

A typed λ -calculus with call-by-name and call-by-value iteration

Herman Geuvers

Radboud University Nijmegen and Eindhoven University of Technology
The Netherlands

Types 2015, Tallinn
May 21, 2015

Contents

- ▶ Church data and Scott data
- ▶ Iteration, case and primitive recursion
- ▶ Call-by-value and Call-by name iteration
- ▶ Examples: addition and storage operator
- ▶ Properties (normalization) and connections

Church numerals

The most well-known Church data type

$$\bar{0} := \lambda x f. x$$

$$\bar{1} := \lambda x f. f x$$

$$\bar{2} := \lambda x f. f (f x)$$

$$\bar{p} := \lambda x f. f^p(x)$$

$$\text{Succ} := \lambda n. \lambda x f. f (n x f)$$

- ▶ The Church data types have **iteration** as basis. The numerals are iterators.
- ▶ **Iteration scheme** for Nat. (Let D be any type.)

$$\frac{d : D \quad f : D \rightarrow D}{\text{It } d f : \text{Nat} \rightarrow D} \quad \text{with} \quad \begin{array}{l} \text{It } d f \bar{0} \quad \rightarrow \quad d \\ \text{It } d f (\text{Succ } x) \quad \rightarrow \quad f (\text{It } d f x) \end{array}$$

- ▶ **Advantage**: quite a bit of **well-founded recursion** for free.
- ▶ **Disadvantage**: no pattern matching built in; predecessor is hard to define; the reduction is actually a conversion

Scott numerals

$$\langle 0 \rangle := \lambda x f . x$$

$$\langle 1 \rangle := \lambda x f . f \langle 0 \rangle$$

$$\langle 2 \rangle := \lambda x f . f \langle 1 \rangle$$

$$\langle n + 1 \rangle := \lambda x f . f \langle n \rangle$$

$$\text{succ} := \lambda p . \lambda x f . f p$$

- ▶ The Scott numerals are **case distinctors**.
- ▶ **Case scheme** for nat. (Let D be any type.)

$$\frac{d : D \quad f : \text{nat} \rightarrow D}{\text{Case } d f : \text{nat} \rightarrow D} \quad \text{with} \quad \begin{array}{l} \text{Case } d f \langle 0 \rangle \quad \rightarrow \quad d \\ \text{Case } d f (\text{succ } x) \quad \rightarrow \quad f x \end{array}$$

- ▶ $\text{Case } d f := \lambda x : \text{nat} . x d f$.
- ▶ **Advantage**: the predecessor can immediately be defined:
 $\text{pred} := \lambda p . p \langle 0 \rangle (\lambda y . y)$.
- ▶ **Disadvantage**: No recursion (which one has to get from somewhere else, e.g. a fixed point-combinator).

Church and Scott for general data types

Given a data type with

- ▶ constructors $\mathbf{c}_1, \dots, \mathbf{c}_k$,
- ▶ with arity $\text{ar}(i)$ of constructor \mathbf{c}_i .

The **Church encoding** is

$$\bar{\mathbf{c}}_i := \lambda x_1 \dots x_{\text{ar}(i)}. \lambda c_1 \dots c_k. c_i (x_1 \vec{c}) \dots (x_{\text{ar}(i)} \vec{c})$$

The **Scott encoding** is

$$\langle \mathbf{c}_i \rangle := \lambda x_1 \dots x_{\text{ar}(i)}. \lambda c_1 \dots c_k. c_i x_1 \dots x_{\text{ar}(i)}$$

Why study Scott data types?

- ▶ No recursion built in \Rightarrow more control over the definable functions. Brunel, Terui 2010; Baillot, De Benedetti, Ronchi della Rocca 2014.
- ▶ Direct access to predecessor. Destructor is constant-time.
- ▶ Combined Church-Scott data types give **Primitive Recursion scheme**. For Nat, these are known as **Parigot numerals**. (See also Stump, Fu 2014)

$$\frac{d : D \quad f : \text{Nat} \rightarrow D \rightarrow D}{\text{Rec } d f : \text{Nat} \rightarrow D} \quad \begin{array}{l} \text{Rec } d f 0 \quad \rightarrow \quad d \\ \text{Rec } d f (\text{Succ } x) \quad \rightarrow \quad f x (\text{Rec } d f x) \end{array}$$

One can define Rec in terms of It, but that's painful. (This is what Kleene found out at the dentist.)

Typing Church and Scott data types

- ▶ Church data types can be typed in polymorphic λ -calculus, $\lambda 2$.
E.g. for Church numbers: $\text{Nat} := \forall X. X \rightarrow (X \rightarrow X) \rightarrow X$.
- ▶ To type Scott data types we need $\lambda 2\mu$: $\lambda 2$ + positive recursive types:
 - ▶ $\mu X. \Phi$ is well-formed if X occurs **positively** in Φ .
 - ▶ Equality on types is the congruence generated from $\mu X. \Phi = \Phi[\mu X. \Phi / X]$.
 - ▶ Additional derivation rule:

$$\frac{\Gamma \vdash M : A \quad A = B}{\Gamma \vdash M : B}$$

E.g. for Scott numerals: $\text{nat} := \mu Y. \forall X. X \rightarrow (Y \rightarrow X) \rightarrow X$.

We will use this as

$$\text{nat} = \forall X. X \rightarrow (\text{nat} \rightarrow X) \rightarrow X.$$

The system $\lambda 2\mu\text{lt}$

The system $\lambda 2\mu\text{lt}$ combines polymorphic lambda calculus with positive recursive types and CBN and CBV iterators.

$$\mathsf{T} := \mathsf{TVar} \mid (\mathsf{T} \rightarrow \mathsf{T}) \mid \mu\mathsf{TVar}.\mathsf{T} \mid \forall\mathsf{TVar}.\mathsf{T},$$

$$\frac{x : A \in \Gamma}{\Gamma \vdash x : A} \qquad \frac{\Gamma \vdash M : A \rightarrow B \quad \Gamma \vdash N : A}{\Gamma \vdash MN : B}$$

$$\frac{\Gamma, x : A \vdash M : B}{\Gamma \vdash \lambda x : A.M : A \rightarrow B} \qquad \frac{\Gamma \vdash x : A}{\Gamma \vdash x : B} \text{ if } A = B$$

$$\frac{\Gamma \vdash M : \forall X.A}{\Gamma \vdash M : A[B/X]} \qquad \frac{\Gamma \vdash M : A}{\Gamma \vdash M : \forall X.A} \text{ if } X \notin \text{FV}(\Gamma)$$

The evaluation relation is weak-head reduction, \rightarrow^{wh} given by

$$\frac{}{(\lambda x.M)P \rightarrow^{wh} M[P/x]} \qquad \frac{M \rightarrow^{wh} N}{MP \rightarrow^{wh} NP}$$

Plus additional typing and reduction rules for Itcbv and Itcbrn .

Call-by-name and call-by-value iteration

To be able to define functions on Scott data types, we add for every data-type two program rules for **iteration**:

- ▶ **Itc_{bn}** for **call-by-name iteration** from data-type D to B ,
Idea: compute first constructor of the output of type B and pass to the appropriate continuation for type B .
- ▶ **Itc_{bv}** for **call-by-value iteration** from data-type D to B .
Idea: compute input value completely and then pass on to a continuation of type $B \rightarrow X$.

So, this function will be of type $\forall X.(B \rightarrow X) \rightarrow (D \rightarrow X)$

To be abbreviated to $\forall X.\neg_X B \rightarrow \neg_X D$

Data types in $\lambda 2\mu\text{lt}$

Data types have the following shape

$$D = \mu T. \forall X. (\Phi^1(T) \rightarrow X) \dots \rightarrow (\Phi^n(T) \rightarrow X) \rightarrow X,$$

with X positive in $\Phi^i(X)$. Phrased differently:

$$D = \forall X. (\Phi^1(D) \rightarrow X) \dots \rightarrow (\Phi^n(D) \rightarrow X) \rightarrow X,$$

Examples:

$$\text{bool} \quad := \quad \forall X. X \rightarrow X \rightarrow X$$

$$\text{nat} \quad := \quad \forall X. X \rightarrow (\text{nat} \rightarrow X) \rightarrow X$$

$$\text{list}_A \quad := \quad \forall X. X \rightarrow (A \rightarrow \text{list}_A \rightarrow X) \rightarrow X$$

$$\text{btree}_A \quad := \quad \forall X. (A \rightarrow X) \rightarrow (\text{btree}_A \rightarrow \text{btree}_A \rightarrow X) \rightarrow X$$

Constructors are defined terms in $\lambda 2\mu\text{lt}$

$$\begin{aligned}\text{nat} &:= \forall X. X \rightarrow (\text{nat} \rightarrow X) \rightarrow X \\ \text{list}_A &:= \forall X. X \rightarrow (A \rightarrow \text{list}_A \rightarrow X) \rightarrow X\end{aligned}$$

The constructors are (without type information with the λ):

$$\begin{aligned}\text{zero} &:= \lambda z. \lambda s. z : \text{nat} \\ \text{succ} &:= \lambda n. \lambda z. \lambda s. s \ n : \text{nat} \rightarrow \text{nat} \\ \text{nil} &:= \lambda n. \lambda c. n : \text{list}_A \\ \text{cons} &:= \lambda a. \lambda l. \lambda n. \lambda c. c \ a \ l : A \rightarrow \text{list}_A \rightarrow \text{list}_A\end{aligned}$$

We write $\langle d \rangle$ for the encoding of a data-element as a term.
So, for $n \in \mathbf{N}$, $\langle n \rangle : \text{nat}$, given by

$$\begin{aligned}\langle 0 \rangle &:= \text{zero} \\ \langle n + 1 \rangle &:= \text{succ} \langle n \rangle\end{aligned}$$

Derivation rule for **Itc**_{bn}, **Call-by-Name** iterator

Given $D = \forall X. (\Phi^1(X) \rightarrow D) \dots \rightarrow (\Phi^n(X) \rightarrow D) \rightarrow X$.

$$\frac{f_1 : \Phi^1(B) \rightarrow B \quad \dots \quad f_n : \Phi^n(B) \rightarrow B}{\text{Itc}_{B}^D f_1 \dots f_n : D \rightarrow B}$$

Plus reduction rule

If B is also a data-type, we have

$B = \forall X. (\Psi^1(X) \rightarrow B) \dots \rightarrow (\Psi^m(X) \rightarrow B) \rightarrow X$.

We could also write the rule for **Itc**_{bn} as follows (for A an arbitrary type):

$$\frac{f_1 : \Phi^1(B) \rightarrow B \quad \dots \quad f_n : \Phi^n(B) \rightarrow B \quad d : D \quad c_1 : \Psi^1(B) \rightarrow A \quad \dots \quad c_m : \Psi^m(B) \rightarrow A}{\text{Itc}_{B}^D f_1 \dots f_n d c_1 \dots c_m : A}$$

Idea: $c_1 : \Psi^1(B) \rightarrow A \dots c_m : \Psi^m(B) \rightarrow A$ are the continuations for data type B , one for each constructor of B .

Derivation rule for Itcbv, Call-by-Value iterator

Idea: $\neg_A B := B \rightarrow A$ is the type of the continuation; $c : \neg_A B$ is called “after the value has been computed”.

Let $D = \forall X. (\Phi^1(X) \rightarrow D) \dots \rightarrow (\Phi^n(X) \rightarrow D) \rightarrow X$.

$$\frac{f_1 : \neg_A B \rightarrow \neg_A \Phi^1(B) \quad \dots \quad f_n : \neg_A B \rightarrow \neg_A \Phi^n(B)}{\text{Itcbv}_B^D f_1 \dots f_n : \neg_A B \rightarrow \neg_A D}$$

Plus reduction rule

This rule can also be phrased in “fully applied form”:

$$\frac{f_1 : \neg_A B \rightarrow \neg_A \Phi^1(B) \quad \dots \quad f_n : \neg_A B \rightarrow \neg_A \Phi^n(B) \quad c : \neg_A B \quad d : D}{\text{Itcbv}_B^D f_1 \dots f_n c d : A}$$

Call-by-name and call-by-value iteration for nat

We show the case for $\text{nat} = \forall X. X \rightarrow (\text{nat} \rightarrow X) \rightarrow X$.

Call-by-name:

$$\frac{f_1 : B \quad f_2 : B \rightarrow B}{\text{Itc}bn \ f_1 \ f_2 : \text{nat} \rightarrow B}$$

with reduction rule:

$$\text{Itc}bn \ f_1 \ f_2 \ n \rightarrow n \ f_1 \ (\lambda x : \text{nat}. f_2 \ (\text{Itc}bn \ f_1 \ f_2 \ x))$$

Call-by-value:

$$\frac{f_1 : \neg_A \neg_A B \quad f_2 : \neg_A B \rightarrow \neg_A B}{\text{Itc}bv \ f_1 \ f_2 : \neg_A B \rightarrow \neg_A \text{nat}}$$

with reduction rule:

$$\text{Itc}bv \ f_1 \ f_2 \ c \ n \rightarrow n \ (f_1 \ c) \ (\text{Itc}bv \ f_1 \ f_2 \ (f_2 \ c)).$$

Note: $f_1 \ c : A$ and $\text{Itc}bv \ f_1 \ f_2 \ (f_2 \ c) : \text{nat} \rightarrow A$.

Example for the case nat: Call-by-Value Addition

$$\frac{f_1 : \neg_A \neg_A \text{nat} \quad f_2 : \neg_A \text{nat} \rightarrow \neg_A \text{nat}}{\text{Itcbv } f_1 f_2 : \neg_A \text{nat} \rightarrow \neg_A \text{nat}}$$

Define:

$$\widehat{\text{succ}} := \lambda c : \neg_A \text{nat}. \lambda n : \text{nat}. c (\text{succ } n) : \neg_A \text{nat} \rightarrow \neg_A \text{nat}$$

$$\hat{n} := \lambda c : \neg_A \text{nat}. c n : \neg_A \neg_A \text{nat} \quad (\text{for } n : \text{nat})$$

Then

$$\text{AddCBV} \quad : \quad \text{nat} \rightarrow \text{nat} \rightarrow \neg_A \neg_A \text{nat}$$

$$\text{AddCBV} \quad := \quad \lambda x y : \text{nat}. \lambda c : \neg_A \text{nat}. \text{Itcbv } \hat{y} \widehat{\text{succ}} c x$$

We have

$$\text{AddCBV } \langle n \rangle \langle m \rangle c \rightsquigarrow c \langle n + m \rangle$$

where $\langle n \rangle$ represents the number n .

Example for the case nat: Call-by-Name Addition

$$\frac{f_1 : \text{nat} \quad f_2 : \text{nat} \rightarrow \text{nat}}{\text{ltcbn } f_1 f_2 : \text{nat} \rightarrow \text{nat}}$$

$\text{AddCBN} := \lambda x y : \text{nat} . \text{ltcbn } y \text{ succ}.$

Then

$$\begin{aligned} \text{AddCBN } \langle n + 1 \rangle \langle m \rangle &\rightarrow \text{succ}(\text{AddCBN } \langle n \rangle \langle m \rangle) \\ &\rightarrow \lambda z s . s(\text{AddCBN } \langle n \rangle \langle m \rangle) \end{aligned}$$

- ▶ Weak-head reduction stops here.
- ▶ $\text{AddCBN } t q$ computes the first ‘pattern’ of the output and passes it on to the appropriate continuation. (In this case s .)

Storage operators

The idea of a storage operator (Krivine, Nour), Stor , for the natural numbers nat is:

For every $n \in \mathbf{N}$ there is a term t_n such that,
given $f : \text{nat} \rightarrow A$ and $M : \text{nat}$ with $M =_{\beta v n} \langle n \rangle$,

$$\text{Stor } M f \rightarrow f t_n.$$

To compute $f M$ call-by-value, we compute $\text{Stor } M f$.
This will first compute M and then feed this into f .

Storage operators in $\lambda 2\mu\text{lt}$

A **storage operator** for data type D is a term

$\text{Stor} : D \rightarrow \forall X. \neg_X \neg_X D$ satisfying, for all M with $M =_{\beta\text{vn}} \langle d \rangle$,

$$\text{Stor } M f \rightarrow f \langle d \rangle.$$

Define $\text{Stor}_{\text{nat}} : \text{nat} \rightarrow \forall X. \neg_X \neg_X \text{nat}$ and

$\text{Unstor}_{\text{nat}} : (\forall X. \neg_X \neg_X \text{nat}) \rightarrow \text{nat}$ as follows.

$$\text{Stor}_{\text{nat}} := \lambda n. \lambda f. \text{Itcbv } \widehat{\text{zero}} \widehat{\text{succ}} f n$$

$$\text{Unstor}_{\text{nat}} := \lambda f. \lambda z s. f (\lambda n. n z s)$$

with $\widehat{\text{zero}} := \lambda c. c \text{ zero}$, $\widehat{\text{succ}} := \lambda c m. c(\text{succ } m)$.

Lemma The term Stor_{nat} is a storage operator for nat and

$\text{Unstor}_{\text{nat}}(\text{Stor}_{\text{nat}} \langle n \rangle) =_{\beta\text{vn}} \langle n \rangle$ for all $n \in \mathbf{N}$.

Strong Normalization

The proof uses the well-known **saturated sets** construction to give an interpretation for types as sets of (untyped) terms.

X is a **saturated set** ($X \in \text{SAT}$) if

1. $X \subseteq \text{SN}$,
2. for x a variable and \vec{P} a sequence of terms, $x\vec{P} \in X$, if $x\vec{P} \in \text{SN}$,
3. X is closed under reduction: if $M \in X$ and $M \rightarrow N$, then $N \in X$,
4. X is closed under **weak-head-redex expansion**: if $M \in X$ and $N \rightarrow^{wh} M$, then $N \in X$,

Question: how to interpret a data type as a saturated set?

- ▶ `nat` consists of specific closed λ -terms: `zero`, `succ`, ...
- ▶ `Itcbv` and `Itcbs` impose closure conditions on `[nat]`, and possibly on the interpretation of other types.

SAT is a complete lattice, so we will define `[nat]` := NAT, where NAT is the **least fixed point** of some monotone $\mathcal{N} : \text{SAT} \rightarrow \text{SAT}$.

Intepreting nat

$[\text{nat}] := \text{NAT}$, where $\text{NAT} := \text{lfp}(\mathcal{N})$, with $\mathcal{N} : \text{SAT} \rightarrow \text{SAT}$ defined by

$$\mathcal{N}(X) := \{M \in \text{Term} \mid M \in \bigcap_{T \in \text{SAT}} T \rightarrow (X \rightarrow T) \rightarrow T \\ \wedge M \text{ satisfies (I) and (II)}\},$$

where (I) and (II) are defined by

- (I) $\forall B, A \in \text{SAT} \forall f_1 \in \neg_A \neg_A B \forall f_2 \in \neg_A B \rightarrow \neg_A B \forall c \in \neg_A B,$
 $\text{Itcbv } f_1 f_2 c M \in A$
- (II) $\forall B \in \text{SAT} \forall f_1 \in B \forall f_2 \in B \rightarrow B, \text{Itc}bn f_1 f_2 M \in B$

Properties of $[\text{nat}] := \text{NAT}$

- ▶ \mathcal{N} is monotone and $X \in \text{SAT} \Rightarrow \mathcal{N}(X) \in \text{SAT}$.
- ▶ NAT satisfies the desired fixed point equation:

$$\text{NAT} = \bigcap_{T \in \text{SAT}} T \rightarrow (\text{NAT} \rightarrow T) \rightarrow T.$$

- ▶ As a consequence $\vdash t : \text{nat} \Rightarrow t \in \text{NAT}$.
- ▶ The SN proof now follows from the soundness of the SAT-model.
- ▶ The proof can be generalised to other data types.

Are Itcbv and Itcbrn needed as primitives?

Equationally, Call-by-value can be defined by Call-by-name.
For the nat case assume CBN:

$$\frac{f_1 : B \quad f_2 : B \rightarrow B}{\text{Itcbrn } f_1 f_2 : \text{nat} \rightarrow B}$$

with reduction rule:

$$\text{Itcbrn } f_1 f_2 n \rightarrow n f_1 (\lambda x : \text{nat}. f_2 (\text{Itcbrn } f_1 f_2 x))$$

Then

$$\frac{\frac{f_1 : \neg_A \neg_A B \quad \frac{f_2 : \neg_A B \rightarrow \neg_A B}{\hat{f}_2 : \neg_A \neg_A B \rightarrow \neg_A \neg_A B}}{\text{Itcbrn } f_1 \hat{f}_2 : \text{nat} \rightarrow \neg_A \neg_A B}}{\lambda c. \lambda n. \text{Itcbrn } f_1 \hat{f}_2 n c : \neg_A B \rightarrow \neg_A \text{nat}}$$

Define this term as Itcbv $f_1 f_2$. This function computes CBV and:

$$\begin{aligned} \text{Itcbv } f_1 f_2 c \text{ zero} &\quad \rightarrow \quad f_1 c \\ \text{Itcbv } f_1 f_2 c (\text{succ } y) &\quad =_{\beta \text{vn}} \quad \text{Itcbv } f_1 f_2 (f_2 c) y. \end{aligned}$$

Can we mediate between Scott nat and Church Nat?

- ▶ From Church to Scott: $F : \text{Nat} \rightarrow \text{nat}$
 $F := \lambda n. n \text{ zero succ}$
- ▶ From Scott to Church: $G : \text{nat} \rightarrow \text{Nat}$
 $G := \text{Itcbn Zero Succ}$
- ▶ From Scott to Church, call-by-value: $H : \neg_A \text{Nat} \rightarrow \neg_A \text{nat}$
 $H := \text{Itcbv } \widehat{\text{Zero}} \widehat{\text{Succ}}$
with $\widehat{\text{Zero}} : \neg_A \neg_A \text{Nat}$, $\widehat{\text{Succ}} : \neg_A \text{Nat} \rightarrow \neg_A \text{Nat}$.

With F , we can transform any function on Scott numerals into a (call-by-name) function on Church numerals:

$$\text{If } f : \text{nat} \rightarrow A \text{ then } f \circ F : \text{Nat} \rightarrow A$$

For encoded numerals we have

$$f(F\bar{n}) =_{\beta_{vn}} f\langle n \rangle$$

So all function representable on nat are representable on Nat.

Can we mediate between Scott nat and Church Nat?

- ▶ From Scott to Church: $G : \text{nat} \rightarrow \text{Nat}$
 $G := \text{Itcbn Zero Succ}$
- ▶ From Scott to Church, call-by-value: $H : \neg_A \text{Nat} \rightarrow \neg_A \text{nat}$
 $H := \text{Itcbv } \widehat{\text{Zero}} \widehat{\text{Succ}}$
with $\widehat{\text{Zero}} : \neg_A \neg_A \text{Nat}$, $\widehat{\text{Succ}} : \neg_A \text{Nat} \rightarrow \neg_A \text{Nat}$.

With H , we can transform any function on Church numerals into a (call-by-value) function on Scott numerals:

$$\text{If } f : \text{Nat} \rightarrow A \text{ then } Hf : \text{nat} \rightarrow A$$

With G , we can transform any function on Church numerals into a (call-by-name) function on Scott numerals:

$$\text{If } f : \text{Nat} \rightarrow A \text{ then } f \circ G : \text{nat} \rightarrow A$$

For encoded numerals we have

$$f(G\langle n \rangle) =_{\beta_{vn}} f\bar{n}$$

So all functions representable on Nat are representable on nat .

Can we define a constant-time predecessor for Church Nat?

Remember that $\text{pred} : \text{nat} \rightarrow \text{nat}$ is defined by

$$\text{pred } n := n \text{ zero } (\lambda x. x).$$

So we can define $\text{Pred} : \text{Nat} \rightarrow \text{Nat}$ by composing

$$\text{Nat} \xrightarrow{F} \text{nat} \xrightarrow{\text{pred}} \text{nat} \xrightarrow{G} \text{Nat}$$

- ▶ NOT in general $\text{Pred}(\text{Succ } t) =_{\beta v n} t$.
- ▶ $\text{Pred} \overline{n+1} \rightarrow \bar{n}$ (for encodings of numbers).
- ▶ $\text{Pred} \overline{n+1} \rightarrow \bar{n}$ in $O(n)$ steps.

Open question: can we improve on this? (Or can we prove it cannot be improved?)

Conclusion and future work

- ▶ Given a fine analysis of CBV and CBN on data types using the Scott data type definition: proper reduction behaviour, confluent, normalizing, typed storage operator.
- ▶ Same representable functions, but (?) better algorithmic behaviour?
- ▶ Flexibly combining CBN and CBV functions
- ▶ Can we define the constant time predecessor on Church Nat in $\lambda 2\mu\text{lt}$?
- ▶ Can we be more refined, by using linear types and taking complexity into account? (cf. Brunel, Terui, many others).