = \text{FV}( (M)_{CL} ) \)
- \( \text{FV}(P) = \text{FV}( (P)_\lambda ) \)

Supports α and β reduction:

- \( (\lambda^* x.P)\ Q \rightarrow^\alpha (P)\ [Q/x] \)
- \( (\lambda^* x.P) = \lambda y. P\ y \cdot x \) (where \( y \notin \text{FV}(P) \))
Equality on combinators

Combinators don’t have an analogue of the η-reduction rule.

• \((SK)_\alpha = (KI)_\alpha\), but SK and KI are both normal forms

To define equality on combinators, we take the least congruence relation satisfying:

• weak reductions, and
• functional extensionality: If \(P \ x = Q \ x\), then \(P = Q\) (where \(x \notin FV(P \ Q)\)).

\[ SK \ x \ y \rightarrow (K \ y) (K \ x) \rightarrow y \leftarrow I \ y \leftarrow K \ I \ x \ y \]

Therefore, \(SK = KI\).

Aside: Hilbert style proof

In Logic and Proof you covered Hilbert style proof:

• Axiom K: \(\forall A. \ A \rightarrow (B \rightarrow A)\)
• Axiom S: \(\forall ABC. (A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))\)
• Modus Ponens: If \(A \rightarrow B\) and \(A\), then \(B\)

Hilbert style proofs correspond to “Typed” combinator terms:

• \(S \ K : \forall AB. ((A \rightarrow B) \rightarrow (A \rightarrow A))\)
• \(S \ K \ K : \forall A. (A \rightarrow A)\)

Logic, Combinators and the \(\lambda\)-calculus are carefully intertwined. See Types course for more details.

Properties

We get the following properties of the translation:

• \(((M)_{CL})_\alpha = M\)
• \(((P)_\alpha)_{CL} = P\)
• \(M = N \leftrightarrow (M)_{CL} = (N)_{CL}\)
• \(P = Q \leftrightarrow (P)_\alpha = (Q)_\alpha\)

Compiling with combinators

The translation given so far is exponential in the number of lambda abstractions.

Add two new combinators

• \(B \ P \ Q \ R \rightarrow_w P \ (Q \ R)\)
• \(C \ P \ Q \ R \rightarrow_w P \ R \ Q\)

Exercise: Encode B and C into just S and K.
Advanced translation

- $\lambda^T x. x = I$
- $\lambda^T x. P = KP \ (x \not\in \text{FV}(P))$
- $\lambda^T x. P x = P \ (x \not\in \text{FV}(P))$
- $\lambda^T x. PQ = B P(\lambda^T x. Q) \ (x \not\in \text{FV}(P) \text{ and } x \in \text{FV}(Q))$
- $\lambda^T x. PQ = C(\lambda^T x. P) Q \ (x \in \text{FV}(P) \text{ and } x \not\in \text{FV}(Q))$
- $\lambda^T x. PQ = S(\lambda^T x. P)(\lambda^T x. Q) \ (x \in \text{FV}(P), x \in \text{FV}(Q))$

(Invented by David Turner)

Example

$$(\lambda^T x. \lambda^T y. y \ x)$$
$$= (\lambda^T x. C (\lambda^T y. y) \ x)$$
$$= (\lambda^T x. C I \ x)$$
$$= C I$$

Compared to $$(\lambda^* x. \lambda^* y. y \ x) = S (K (S I)) (S (K K) I)$$

Translation with $\lambda^*$ is exponential, while $\lambda^T$ is only quadratic.

Combinators as graphs

To enable lazy reduction, consider combinator terms as graphs.

$S$ reduction creates two pointers to the same subterm.

Let’s consider

- let $sqr$ $x$ $=$ $\text{mult}$ $x$ $x$ in $sqr$ $5$ $=$ $(\lambda f \ S f) (\lambda m. \text{mult} \ m \ m)$

this translates to

- $C I S \ (\lambda m. \text{mult} \ m)$

Exercise: Show this translation.
C I 5 (S mult I)

Reduction: I

Reduction: K

Reduction: S
Reduction: B

Recursion

Reduction: C

Comments

If 5 was actually a more complex calculation, would only have to perform it once.

Lazy languages such as Haskell, don’t use this method.

Could we have done graphs of \( \lambda \)-terms? No. Substitution messes up sharing.

Example using recursion in Paulson’s notes.
Simply typed $\lambda$-calculus

Types

$$\tau ::= \text{int} \mid \tau \to \tau$$

Syntactic convention

$$\tau_1 \to \tau_2 \to \tau_3 = \tau_1 \to (\tau_2 \to \tau_3)$$

Simplifies types of curried functions.

Type checking

We check

• $M \, N : \tau$ iff $\exists \, \tau' . M : \tau' \to \tau$ and $N : \tau'$
• $\lambda x . M : \tau \to \tau'$ iff $\exists \, \tau . \text{if } x : \tau \text{ then } M : \tau'$$
• $n : \text{int}$

Semantics course covers this more formally, and types course next year in considerably more detail.
Types help find terms

Consider type \( (\tau_1 \rightarrow \tau_2 \rightarrow \tau_3) \rightarrow \tau_2 \rightarrow \tau_1 \rightarrow \tau_3 \)

Term \( \lambda f.M \) where \( f : (\tau_1 \rightarrow \tau_2 \rightarrow \tau_3) \) and \( M : \tau_2 \rightarrow \tau_1 \rightarrow \tau_3 \)

Therefore \( M = \lambda xy . N \) where \( x : \tau_2 \), \( y : \tau_1 \) and \( N : \tau_3 \).

Therefore \( N = f y x \)

Therefore \( \lambda fxy. f y x : (\tau_1 \rightarrow \tau_2 \rightarrow \tau_3) \rightarrow \tau_2 \rightarrow \tau_1 \rightarrow \tau_3 \)

Polymorphism and inference

ML type system supports polymorphism:

\[
\tau ::= \alpha \mid \forall \alpha. \tau \mid \ldots
\]

Types can be inferred using unification.

Overview

Encode evaluation order.

Encode control flow commands: for example Exit, exceptions, and goto.

Enables backtracking algorithms easily.

Key concept:

- don’t return, pass result to continuation. (This is what you did with the MIPS JAL (Jump And Link) instruction.)
Call-by-value

Definition:
1. \([x]_v(k) = k \, x\)
2. \([c]_v(k) = k \, c\)
3. \([\lambda x. M]_v(k) = k \, (\lambda (x, k'). \, [M](k'))\)
4. \([MN]_v(k) = [M](\lambda m. \, [N](\lambda n. \, m(n,k)))\)

Intuition:
- \([M]_v(k)\) means evaluate \(M\) and then pass the result to \(k\).
- \(k\) is what to do next.
Pairs not essential, but make the translation simpler.

Example: CBV

\[
[\lambda x. y]_v(k) \\
= k \, (\lambda (x, k'). \, [y]_v(k')) \\
= k \, (\lambda (x, k'). \, k' \, y)
\]

\[
[\lambda x. y \, z]_v(k) \\
= [\lambda x. y]_v(\lambda m. \, [z]_v(\lambda n. \, m(n,k))) \\
= [\lambda x. y]_v(\lambda m. \, (\lambda n. \, m(n,k)) \, z) \\
= (\lambda m. \, (\lambda n. \, m(n,k)) \, z) \, (\lambda (x, k'). \, k' \, y) \\
= (\lambda n. \, (\lambda (x, k'). \, k' \, y)) \, (n,k)) \, z \\
= k \, y
\]

Call-by-name

Definition:
- \([x]_n(k) = x \, k\)
- \([c]_n(k) = k \, c\)
- \([\lambda x. M]_n(k) = k \, (\lambda (x, k'). \, [M]_n(k'))\)
- \([MN]_n(k) = [M](\lambda m. \, m(k',[N]_n(k'))\)

Only application and variable are different. Don’t have to evaluate \(N\) before putting it into \(M\).

Mistake

CBN and CBV

For any closed term \(M\) \((FV(M) = \{\})\)
- \(M\) terminates with value \(v\) in the CBV \(\lambda\)-calculus, iff \([M]_v(\lambda x. x)\) terminates in both the CBV and CBN \(\lambda\)-calculus with value \(v\).
- \(M\) terminates with value \(v\) in the CBN \(\lambda\)-calculus, iff \([M]_n(\lambda x. x)\) terminates in both the CBV and CBN \(\lambda\)-calculus with value \(v\).
Encoding control

Consider trying to add an Exit instruction to the $\lambda$-calculus.

- Exit $M \rightarrow$ Exit (CBN and CBV)
- $M$ Exit $\rightarrow$ Exit (Just CBV)

When we encounter Exit execution is stopped.

- $(\lambda x. y)$ Exit $\rightarrow$ Exit (CBV)
- $(\lambda x. y)$ Exit $\rightarrow$ $y$ (CBN)

Encode as

- $\llbracket$ Exit $\rrbracket (k) = ()$ (Both CBV and CBN)

Example CBV

$$\llbracket (\lambda x. y) \text{ Exit} \rrbracket (k)$$
$$= \llbracket \lambda x. y \rrbracket (\lambda m. \llbracket \text{Exit} \rrbracket (\lambda n. m(n,k)))$$
$$= \llbracket \lambda x. y \rrbracket (\lambda m. ())$$
$$= (\lambda m. ()) (\lambda (x,k'). k' y)$$
$$\rightarrow (\lambda m. () (k'))$$

Example CBN

$$\llbracket (\lambda x. y) \text{ Exit} \rrbracket (k)$$
$$= \llbracket \lambda x. y \rrbracket (\lambda m. \llbracket \text{Exit} \rrbracket (\lambda n. m(n,k)))$$

Order of evaluation

With CBV we can consider two orders of evaluation:

Function first:
$$\llbracket M \ N \rrbracket_{\nu2}(k) = \llbracket M \rrbracket (\lambda m. \llbracket N \rrbracket (\lambda n. m(n,k)))$$

Argument first:
$$\llbracket M \ N \rrbracket_{\nu1}(k) = \llbracket N \rrbracket (\lambda n. \llbracket M \rrbracket (\lambda m. m(n,k)))$$
Example

Consider having two Exit expressions
- \([\text{Exit}_1]\) (k) = 1
- \([\text{Exit}_2]\) (k) = 2

Now, we can observe the two different translations by considering Exit_1 Exit_2:
- \([\text{Exit}_1 \text{Exit}_2]_v(k) = [\text{Exit}_1]_v(k)\) (Function first)
- \([\text{Exit}_1 \text{Exit}_2]_v(k) = [\text{Exit}_2]_v(k)\) (Argument first)

Example (continued)

\([\text{Exit}_1 \text{Exit}_2]_v(k) = [\text{Exit}_1] (\lambda m. [\text{Exit}_2] (\lambda n. m(n,k))) = 1 = [\text{Exit}_1](k)\)

\([\text{Exit}_1 \text{Exit}_2]_v(k) = [\text{Exit}_2] (\lambda n. [\text{Exit}_1] (\lambda m. m(n,k))) = 2 = [\text{Exit}_2](k)\)

Typed translation: CBV

Consider types:
\[ \tau ::= b \mid \tau \to \tau \mid \bot \]

Here b is for base types of constants, \(\bot\) for continuation return type.

We translate:
- \([\tau_1 \to \tau_2]_v = ([\tau_1]_v \to ([\tau_2]_v \to \bot)) \to \bot\)
- \([b]_v = b\)

If \(M : \tau\) then \(\lambda k. [M]_v(k) : ([\tau]_v \to \bot) \to \bot\)

Sometimes, we write \(T\) \(\tau\) for \((\tau \to \bot) \to \bot\)

Types guide translation

For function translation: Assume
- \(k : ([\tau_1 \to \tau_2]_v \to \bot)\)
- \(\lambda x.M : \tau_1 \to \tau_2\), hence \([M]_v : ([\tau_2]_v \to \bot) \to \bot\) if \(x : [\tau_1]_v\)

Find \(N\) such that \(k N : \bot\) therefore \(N : [\tau_1 \to \tau_2]_v\)

So, \(N = \lambda(x,k'). L\), where \(L : \bot\) if
- \(x : [\tau_1]_v\) and
- \(k' : [\tau_2]_v \to \bot\)

Therefore \(L = [M]_v(k')\)
- \([\lambda x.M]_v(k) = k (\lambda(x,k'). [M]_v(k'))\)
Types guide translation

Application translation (MN): Assume

• $k : (\tau_2 \rightarrow \perp)$
• $M : \tau_1 \rightarrow \tau_2$, hence $\llbracket M \rrbracket_v : (\llbracket \tau_1 \rightarrow \tau_2 \rrbracket_v \rightarrow \perp) \rightarrow \perp$
• $N : \tau_1$, hence $\llbracket N \rrbracket_v : (\llbracket \tau_1 \rrbracket_v \rightarrow \perp) \rightarrow \perp$

Find $L$ such that $\llbracket M \rrbracket_v L : \perp$ therefore $L : \llbracket \tau_1 \rightarrow \tau_2 \rrbracket_v \rightarrow \perp$

So, $L = \lambda m. L_1$, where $L_1 : \perp$ if $m : \llbracket \tau_1 \rightarrow \tau_2 \rrbracket_v = [\tau_1]_v ^* ([\tau_2]_v \rightarrow \perp) \rightarrow \perp$

Find $L_2$ such that $\llbracket N \rrbracket_v L_2 : \perp$ therefore $L_2 : \llbracket \tau_1 \rrbracket_v \rightarrow \perp$

Therefore $L_2 = \lambda n. L_3$ where $L_3 : \perp$ if $n : [\tau_1]_v$

Therefore $L_3 = m (n,k)$

$\llbracket M N \rrbracket_v (k) = \llbracket M \rrbracket (\lambda m. \llbracket N \rrbracket (\lambda n. m (n,k)))$

Other encodings

We can encode other control structures:

• Exceptions (2 continuations: normal and exception)
• Breaks and continues in loops (3 continuations: normal, break, and continue)
• Goto, jumps and labels
• call/cc (passing continuations into programs)
• backtracking

Exercises

• Find an example that evaluates differently for each of the three encodings, and demonstrate this.
• How would you perform a type call–by–name translation?

$\llbracket \tau_1 \rightarrow \tau_2 \rrbracket_n = ((T [\tau_1]_n)^* ([\tau_2]_n \rightarrow \perp)) \rightarrow \perp$

Aside: backtracking

Continuations can be a powerful way to implement backtracking algorithms. (The following is due to Olivier Danvy.)

Consider implementing regular expression pattern matcher in ML:

datatype re =
  Char of char (* “c” *)
| Seq of re * re (* re1 ; re2 *)
| Alt of re * re (* re1 | re2 *)
| Star of re * re (* re1 * *)
Implementation

Plan: use continuations to enable backtracking:

```haskell
fun "c" (a::xs) k = if a=c then (k xs) else false
| f "c" [] k = false
| f (re1 ; re2) xs k = f re1 xs (λys. f re2 ys k)
| f (re1 | re2) xs k = (f re1 xs k) orelse (f re2 xs k)
| f (re1 *) xs k =
    (k xs) orelse (f (re1 ; re1*) xs k)
```

Exercise: execute

```haskell
f ("a" ; "a" ; "a") ["a", "a"] (λxs. xs=[])
```

Example execution

```haskell
f ("a" ; "a" | "a") ["a", "a"] (λxs. xs=[])
```

Exercise

How could you extend this to

- count the number of matches; and
- allow matches that don’t consume the whole string?

Remove use of orelse by building a list of continuations for backtracking.

Comments

Not the most efficient regular expression pattern matching, but very concise code.

This style can implement efficient lazy pattern matchers or unification algorithms.
Now, we can consider extending the λ-calculus with:
- Assignment: $M := N$
- Read: $!M$

How can we do this by encoding?

Evaluation order made explicit (CPS transform).

Parameter used to carry state around.

We use the following encoding of state functions:
- SET: $s \ x \ y = \lambda z. IF \ z=x \ THEN \ y \ ELSE \ s \ x$
- GET: $s \ x = s \ x$

Note that, we ignore allocation in this encoding.
CPS and State

Definition: (This is a CBN translation.)
- \([x]_{\text{CPS}}(k,s) = x(k,s)\)
- \([c]_{\text{CPS}}(k,s) = k(c,s)\)
- \([\lambda x.M]_{\text{CPS}}(k,s) = \lambda i.(\lambda (x,k',s). [M](k',s))\)
- \([MN]_{\text{CPS}}(k,s) = [M][\lambda (m,s'). m([N]_{\text{CPS}}(k,s)),s)\)
- \([!M]_{\text{CPS}}(k,s) = [M][\lambda (v,s'). k\text{(GET }s'\text{ v }s'),s)\)
- \([\text{[M]}:=N]_{\text{CPS}}(k,s) = [M][\lambda (v,s'). ((\text{SET }s'\text{ v }[N]),s)\)

Let notation

We can simplify the presentation by using let notation:
- \([M](\lambda (m,s'). \lambda (n,s''). n(m,k,s'')). s'),s\)
  would be
  \(\text{let}_{\text{CPS}}\ (m,s') = [M](s)\) in
  \(\text{let}_{\text{CPS}}\ (n,s'') = [N](s')\) in \(m(n,k,s'')\)
where
- \(\text{let}_{\text{CPS}}\ (m,s') = [M](s)\) in \(T\)
  means
- \([M](\lambda (m,s'). T, s)\)

Note: This differs from the use of “let” in ISWIM.

CPS and State

Definition: (This is a CBV translation.)
- \([x]_{\text{CBV}}(k,s) = k(x,s)\)
- \([c]_{\text{CBV}}(k,s) = k(c,s)\)
- \([\lambda x.M]_{\text{CBV}}(k,s) = \lambda i.(\lambda (x,k',s'). [M](k',s')). s)\)
- \([MN]_{\text{CBV}}(k,s) = [M][\lambda (m,s'). m([N]_{\text{CBV}}(k,s)),s)\)
- \([!M]_{\text{CBV}}(k,s) = [M][\lambda (v,s'). k\text{(GET }s'\text{ v }s'),s)\)
- \([\text{[M]}:=N]_{\text{CBV}}(k,s) = [M][\lambda (v,s'). ((\text{SET }s'\text{ v }[N]),s)\)

Exercises

- Extend encoding with sequential composition \(M:N\)
- Translate: \([x]::=1; !x\)
- Translate: \((x.y.z)([x]::=(!x+1))\)
- Alter translation for CBV (assignment, application and variable cases need changing.)
- Redo translations above.

Mistake
It's getting complicated

Common theme, we are threading “stuff” through the evaluation:
• continuations
• state

If we add new things, for example IO and exceptions, we will need even more parameters.
Can we abstract the idea of threading “stuff” through evaluation?

Monad (Haskell)

Haskell provides a syntax and type system for threading “effects” through code.

Two required operations
• return : \( \tau \rightarrow \text{T}\tau \)
• \( \ggg : \text{T}\tau \rightarrow (\tau \rightarrow \text{T}\tau') \rightarrow \text{T}\tau' \) [bind]

Option/Maybe Monad

Types
• Option \( \tau \)

Definition
• Option \( \tau = \text{unit} + \tau \)

Operations
• return : \( \tau \rightarrow \text{Option } \tau \)
  return M = Some M
• \( \ggg : \text{Option } \tau \rightarrow (\tau \rightarrow \text{Option } \tau') \rightarrow \text{Option } \tau' \)
  \( \lambda xy. \text{case } x \text{ of None }=> \text{None} \mid \text{Some z }=> y \ z \)

Example

Imagine findx and findy are of type \( \text{unit } \rightarrow \text{Option } \tau \)

findx() \( \ggg \) \( \lambda x. \)

findy() \( \ggg \) \( \lambda y. \)

return (x,y)

This code is of type Option \( (\tau + \tau) \).

ML code:
  case findx() of
    None => None
  | Some x => case findy() of None => None
    | Some y => Some (x,y)
**Do notation**

```haskell
findx() >>= \x. 
findy() >>= \y. 
return (x,y)
```

Haskell has syntax to make this even cleaner:

```haskell
do {
  x <- findx();
  y <- findy();
  return (x,y)
}
```

**State monad**

**Types**
- State τ

**Definition**
- State τ = s → s * τ (s is some type for representing state, i.e. partial functions)

**Operations**
- return : τ → State τ
- >>= : (State τ → (x → State τ') → State τ' ) (infix)
- set : Loc → Int → State ()
- get : Loc → State Int
- new : () → State Loc

**Example: swap**

Assume x and y are two locations.

```haskell
do {
  z <- get x;
  w <- get y;
  set y z;
  set x w
}
```

**Haskell**

Read up on Haskell if this interests you.
Concluding remarks

Where this course sits

Summary

“Everything” can be encoded into the $\lambda$-calculus.

- Caveat: not concurrency!

Should we encode everything into $\lambda$-calculus?

The end