

Lectures on The Lambda Calculus (II)

Masahiko Sato
Graduate School of Informatics, Kyoto University

Autumn school “Proof and Computation”
Fischbachau, Germany
October 6, 2016

Plan of the lectures

- I Background history, philosophy and *main idea*.
- II The free algebra \mathbb{T} of *threads*
- III The free algebra \mathbb{L} of *\mathbb{L} -expressions*. Church-Rosser Theorem and the pushout property.

These lectures are based on my work in progress.

de Bruijn algebra \mathbb{D} vs. our algebra \mathbb{L}

The de Bruijn algebra enjoys the following equation:

$$\mathbb{D} = \mathbb{N} + \lambda\mathbb{D} + (\mathbb{D} \mathbb{D})$$

We define the algebra \mathbb{L} of \mathbb{L} -expressions by the following two equations.

$$\mathbb{T} = \mathbb{N} + \lambda\mathbb{T}, \quad \mathbb{L} = \mathbb{T} + (\mathbb{L} \mathbb{L})^{\mathbb{N}}$$

This is an instance of the following algebra which depends on algebra τ :

$$\mathbb{L}_{\tau} = \tau + (\mathbb{L}_{\tau} \mathbb{L}_{\tau})^{\mathbb{N}}$$

By putting $\tau = \mathbb{T}$ we obtain \mathbb{L} .

de Bruijn algebra \mathbb{D} vs. our algebra \mathbb{L} (cont.)

$$\mathbb{L}_\tau = \tau + (\mathbb{L}_\tau \mathbb{L}_\tau)^\mathbb{N}$$

By putting $\tau = \mathbb{P}$ (closed threads), we obtain \mathbb{L}_0 consisting exactly of *closed \mathbb{L} -expressions* as follows.

$$\mathbb{L}_0 = \mathbb{P} + (\mathbb{L}_0 \mathbb{L}_0)^\mathbb{N}$$

In \mathbb{D} , it is not as easy as in our case. One can only get \mathbb{D}_0 , consisting of closed de Bruijn terms, by solving the following infinite family of equations. We put

$\mathbb{N}_i := \{n \in \mathbb{N} \mid n < i\}$ ($i \in \mathbb{N}$).

$$\mathbb{D}_0 = \mathbb{N}_0 + \lambda \mathbb{D}_1 + (\mathbb{D}_0 \mathbb{D}_0),$$

$$\mathbb{D}_1 = \mathbb{N}_1 + \lambda \mathbb{D}_2 + (\mathbb{D}_1 \mathbb{D}_1),$$

$$\mathbb{D}_2 = \mathbb{N}_2 + \lambda \mathbb{D}_3 + (\mathbb{D}_2 \mathbb{D}_2),$$

...

Embedding of de Bruijn algebra \mathbb{D} into our algebra \mathbb{L}

We can save \mathbb{D} from this situation by embedding \mathbb{D} into our algebra \mathbb{L} by defining the embedding function

$$[\cdot] : \mathbb{D} \rightarrow \mathbb{L}$$

as follows.

$$\begin{aligned} [n] &:= n, \\ [\lambda D] &:= \lambda [D], \\ [(D E)] &:= ([D] [E]). \end{aligned}$$

What is this function?

Embedding of de Bruijn algebra \mathbb{D} into our algebra \mathbb{L}

We can save \mathbb{D} from this situation by embedding \mathbb{D} into our algebra \mathbb{L} by defining the embedding function

$$[\cdot] : \mathbb{D} \rightarrow \mathbb{L}$$

as follows.

$$\begin{aligned} [n] &:= n, \\ [\lambda D] &:= \lambda [D], \\ [(D E)] &:= ([D] [E]). \end{aligned}$$

What is this function?

The identity function! \mathbb{D} is indeed a subset of \mathbb{L} . (So far we can apply λ only to threads. But we will extend it to be applicable to any \mathbb{L} -expression.)

The data structure of threads

We can view the algebra \mathbb{T} in the following two ways.

$$\mathbb{T} = \mathbb{N} + \lambda\mathbb{T} \quad \text{or} \quad \mathbb{T} = \mathbb{N} \times \mathbb{N}$$

In the first view, a typical element of \mathbb{T} can be written as $\lambda^i k$. This element is obtained from k by applying the constructor λ to k i times.

In the second view, the same element can be written i/k . From abstract syntax point of view, they are just two different notation (written in two different syntax) for the same *thread*.

For example, $k = \lambda^0 k$ in the first view, corresponds to $0/k$ in the second view. For this reason we will also write k for $0/k$.

The datatype \mathbb{T}

For technical reason, we will officially define \mathbb{T} by the following inductive definition, taking the second view above.

$$\frac{i \in \mathbb{N} \quad k \in \mathbb{N}}{i/k \in \mathbb{T}} \text{Thrd}$$

We will use q, r, s, t as meta variables ranging over threads.

Now, any thread t can be uniquely written $t = i/k$. In this case we say that *height* of t , written $\text{Ht}(t)$ is i and *depth* of t , written $\text{Dp}(t)$, is k .

λ as an operator on \mathbb{T}

We do not have λ in \mathbb{T} , but we can *define* it as an operator on \mathbb{T} :

$$\lambda \frac{i}{k} := \frac{i'}{k}.$$

In general, we define $\lambda^n : \mathbb{T} \rightarrow \mathbb{T}$ by

$$\lambda^n \frac{i}{k} := \frac{i + n}{k}.$$

So, $\lambda^n t$ increases height of t by n keeping its depth. For example, we have:

$$\lambda^i k = \lambda^i \frac{0}{k} = \frac{0 + i}{k} = i/k$$

Closed and open threads

A thread i/k is defined to be *closed* if $i > k$, and it is defined to be *open* if $i \leq k$.

Since $i/k = \lambda^i k$ we may visualize it as follows.

$$\lambda_{i-1} \lambda_{i-2} \cdots \lambda_1 \lambda_0 k$$

So, recalling the de Bruijn notation, we see that it is a closed term if and only if $i > k$.

Closed threads are also called *projections*. We write \mathbb{P} for the set $\{t \in \mathbb{T} \mid t \text{ is closed}\}$ of propositions.

Classification of \mathbb{T} and \mathbb{P} by height

We put

$$\mathbb{T}^n := \{t \in \mathbb{T} \mid \text{Ht}(t) \geq n\},$$

$$\mathbb{P}^n := \{t \in \mathbb{P} \mid \text{Ht}(t) \geq n\}.$$

We have:

$$\mathbb{T} = \mathbb{T}^0 \supsetneq \mathbb{T}^1 \supsetneq \mathbb{T}^2 \dots$$

$$\mathbb{P} = \mathbb{P}^0 \supsetneq \mathbb{P}^1 \supsetneq \mathbb{P}^2 \dots$$

We note that $\lambda^i : \mathbb{T}^n \rightarrow \mathbb{T}^{n+i}$ and $\lambda^i : \mathbb{P}^n \rightarrow \mathbb{P}^{n+i}$ (since \mathbb{P} is closed under application of λ). So, it is natural to write $\lambda^i \mathbb{T}^n$ for \mathbb{T}^{n+i} and $\lambda^i \mathbb{P}^n$ for \mathbb{P}^{n+i}

Closing and opening

Recall that:

$$\lambda : \mathbb{T}^n \rightarrow \mathbb{T}^{n+1} \quad (n \in \mathbb{N})$$

Not only λ has this arity, it is also a *bijjective* operator.

So it has its inverse

$$\bar{\lambda} : \mathbb{T}^{n+1} \rightarrow \mathbb{T}^n \quad (n \in \mathbb{N})$$

with the property $\bar{\lambda}\lambda t = t$ for all $t \in \mathbb{T}$ and $\lambda\bar{\lambda}t = t$ for all $t \in \lambda\mathbb{T}$.

Note that

$$\lambda : \mathbb{T} \rightarrow \lambda\mathbb{T} \quad \text{and} \quad \bar{\lambda} : \lambda\mathbb{T} \rightarrow \mathbb{T}$$

Given any thread, by applying λ sufficiently many times, it becomes a closed thread. So we will call λ a *closing* operator. Similarly $\bar{\lambda}$ will be called an *opening* operator.

Instantiation operation

We wish to define the *instantiation* operation which is a binary function of the form:

$$\langle \lambda \cdot \cdot \rangle : \lambda\mathbb{T} \times \mathbb{T} \rightarrow \mathbb{T}$$

This form imposes a natural condition that $\langle \cdot t \rangle$ is meaningful only if the first argument \cdot is of the form λr .

So, for any threads r and t , we wish to know what $\langle \lambda r t \rangle$ means. Our intuition is that it means the result of *applying* the function λr to its argument t .

Our idea is to define it by defining yet another binary function of the form:

$$\cdot \leftarrow \cdot : \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{T},$$

and then put:

$$\langle \lambda r t \rangle := r \leftarrow t.$$

Filling operation

We wish to define the *filling* operation which is a binary function of the form:

$$\cdot \leftarrow \cdot : \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{T}$$

So, for any threads r and t , we wish to know what $r \leftarrow t$ means.

Our idea is to define it by case analysis on the form of r . Namely, we say that r is *balanced* if $\text{Ht}(r) = \text{Dp}(r)$, and define the filling operation according as r is balanced or not.

Filling operation: Balanced case

In this case, $r = \lambda^i k$ where $i = \text{Ht}(r) = \text{Dp}(r) = k$.

Filling *succeeds* in this case, and we put:

$$r \leftarrow t := \uparrow^r t.$$

Here, $\uparrow^r : \mathbb{T} \rightarrow \mathbb{T}$ is a *lifting* operation defined by:

$$\uparrow^r \frac{j}{\ell} := \begin{cases} \frac{j+k}{\ell} & \text{if } j > \ell, \\ \frac{j+k}{\ell+k} & \text{if } j \leq \ell. \end{cases}$$

Note that for any $t \in \mathbb{T}$, $\uparrow^r t$ is closed (open) iff t is closed (open), and $\uparrow^r t = \lambda^{\text{Ht}(r)} t$ if t is closed. Also:

$$\uparrow^r : \mathbb{T}^n \rightarrow \mathbb{T}^{n+\text{Ht}(r)}.$$

Filling operation: Unbalanced case

In this case, $r = \lambda^i k$ where $i = \text{Ht}(r) \neq \text{Dp}(r) = k$.
Filling *fails* in this case, and we put:

$$r \leftarrow t := \Downarrow r.$$

Here, $\Downarrow : \mathbb{T} \rightarrow \mathbb{T}$ is a *lowering* operation defined by:

$$\Downarrow \frac{i}{k} := \begin{cases} \frac{i}{k} & \text{if } i \geq k, \\ \frac{i}{k-1} & \text{if } i < k. \end{cases}$$

The lowering function lowers depth of r by one only when r is open and $\text{Dp}(r) > 0$. Note that for any $r \in \mathbb{T}$, $\Downarrow r$ is closed (open) iff r is closed (open). Also:

$$\Downarrow : \mathbb{T}^n \rightarrow \mathbb{T}^n.$$

Filling operation (cont.)

Combining the balanced and unbalanced cases, we get the following definition of filling operation.

$$r \leftarrow t := \begin{cases} \uparrow^r t & \text{if } r \text{ is balanced,} \\ \downarrow r & \text{if } r \text{ is unbalanced.} \end{cases}$$

We can spell out the explicit definition as follows.

$$\frac{i}{k} \leftarrow \frac{j}{\ell} := \begin{cases} \frac{i}{k} & \text{if } i > k, \\ \frac{j+k}{\ell} & \text{if } i = k \text{ and } j > \ell, \\ \frac{j+k}{\ell+k} & \text{if } i = k \text{ and } j \leq \ell, \\ \frac{i}{k-1} & \text{if } i < k. \end{cases}$$

Definition of instantiation operation

Our plan was to define instantiation operation with arity:

$$\langle \lambda \cdot \cdot \rangle : \lambda\mathbb{T} \times \mathbb{T} \rightarrow \mathbb{T}$$

in terms of the filling operation with arity

$$\cdot \leftarrow \cdot : \mathbb{T} \times \mathbb{T} \rightarrow \mathbb{T}$$

by putting

$$\langle \lambda r t \rangle := r \leftarrow t.$$

Here, we define instantiation as follows.

Definition (Instantiation $\langle \cdot \cdot \rangle : \lambda\mathbb{T} \times \mathbb{T} \rightarrow \mathbb{T}$)

$$\langle r t \rangle := \bar{\lambda}r \leftarrow t$$

Definition of instantiation operation (cont.)

Putting $r = i/k$ ($i > 0$) and $t = j/\ell$ we have:

$$\left\langle \frac{i}{k} \frac{j}{\ell} \right\rangle := \frac{i-1}{k} \leftarrow \frac{j}{\ell} = \begin{cases} \frac{i-1}{k} & \text{if } i-1 > k, \\ \frac{j+k}{\ell} & \text{if } i-1 = k \text{ and } j > \ell, \\ \frac{j+k}{\ell+k} & \text{if } i-1 = k \text{ and } j \leq \ell, \\ \frac{i-1}{k-1} & \text{if } i-1 < k. \end{cases}$$

Examples of instantiation operation: Case 1

We wish to see the informal correctness of our definition of instantiation based on our intuitive understanding of threads.

Here, we consider the following case:

$$i > 0 \text{ and } i - 1 > k$$

$$\langle \frac{i}{k} r \rangle := \frac{i-1}{k} \leftarrow r = \frac{i-1}{k}$$

Let us say that $i = 4$ and $k = 2$, so that we have

$$\langle \lambda^4 2 r \rangle = \lambda^3 2 \leftarrow r = \lambda^3 2$$

Or, equivalently:

$$\langle \lambda^4 2 r \rangle = \langle \lambda_{xyzuy} r \rangle = \lambda_{yzuy} = \lambda^3 2$$

Examples of instantiation operation: Case 2

Here, we consider the following case:

$$i > 0, i - 1 = k \text{ and } j > l$$

$$\left\langle \frac{i}{k} \frac{j}{\ell} \right\rangle := \frac{i-1}{k} \leftarrow \frac{j}{\ell} = \frac{j+k}{\ell}$$

Let us say that $i = 3$, $k = 2$, $j = 1$ and $\ell = 0$ so that we have

$$\langle \lambda^3 2 \lambda^1 0 \rangle = \lambda^2 2 \leftarrow \lambda^1 0 = \lambda^3 0$$

Or, equivalently:

$$\langle \lambda^3 2 \lambda^1 0 \rangle = \langle \lambda_{xyz} x \lambda_u u \rangle = \lambda_{yz} \lambda_u u = \lambda_{yzu} u = \lambda^3 0$$

Examples of instantiation operation: Case 3

Here, we consider the following case:

$$i > 0, i - 1 = k \text{ and } j \leq l$$

$$\langle \frac{i}{k} \frac{j}{\ell} \rangle := \frac{i-1}{k} \leftarrow \frac{j}{\ell} = \frac{j+k}{\ell+k}$$

Let us say that $i = 3$, $k = 2$, $j = 1$ and $\ell = 1$ so that we have

$$\langle \lambda^3 2 \lambda^1 1 \rangle = \lambda^2 2 \leftarrow \lambda^1 1 = \lambda^3 3$$

Or, equivalently:

$$\langle \lambda^3 2 \lambda^1 1 \rangle = \langle \lambda_{xyz} x \lambda_u 1 \rangle = \lambda_{yz} \lambda_u 3 = \lambda_{yz} u 3 = \lambda^3 3$$

We changed **1** to **3** to avoid capturing by λ_{yz} .

Examples of instantiation operation: Case 4

Here, we consider the following case:

$$i > 0 \text{ and } i - 1 < k$$

$$\langle \frac{i}{k} r \rangle := \frac{i-1}{k} \leftarrow r = \frac{i-1}{k-1}$$

Let us say that $i = 2$ and $k = 2$, so that we have

$$\langle \lambda^2 2 r \rangle = \lambda^1 2 \leftarrow r = \lambda^1 1$$

Or, equivalently:

$$\langle \lambda^2 2 r \rangle = \langle \lambda_{xy} 2 r \rangle = \lambda_y 1 = \lambda^1 1$$

We changed **2** to **1** since it is not in the scope of λ_x anymore.

Instantiation under λ

Consider $\lambda_z(\lambda_{xy}(x y) z)$. In the traditional λ -calculus, we can convert the underlined β -redex using the ξ -rule as follows.

$$\frac{\overline{(\lambda_{xy}(x y) z) \rightarrow_{\beta} \lambda_y(z y)}^{\beta}}{\lambda_z(\lambda_{xy}(x y) z) \rightarrow_{\beta} \lambda_{zy}(z y)}^{\xi}$$

In our calculus, we wish to eliminate the ξ -rule, by extending the β -rule so that we can reduce the inner β -redex directly as shown below. Here, we note that

$$\lambda_z(\lambda_{xy}(x y) z) = ((\lambda^3 1 \lambda^3 0)^3 \lambda^1 0)^1$$

$$\lambda_{zy}(z y) = (\lambda^2 1 \lambda^2 0)^2$$

$$((\lambda^3 1 \lambda^3 0)^3 \lambda^1 0)^1 \rightarrow_{\beta} \langle (\lambda^3 1 \lambda^3 0)^3 \lambda^1 0 \rangle^1 = (\lambda^2 1 \lambda^2 0)^2$$

Instantiation at level n

What we will do here is to generalize the instantiation operation $\langle r \ t \rangle$ (which operates on $r \in \lambda\mathbb{T}$ and $t \in \mathbb{T}$) to $\langle r \ t \rangle^n$ with arity:

$$\langle \cdot \cdot \rangle^n : \lambda\mathbb{T}^n \times \mathbb{T}^n \rightarrow \mathbb{T}^n$$

We define this operation so that the following diagram commutes:

$$\begin{array}{ccc} \lambda\mathbb{T}^n \times \mathbb{T}^n & \xrightarrow{\langle \cdot \cdot \rangle^n} & \mathbb{T}^n \\ \bar{\lambda}^n \times \bar{\lambda}^n \downarrow & & \uparrow \lambda^n \\ \lambda\mathbb{T} \times \mathbb{T} & \xrightarrow{\langle \cdot \cdot \rangle} & \mathbb{T} \end{array}$$

Namely,

$$\langle r \ t \rangle^n := \lambda^n \langle \bar{\lambda}^n r \ \bar{\lambda}^n t \rangle$$

Instantiation Lemma

Lemma (Instantiation Lemma for threads)

$$n < m, r \in \mathbb{T}^{m+1}, s \in \mathbb{T}^m, t \in \mathbb{T}^n \vdash \\ \langle \langle r \ s \rangle^m t \rangle^n = \langle \langle r \ t \rangle^n \langle s \ t \rangle^n \rangle^{m-1}.$$

The above lemma is derivable from the following lemma.

Lemma (special case of the above lemma)

$$0 < m, r \in \mathbb{T}^{m+1}, s \in \mathbb{T}^m, t \in \mathbb{T} \vdash \\ \langle \langle r \ s \rangle^m t \rangle = \langle \langle r \ t \rangle \langle s \ t \rangle \rangle^{m-1}.$$

Substitution and Instantiation

$x \neq y, x \notin \text{FV}(M) \vdash$

$$K[x := L][y := M] = K[y := M][x := L[y := M]].$$

$$K \in \mathbb{T}^2, L \in \mathbb{T}^1, M \in \mathbb{T} \vdash \langle\langle K L \rangle^1 M\rangle = \langle\langle K M \rangle \langle L M \rangle\rangle.$$

Or, equivalently,

$$\langle\langle \lambda^2 K \lambda^1 L \rangle^1 M\rangle = \langle\langle \lambda^2 K M \rangle \langle \lambda^1 L M \rangle\rangle.$$

We can see that Instantiation operation naturally represents β -conversion rule as an algebraic operation.