The Conservation Theorem revisited

Philippe de Groote

INRIA-Lorraine – CRIN – CNRS
Campus Scientifique - B.P. 239
54506 Vandœuvre-lès-Nancy Cedex – FRANCE

Abstract. This paper describes a method of proving strong normalization based on an extension of the conservation theorem. We introduce a structural notion of reduction that we call β_S, and we prove that any λ-term that has a $\beta_I\beta_S$-normal form is strongly β-normalizable. We show how to use this result to prove the strong normalization of different typed λ-calculi.

1 Introduction

We present a method of proving strong normalization for several typed λ-calculi. This method is based on a extension of the conservation theorem.

The conservation theorem for λI [1, Chap. 11, §3.]), says that all the β-reducts of a λI-term that is not strongly β-normalizable are not strongly β-normalizable. As a corollary, any λI-term that has a β-normal form is strongly β-normalizable. This is one of the few results relating strong normalization to (weak) normalization.

The above property, of course, fails for λK-terms. The conservation theorem, when formulated for λK, concerns only the β_I-reducts of the λK-terms and not all its β-reducts. Hence, we lose the corollary and we cannot use the theorem to turn a proof of normalization into a proof of strong normalization.

In this paper, we state and prove a version of the corollary that holds for λK-terms. To this end, we introduce a new notion of reduction (β_S) that allows the contraction of the β_K-redexes to be delayed. We prove that any λ-term that has a $\beta_I\beta_S$-normal form is strongly $\beta_I\beta_S$-normalizable. Then, it turns out by postponement that any λ-term that has a $\beta_I\beta_S$-normal form is strongly β-normalizable. This is the central result of the paper and we show how to use it to prove the strong normalization of different typed λ-calculi.

The typed calculi we consider are calculi à la Church [3]. Yet we do not consider that the type discipline is part of the term formation rules. We rather consider that the typing rules allow well-typed terms to be singled out from the set of raw terms. Therefore our technical framework is the untyped λ-calculus and our strong normalization proofs rely on the so-called erasing trick. Nevertheless, we show in Section 6 how to extend our results to Barendregt's set $\mathcal{T}$ of pseudo-terms.

The paper is organized as follows. Section 2 is reminder of well-known definitions concerning the untyped λ-calculus. In Section 3, we introduce the notion of reduction β_S and we establish the postponement property of β_K-contractions with respect to β_S- and β_I-contractions. In Section 4, we prove the conservation theorem for $\beta_I\beta_S$-reductions. To this end, we use labeled λ-terms à la Lévy. In Section 5, we obtain, as a corollary of the postponement and conservation properties, that $\beta_I\beta_S$-normalization implies strong β-normalization. This result is used in section Section 6 to prove the strong normalization of Church's simply typed λ-calculus. Then we show how to extend the proof to Barendregt's λ-cube. We present our conclusions in Section 7.

In:
M.Bezem and J.F.Groote (Eds)
Proceedings of the International Conference on Typed Lambda Calculi and Applications,
Lecture Notes in Computer Science, Vol. 664, Springer-Verlag (1993), pp. 163-178.

2 Basic Definitions

In this section we remind the reader of some basic notions about the type-free λ-calculus. The definitions we give, which are taken from [1], concern mainly the concepts of reduction and normalization. The reader familiar with this material may proceed directly to Section 3.

Type-free λ-terms are built up on an infinite numerable set of variable $\mathcal{V}$ according to the following definition.

Definition 1. *The set Λ of λ-terms is inductively defined as follows:*

i. $x \in \mathcal{V} \Rightarrow x \in \Lambda$,
ii. $x \in \mathcal{V}, M \in \Lambda \Rightarrow \lambda x.\, M \in \Lambda$,
iii. $M, N \in \Lambda \Rightarrow (M\, N) \in \Lambda$.

The symbol λ is a binding operator and the notions of free and bound occurrences of a variable are as usual in logic. In particular, the free occurrences of x in M are bound in $\lambda x.\, M$. The set of variables occurring free in a λ-term M is denoted $\mathrm{FV}(M)$. The λ-terms that can be transformed into each other by renaming their bound variables are identified. We also consider that some variable convention prevents us from clashes of variables (see [1, page 26], also [5] for a formal treatment).

Definition 2. *Any binary relation $R \subset \Lambda \times \Lambda$ is called a notion of reduction. If R is a notion of reduction and $(M, N) \in R$, M is called a R-redex and N is called the contractum of M.*

Given some notion of reduction R, one defines the following binary relations between λ-terms: the relation of R-contraction ($\rightarrow_R$), the relation of R-reduction ($\twoheadrightarrow_R$), the relation of strict R-reduction ($\xrightarrow{+}_R$), and the relation of R-conversion ($\twoheadleftarrow\!\!\twoheadrightarrow_R$).

Definition 3. *Let R be a notion of reduction. The corresponding contraction relation is the least relation containing R, and compatible with the λ-term formation rules. This relation is inductively defined by the following rules:*

$$i.\quad M \rightarrow_R N \quad \text{if } (M, N) \in R, \qquad ii.\quad \frac{M \rightarrow_R N}{\lambda x.\, M \rightarrow_R \lambda x.\, N},$$

$$iii.\quad \frac{M \rightarrow_R N}{(M\, O) \rightarrow_R (N\, O)}, \qquad iv.\quad \frac{M \rightarrow_R N}{(O\, M) \rightarrow_R (O\, N)}.$$

The relation of R-reduction ($\twoheadrightarrow_R$), is the transitive, reflexive closure of the relation of R-contraction; the relation of strict R-reduction ($\xrightarrow{+}_R$) is the transitive closure of the relation of R-contraction; the relation of R-conversion ($\twoheadleftarrow\!\!\twoheadrightarrow_R$) is the transitive, reflexive, symmetric closure of the relation of R-contraction.

Any notion of reduction induces also the corresponding concepts of normal form, normalization, and strong normalization.

Definition 4. *Let R be a notion of reduction. A λ-term M is called a R-normal form if and only if there does not exist $N \in \Lambda$ such that $M \rightarrow_R N$.*

Definition 5. *Let R be a notion of reduction. A λ-term M is called R-normalizable if and only if there exists a R-normal form N such that $M \twoheadrightarrow_R N$.*

Definition 6. *Let R be a notion of reduction. A λ-term M is called strongly R-normalizable if and only if there exists an upper bound to the length n of any sequence of R-contractions starting in M:*

$$M \equiv M_0 \to_R M_1 \to_R \ldots \to_R M_n.$$

$M[x{:=}N]$ denotes the result of substituting a λ-term N for the free occurrences of a variable x in a λ-term M. This operation is defined as follows.

Definition 7.

i. $x[x{:=}N] \equiv N$,
ii. $y[x{:=}N] \equiv y \quad$ *if* $x \not\equiv y$,
iii. $\lambda y.\, M[x{:=}N] \equiv \lambda y.\, M[x{:=}N]$,
iv. $(M\,O)[x{:=}N] \equiv (M[x{:=}N]\,O[x{:=}N])$.

The principal notion of reduction of the λ-calculus is the notion of reduction β.

Definition 8. *The notion of reduction β is defined by the following contraction rule:*

$\beta: \quad (\lambda x.\, M\, N) \to M[x{:=}N]$.

3 The Notions of Reduction β_I, β_K, and β_S

A β-redex $(\lambda x.\, M\, N)$ is called an I-redex if $x \in \mathrm{FV}(M)$ and a K-redex otherwise. This distinction allows the notion of β-reduction to be split into the two notions of β_I- and β_K-reduction.

Definition 9. *The notions of reduction β_I and β_K are respectively defined by the following contraction rules:*

$\beta_I: \quad (\lambda x.\, M\, N) \to M[x{:=}N] \quad$ *if* $x \in \mathrm{FV}(M)$,
$\beta_K: \quad (\lambda x.\, M\, N) \to M \quad$ *if* $x \notin \mathrm{FV}(M)$.

A possible strategy to prove strong normalization when dealing with two different notions of reduction is to take advantage of some postponement property. For instance, the postponement of η-contractions with respect to β-contractions is a well known property (see [1, page 386]) from which it follows that any strongly β-normalizable λ-term is strongly $\beta\eta$-normalizable.

The postponement strategy may be used when two notions of reduction, say R_1 and R_2, are such that

1. any term of interest is strongly R_1-normalizable,
2. any term of interest is strongly R_2-normalizable,
3. the contraction of a R_2-redex cannot create a R_1-redex.

In the case of β_I and β_K, Conditions 1 and 2 are satisfied. Indeed, we have that

1. any $M \in \Lambda$ that is β_I-normalizable is strongly β_I-normalizable—this is an easy consequence of the conservation theorem for λI (see [1, Chap. 11, §3.]),
2. any $M \in \Lambda$ is strongly β_K-normalizable—this is obvious because the length of any β_K-contractum is strictly less than the length of the corresponding β_K-redex.

Unfortunately, neither β_I-contractions nor β_K-contractions may be postponed. This is shown by the following counterexamples:

$$
\begin{array}{ll}
(\lambda x.\,(x\,M)\,\lambda y.\,z) \to_{\beta_I} (\lambda y.\,z\,M) & \text{a } \beta_K\text{-redex is created,} \\
((\lambda y.\,\lambda x.\,x\,M)\,N) \to_{\beta_K} (\lambda x.\,x\,N) & \text{a } \beta_I\text{-redex is created.}
\end{array}
$$

In order to fix this problem, we are going to introduce a third notion of reduction, namely β_S. This notion of reduction is such that

1. if $M \twoheadrightarrow_{\beta_S} N$ then $M \twoheadleftarrow\!\!\twoheadrightarrow_{\beta} N$,
2. β_K-contractions may be postponed with respect of $\beta_I\beta_S$-contractions,
3. the conservation theorem holds for β_I and β_S.

The first of these three properties is immediate. The second one and the third one will be established respectively as Lemma 3 and Theorem 4.

Definition 10. *The notion of reduction β_S is defined by the following contraction rule:*

$$\beta_S: \quad ((\lambda x.\,M\,N)\,O) \to (\lambda x.\,(M\,O)\,N) \quad \textit{if } x \notin \mathrm{FV}(M).$$

Notice that we have by the variable convention that $x \notin \mathrm{FV}(O)$.

Lemma 1. (postponement of β_K-contractions) *Let $R \in \{\beta_I, \beta_S\}$. Let $M, N, O \in \Lambda$ be such that*

$$M \to_{\beta_K} N \quad \textit{and} \quad N \to_R O$$

Then there exists $P \in \Lambda$ such that

$$M \xrightarrow{+}_{\beta_I\beta_S} P \quad \textit{and} \quad P \twoheadrightarrow_{\beta_K} O$$

Proof. The proof is by induction on the derivation of $M \to_{\beta_K} N$, distinguishing subcases according to the way $N \to_R O$. The details are given in Appendix A. □

4 The Conservation Theorem for β_I and β_S

In this section we establish the main technical result of this paper, namely that any $\beta_I\beta_S$-normalizable λ-term is strongly $\beta_I\beta_S$-normalizable. To this end we use labeled λ-terms à la Lévy (see [1, Chap. 14]).

Definition 11. *The set $\Lambda^{\mathbb{N}}$ of labeled λ-terms is inductively defined as follows:*

i. $n \in \mathbb{N}, x \in \mathcal{V} \Rightarrow (x)^n \in \Lambda^{\mathbb{N}}$,
ii. $n \in \mathbb{N}, x \in \mathcal{V}, M \in \Lambda^{\mathbb{N}} \Rightarrow \lambda x.\,M^n \in \Lambda^{\mathbb{N}}$,
iii. $n \in \mathbb{N}, M \in \Lambda^{\mathbb{N}}, N \in \Lambda^{\mathbb{N}} \Rightarrow (M\,N)^n \in \Lambda^{\mathbb{N}}$.

We use $M, N, O, \ldots$ as metavariable ranging on labeled λ-terms. We use the notation M^n or $(M)^n$ to stress that the outermost label of a labeled λ-term M is n. Thus, according to this last convention, the meta-expressions M, M^n, and $(M)^n$ used in a same context stand exactly for the same labeled λ-term. We also write M^{+m} or $(M)^{+m}$ for the labeled λ-term obtained by adding m to the outermost label of a labeled λ-term M. Thus, if the outermost label of M is n, then $(M)^{+m}$ denotes the same term than $(M)^{n+m}$.

Let $M \in \Lambda^{\mathbb{N}}$. We write $|M|$ for the (unlabeled) λ-term obtained by erasing all the labels in M. Therefore, for $M \in \Lambda^{\mathbb{N}}$, we have $|M| \in \Lambda$.

Now, let $M \in \Lambda$. We identify M with the labeled λ-term M' such that (i) $|M'| \equiv M$, (ii) all the labels in M' are 0. Therefore, we have that $\Lambda \subset \Lambda^{\mathbb{N}}$.

Labels will be used as counters to record the number of contracted redexes when reducing a term. This idea motivates the next two definitions. The first one generalizes the operation of substitution on labeled λ-terms. The second one introduces labeled versions of the notions of reduction β_I and β_S.

Definition 12. *The substitution $M[x{:=}N]$ of a labeled λ-term M for the free occurrences of a variable x in a labeled λ-term N is defined as follows.*

i. $(x)^m[x{:=}N^n] \equiv (N)^{m+n}$,
ii. $(y)^m[x{:=}N^n] \equiv (y)^m \qquad$ *if* $x \not\equiv y$,
iii. $\lambda y.\, M^m[x{:=}N^n] \equiv \lambda y.\, M[x{:=}N^n]^m$,
iv. $(M\,O)^m[x{:=}N^n] \equiv (M[x{:=}N^n]\,O[x{:=}N^n])^m$.

The labeled versions of the notions of reduction β_I and β_S are called respectively β_I^+ and β_S^+.

Definition 13. *The notions of reduction β_I^+ and β_S^+ are respectively defined by the following contraction rules:*

$$\begin{array}{lll} \beta_I^+ : & (\lambda x.\, M^m\, N)^n \to (M[x{:=}N])^{+(m+n+1)} & \text{if } x \in \mathrm{FV}(M), \\ \beta_S^+ : & ((\lambda x.\, M^m\, N)^n\, O)^o \to (\lambda x.\, (M\, O)\, N)^{m+n+o+1} & \text{if } x \notin \mathrm{FV}(M). \end{array}$$

To adapt the definition of contraction is straightforward.

Definition 14. *Let $R \subset \Lambda^{\mathbb{N}} \times \Lambda^{\mathbb{N}}$. The relation of R-contraction is inductively defined by the following rules:*

$$i.\quad M \to_R N \quad \text{if } (M,N) \in R, \qquad ii.\quad \frac{M \to_R N}{\lambda x.\, M^n \to_R \lambda x.\, N^n},$$

$$iii.\quad \frac{M \to_R N}{(M\,O)^n \to_R (N\,O)^n}, \qquad iv.\quad \frac{M \to_R N}{(O\,M)^n \to_R (O\,N)^n}.$$

The definitions of R-reduction, strict R-reduction, and R-conversion are unchanged.

The notion of reduction $\beta_I^+ \cup \beta_S^+$ satisfies the Church-Rosser property.

Theorem 1. (Church-Rosser) *Let $M, N, O \in \Lambda^{\mathbb{N}}$ be such that*

$$M \twoheadrightarrow_{\beta_I^+ \beta_S^+} N \text{ and } M \twoheadrightarrow_{\beta_I^+ \beta_S^+} O.$$

Then there exists $P \in \Lambda^{\mathbb{N}}$ such that

$$N \twoheadrightarrow_{\beta_I^+ \beta_S^+} P \text{ and } O \twoheadrightarrow_{\beta_I^+ \beta_S^+} P.$$

Proof. *One uses the lemma of Hindley-Rossen: first one establishes that β_I^+ and β_S^+ are individually Church-Rosser; then one shows that β_I^+ and β_S^+ commute. This can be done using the method of Tait and Martin-Löf. The details are given in Appendix B.* ☐

The next step is to define the weight of a term as the sum of all its labels.

Definition 15. *The weight $\Theta[M]$ of a labeled λ-term M is defined as follows.*

i. $\Theta[(x)^n] = n$,
ii. $\Theta[\lambda y.\, M^n] = n + \Theta[M]$,
iii. $\Theta[(M\, N)^n] = n + \Theta[M] + \Theta[N]$.

We are now in the position of proving the conservation theorem for β_I and β_S. The proof consists of three easy lemmas.

Lemma 2. *Let $R \in \{\beta_I^+, \beta_S^+\}$, and let $M, N \in \Lambda^{\mathbb{N}}$ be such that $M \xrightarrow{+}_R N$. Then $\Theta[M] < \Theta[N]$.*

Proof. *The statement is proven for one-step reduction by a straightforward induction on the definition of contraction. Notice, in the case of β_I^+, the part played by the proviso $x \in \mathrm{FV}(M)$.* □

Lemma 3. (Conservation for β_I^+ and β_S^+) *Let $M \in \Lambda^{\mathbb{N}}$ be $\beta_I^+\beta_S^+$-normalizable. Then M is strongly $\beta_I^+\beta_S^+$-normalizable.*

Proof. *Since M is $\beta_I^+\beta_S^+$-normalizable, it has at least one $\beta_I^+\beta_S^+$-normal form. On the other hand, by the Church-Rosser property, M has at most one $\beta_I^+\beta_S^+$-normal form. So, let M^* be the unique $\beta_I^+\beta_S^+$-normal form of M. Then, according to Lemma 4, we have that the length of any sequence of $\beta_I^+\beta_S^+$-reduction starting in M is bounded by $\Theta[M^*] - \Theta[M]$.* □

Lemma 4. *Let $M, N \in \Lambda$ be such that $M \to_{\beta_I\beta_S} N$. Then there exist $M^*, N^* \in \Lambda^{\mathbb{N}}$ such that $|M^*| \equiv M$, $|N^*| \equiv N$, and $M^* \to_{\beta_I^+\beta_S^+} N^*$.*

Proof. *Follows from the fact that $\Lambda \subset \Lambda^{\mathbb{N}}$ and that, if $P \to_{\beta_I^+\beta_S^+} Q$, then $|P| \to_{\beta_I\beta_S} |Q|$.* □

Theorem 2. (Conservation for β_I and β_S) *Let $M \in \Lambda$ be $\beta_I\beta_S$-normalizable. Then M is strongly $\beta_I\beta_S$-normalizable.*

Proof. *Follows from Lemma 4 and Lemma 4.* □

5 Main Result

Our goal is to take advantage of the result established in the previous section when dealing with the notion of reduction β.

Theorem 4 may be seen as a generalized version of the conservation theorem. Indeed the usual conservation theorem [1, Chap. 11, §3.] appears as a particular case of Theorem 4. Nevertheless, we may go further in the generalization and state the following theorem, which is the central result of this paper.

Theorem 3. (Generalized conservation) *Let $M \in \Lambda$ be $\beta_I\beta_S$-normalizable. Then M is strongly β-normalizable.*

Proof. *Suppose that there exists a λ-term $M \in \Lambda$ that is $\beta_I\beta_S$-normalizable but that is not strongly β-normalizable. Then there exists an infinite sequence of terms $M_0, M_1, M_2, \ldots \in \Lambda$ such that*

$$M \equiv M_0 \text{ and } \forall i \in \mathbb{N}, M_i \to_\beta M_{i+1}.$$

This sequence must be such that

$$\forall k \in \mathbb{N}, \exists l \in \mathbb{N}, l \geq k \text{ and } M_l \rightarrow_{\beta_I} M_{l+1}$$

otherwise there would exist an infinite sequence of β_K-contractions, which is absurd. But then, by Lemma 3, it is possible to construct an infinite sequence of $\beta_I\beta_S$-contractions starting in M, and this contradicts the fact that, by Theorem 4, M is strongly $\beta_I\beta_S$-normalizable. □

6 Application to Typed λ-Calculi

In this section we show how to use Theorem 5 to prove the strong normalization of several typed λ-calculi. We first establish the strong normalization of Church's simply typed λ-calculus. Then we discuss how to extend the proof to Barendregt's λ-cube by following the edges of the cube in the three possible directions.

6.1 Church's Simply Typed λ-Calculus

We first define the raw syntax of simple types and simply typed λ-terms.

Let $\mathcal{A}$ be a set of symbols called atomic types. The set $\mathcal{S}$ of simple types is inductively defined as follows.

Definition 16.

i. $a \in \mathcal{A} \Rightarrow a \in \mathcal{S}$,
ii. $\alpha, \beta \in \mathcal{S} \Rightarrow (\alpha \rightarrow \beta) \in \mathcal{S}$

Let $\mathcal{V}$ be the set of variables of type α. We define the set $\Lambda^{\rightarrow}$ of raw simply typed λ-terms.

Definition 17. *The set $\Lambda^{\rightarrow}$ of raw simply typed λ-terms is inductively defined as follows:*

i. $x \in \mathcal{V} \Rightarrow x \in \Lambda^{\rightarrow}$,
ii. $x \in \mathcal{V}, \alpha \in \mathcal{S}, M \in \Lambda^{\rightarrow} \Rightarrow (\lambda x\!:\!\alpha.\, M) \in \Lambda^{\rightarrow}$,
iii. $M, N \in \Lambda^{\rightarrow} \Rightarrow (M\, N) \in \Lambda^{\rightarrow}$.

If $M \in \Lambda^{\rightarrow}$ and $\alpha \in \mathcal{S}$, an expression of the form $M : \alpha$ is called a statement. M is called the subject of the statement and α is called the predicate. A statement whose subject is a variable is called a declaration. A sequence of declarations whose subjects are all distinct is called a typing context. We will use $\Gamma, \Delta, \ldots$ as metavariables ranging over typing contexts.

The notion of well-typed term is defined by providing a proof system to derive typing judgements of the shape

$$\Gamma \vdash M : \alpha$$

where Γ is a typing context, $M \in \Lambda^{\rightarrow}$, and $\alpha \in \mathcal{S}$.

Definition 18.

$$\Gamma \vdash x : \alpha \quad \text{if } x : \alpha \in \Gamma \qquad \text{(variable)}$$

$$\frac{\Gamma, x : \alpha \vdash M : \beta}{\Gamma \vdash (\lambda x \!:\! \alpha.\, M) : (\alpha \to \beta)} \qquad \textit{(abstraction)}$$

$$\frac{\Gamma \vdash M : (\alpha \to \beta) \qquad \Gamma \vdash N : \alpha}{\Gamma \vdash (M\, N) : \beta} \qquad \textit{(application)}$$

We now establish the normalization of Church's simply typed λ-calculus by giving a proof due to Turing [9].

Theorem 4. (Normalization) *Let Γ be a context, and let $M \in \Lambda^{\to}$ and $\alpha \in \mathcal{S}$ be such that*

$$\Gamma \vdash M : \alpha$$

then M has a β-normal form.

Proof. One defines the order of a β-redex $((\lambda x \!:\! \alpha.\, M)\, N)$ as the length of the type assigned to $(\lambda x \!:\! \alpha.\, M)$. Now, consider some β-contraction $P \to_\beta Q$ where the contracted redex is $((\lambda x \!:\! \alpha.\, M)\, N)$. The redexes in Q are of six kinds:

1. *redexes occurring in P disjointly with $((\lambda x \!:\! \alpha.\, M)\, N)$; these redexes are unchanged;*
2. *redexes occurring in M, and possibly modified by the substitution $M[x{:=}N]$; their orders are unchanged;*
3. *new redexes $((\lambda y \!:\! \alpha_1.\, N_1)\, O_i)$, if $N \equiv (\lambda y \!:\! \alpha_1.\, N_1)$ and $(x\, O_i)$ occurs in M; the order of these redexes is the length of the type assigned to N, which is less than the order of $((\lambda x \!:\! \alpha.\, M)\, N)$;*
4. *a new redex $((\lambda y \!:\! \alpha_1.\, M_1[x{:=}N])\, O)$, if $M \equiv (\lambda y \!:\! \alpha_1.\, M_1)$ and $((\lambda x \!:\! \alpha.\, M)\, N)$ occurs in P as the left subterm of an application $(((\lambda x \!:\! \alpha.\, M)\, N)\, O)$; the order of this redex is the length of the type assigned to M, which is less than the order of $((\lambda x \!:\! \alpha.\, M)\, N)$;*
5. *a new redex $((\lambda y \!:\! \alpha_1.\, N_1)\, O)$, if $N \equiv (\lambda y \!:\! \alpha_1.\, N_1)$, $M \equiv x$, and $((\lambda x \!:\! \alpha.\, x)\, N)$ occurs in P as the left subterm of an application $(((\lambda x \!:\! \alpha.\, x)\, N)\, O)$; the order of this redex is the length of the type assigned to N, which is less than the order of $((\lambda x \!:\! \alpha.\, M)\, N)$;*
6. *redexes occurring in N and possibly multiplied by the substitution $M[x{:=}N]$; their orders are kept unchanged.*

The normalization procedure runs as follows: contract one of the β-redexes of highest order that is innermost, and repeat this process until no more redex remains.

Since the contracted redex, say $((\lambda x \!:\! \alpha.\, M)\, N)$, is chosen innermost, the order of any redex occurring in N is strictly less than the order of the contracted redex. Therefore, at each step, the orders of the created or multiplied redexes (of kind 3, 4, 5, and 6) are strictly less than the order of the redex that disappears. The theorem follows by an induction up to ω^2. □

Normalization concerns only the existence of normal forms. Hence, to prove a normalization theorem, it sufficient to exhibit one normalization procedure. Strong normalization is, in general, more complex to establish. One has to show that any reduction strategy is normalizing. This is why Theorem 5 is interesting: it allows one to reduce a proof of strong normalization to a proof of normalization. This is illustrated by the proof of the next theorem.

Theorem 5. (Strong normalization) *Let Γ be a context, and let $M \in \Lambda^{\to}$ and $\alpha \in \mathcal{S}$ be such that*

$$\Gamma \vdash M : \alpha$$

then M is strongly β-normalizable.

Proof. According to Theorem 5, it is sufficient to show that M has a $\beta_I\beta_S$-normal form.

First one defines the order of a β_S-redex $(((\lambda x\!:\!\alpha.\,M)\,N)\,O)$ as the length of the type assigned to $(\lambda x\!:\!\alpha.\,M)$. Then, it is straightforward to replay the proof of Theorem 6.1. $\square$

A technical advantage of the above proof, as opposed to a proof à la Tait (see [15, App. 2]) is that our proof is arithmetizable. We will come back to this point in the conclusions.

6.2 Barendregt's λ-Cube

Barendregt's λ-cube is a system of eight typed λ-calculi ordered by inclusion. The simplest of these calculi is Church's simply typed λ-calculus and the more complex is Coquand's Constructions [4]. The others systems include Girard's system F [12, 13], its higher order version F_ω, and the system LF [14], which is a variant of AUT-QE [6].

A complete description of the λ-cube, together with examples, may be found in [3, 2]. We just give briefly the main definitions.

The set $\mathcal{T}$ of raw types and of raw typed λ-terms is defined at once. This set is called the set of pseudo-terms.

Definition 19. *The set $\mathcal{T}$ of pseudo-terms is inductively defined as follows:*

i. $\square \in \mathcal{T}$,
ii. $* \in \mathcal{T}$,
iii. $x \in \mathcal{V} \Rightarrow x \in \mathcal{T}$,
iv. $x \in \mathcal{V}, M, N \in \mathcal{T} \Rightarrow (\lambda x\!:\!M.\,N) \in \mathcal{T}$,
v. $x \in \mathcal{V}, M, N \in \mathcal{T} \Rightarrow (\Pi x\!:\!M.\,N) \in \mathcal{T}$,
vi. $M, N \in \mathcal{T} \Rightarrow (M\,N) \in \mathcal{T}$.

The two constants $\square$ and $*$ are called sorts. We let s, s_1 and s_2 range over sorts. The statements are of the form $M : N$, with $M, N \in \mathcal{T}$. The notions of declaration and typing context are defined as previously. Type assignment is defined as follows.

Definition 20.

$$\textit{i.}\quad \vdash * : \square, \qquad\qquad \textit{ii.}\quad \frac{\Gamma \vdash M : s}{\Gamma, x : M \vdash x : M},$$

$$\textit{iii.}\quad \frac{\Gamma \vdash M : N \qquad \Gamma \vdash O : s}{\Gamma, x : O \vdash M : N}, \qquad \textit{iv.}\quad \frac{\Gamma \vdash M : (\Pi x\!:\!O.\,P) \qquad \Gamma \vdash N : O}{\Gamma \vdash (M\,N) : P[x\!:=\!N]},$$

$$\textit{v.}\quad \frac{\Gamma \vdash M : N \qquad \Gamma \vdash O : s}{\Gamma \vdash M : O} \quad \textit{if } N \twoheadleftarrow\!\!\twoheadrightarrow_\beta O,$$

$$\textit{vi.}\quad \frac{\Gamma \vdash M : s_1 \qquad \Gamma, x : M \vdash N : s_2}{\Gamma \vdash (\Pi x\!:\!M.\,N) : s_2},$$

$$\textit{vii.}\quad \frac{\Gamma \vdash M : s_1 \qquad \Gamma, x : M \vdash N : O \qquad \Gamma, x : M \vdash O : s_2}{\Gamma \vdash (\lambda x\!:\!M.\,N) : (\Pi x\!:\!M.\,O)}.$$

Rules vi and vii are parametrized by the pair of sorts (s_1, s_2). By taking this pair to be $(*, *)$, one gets a new formulation of Church's simply typed λ-calculus ($\lambda^{\rightarrow}$). To see this, one defines $(\alpha \rightarrow \beta)$ as $(\Pi x\!:\!\alpha.\,\beta)$ where the variable x does not occur in β. For a precise correspondence between this new formulation and Definition 6.1, see [3].

By taking (s_1, s_2) to be $(\square, *)$, $(\square, \square)$, or $(*, \square)$, one gets respectively *terms depending on types*, *types depending on types*, and *types depending on terms*. This corresponds to the three possible directions of the edges of the λ-cube.

Terms Depending on Types. By allowing for terms depending on types, one obtains systems containing Girard's system F. Therefore it is vain to seek for some arithmetizable proof [13]. The method of reducibility candidates [8] is somehow required.

This method, due to Girard, extends Tait's technique and yields strong normalization proofs. Therefore, it may be unclear how one can take advantage of Theorem 5 in this context. Nevertheless, when one is simply interested in normalization (as opposed to strong normalization), there exists a simplification of the method, which is due to Scedrov [18].

In the course of (one version of) the strong normalization proof of system F, one defines the notion of saturated set [8]. A saturated set S is a set of strongly normalizable λ-terms such that[1]:

1. if x is a variable and $M_1, \dots, M_n$ are strongly normalizable λ-terms, then $((x\, M_1)\, \dots M_n) \in S$,
2. if N_1 is a strongly normalizable λ-term and if $((M[x{:=}N_1]\, N_2)\, \dots N_n) \in S$, then $(((\lambda x.\, M\, N_1)\, N_2)\, \dots N_n) \in S$.

Scedrov notices that normalizable λ-terms, as opposed to strongly normalizable λ-terms, are closed under β-expansion. Therefore, when adapting the above definition to the case of normalization, one may drop Condition 2. The same idea may be used to establish $\beta_I\beta_S$-normalization.

Types Depending on Types. Our proof technique relies on the so-called erasing trick. To prove that a typed λ-term M is strongly normalizable, we prove that the untyped λ-term that is obtained by erasing the type information from M is strongly-normalizable.

Types depending on types introduces redexes at the type level. For instance, one may derive the following:

$$\begin{aligned} &\vdash (\Pi x{:}*.\, *) : \Box \\ &\vdash (\lambda a{:}*.\, a) : (\Pi x{:}*.\, *) \\ b : * &\vdash (\lambda x{:}((\lambda a{:}*.\, a)\, b).\, x) : (\Pi x{:}b.\, b) \end{aligned}$$

Nevertheless, the erasing trick may still be used because the strong normalization at the type level may be established independently of the strong normalization at the term level. This is even true for F_ω [8].

Types Depending on Terms. With types depending on terms, the erasing trick fails. Thus we cannot reason in the framework of untyped λ-calculus any more. We have to work with the complete set of pseudo-terms $\mathcal{T}$.

To adapt Theorem 5 to the set $\mathcal{T}$ is not straightforward. The problem is that even the usual version of the conservation theorem fails for $\mathcal{T}$. There exist pseudo-terms that are β_I-normalizable but not strongly β_I-normalizable as shown by the following counterexample:

$$((\lambda x{:}\Omega.\, x)\, y) \to_{\beta_I} y$$

where $\Omega \equiv ((\lambda x{:}y.\, x\, x)\, (\lambda x{:}y.\, x\, x))$.

Nevertheless, a version of Theorem 5 may be stated for pseudo-terms by following an idea due to Nederpelt [17]. This idea consists in adapting the notion of reduction β_I as follows.

[1] The conditions that we give are due to Mitchell. For a comparison between the different versions of Girard's method see the comprehensive article of Gallier [8].

$\beta_I^{\mathcal{T}}$: $((\lambda x{:}M.\, N)\, O) \to ((\lambda x{:}M.\, N[x{:=}O])\, O)$ if $x \in \mathrm{FV}(M)$.

On the other hand, our notion of reduction β_S may be kept unchanged (except for the syntax).

$\beta_S^{\mathcal{T}}$: $(((\lambda x{:}M.\, N)\, O)\, P) \to ((\lambda x{:}M.\, (N\, P))\, O)$ if $x \notin \mathrm{FV}(N)$.

With this new definitions, Theorem 5 holds for pseudo-terms.

7 Conclusions

The technique we have developed in this paper yields rather transparent proofs of strong normalization (transparency, of course, is chiefly a matter of style—or even a matter of taste).

As we already mentioned, the strong normalization proof that we have given for Church's simply typed λ-calculus is arithmetizable. Other arithmetizable proofs exist, among which the one by Gandy [10] is may be the best known. Gandy's proof is based on a semantic interpretation of the simply typed λ-terms and is, therefore, quite different form our proof. Nevertheless, it is interesting to note that he transforms β_K-redexes into β_I-redexes.

The untyped λ-terms typable in Nederpelt's calculus correspond exactly to the simply typable λ-terms [7]. Therefore Nederpelt's proof [17] may also be seen as an arithmetizable proof for Church's simply typed λ-calculus. As we explained in Section 6, our proof technique is close to the one of Nederpelt. The main difference is that Nederpelt does not use the notion of reduction β_S, but generalizes further the notion of reduction $\beta_I^{\mathcal{T}}$. This yields some technical complications when proving the Church-Rosser property. Another difference is that Nederpelt's proof is tailormade for his own calculus. Nevertheless, Nederpelt says in his thesis [17] that his proof may be turn in a general method of proving strong normalization from normalization. This statement is worked out by Klop in [16]. Indeed Klop provides a generalization of Nederpelt's method for a large class of reduction systems. The present work may also be seen as an exploration of Nederpelt's statement.

Van Daalen, in his thesis [19], gives also an arithmetizable proof of strong normalization for simply typed λ-calculus. His proof, which is totally syntactic, does not use any other notion of reduction than β, and is based on a strong substitution lemma.

We have briefly explained how to extend the strong normalization proof of simply typed λ-calculus to Barendregt's λ-cube. The ideas that we have developed are immediately applicable to the direct successors of $\lambda^{\to}$, namely $\lambda 2$, $\lambda\underline{\omega}$, and λP. To put these ideas together into a modular proof of strong normalization for the calculus of Constructions (in the spirit of [11]) will be the subject of future work.

References

1. H.P. Barendregt. *The lambda calculus, its syntax and semantics.* North-Holland, revised edition, 1984.
2. H.P. Barendregt. Introduction to Generalised Type Systems. *Journal of Functional Programming*, 1(2):125–154, 1991.
3. H.P. Barendregt. Lambda calculi with types. In S. Abramsky, D. Gabbai, and T. Maibaum, editors, *Handbook of Logic in Computer Science*. Oxford University Press, 1992.
4. Th. Coquand. Metamathematical investigations of a calculus of constructions. In P. Odifreddi, editor, *Logic and Computer Science*, pages 91–122. Academic Press, 1990.

5. N.G. de Bruijn. Lambda calculus notations with nameless dummies, a tool for automatic formula manipulation, with an application to the Church-Rosser theorem. *Indigationes Mathematicae*, 34:381–392, 1972.
6. N.G. de Bruijn. A survey of the project Automath. In J.P. Seldin and J.R. Hindley, editors, *to H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, pages 579–606. Academic Press, 1980.
7. Ph. de Groote. *Définition et Propriétés d'un métacalcul de représentation de théories.* Thèse de doctorat, Université Catholique de Louvain, Faculté des Sciences Appliquées, 1991.
8. J.H. Gallier. On Girard's "Candidats de Réductibilité". In P. Odifreddi, editor, *Logic and Computer Science*, pages 123–203. Academic Press, 1990.
9. R.O. Gandy. An early proof of normalization by A.M. Turing. In J. P. Seldin and J. R. Hindley, editors, *to H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, pages 453–455. Academic Press, 1980.
10. R.O. Gandy. Proofs of strong normalization. In J. P. Seldin and J. R. Hindley, editors, *to H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*, pages 457–478. Academic Press, 1980.
11. H. Geuvers and M.-J. Nederhof. Modular proof of strong normalization for the calculus of construction. *Journal of Functional Programming*, 1(2):155–189, 1991.
12. J.-Y. Girard. The system F of variable types, fifteen years later. *Theoretical Computer Science*, 45:159–192, 1986.
13. J.-Y. Girard, Y. Lafont, and P. Taylor. *Proofs and Types*, volume 7 of *Cambridge Tracts in Theoretical Computer Science.* Cambridge University Press, 1989.
14. R. Harper, F. Honsel, and G. Plotkin. A framework for defining logics. In *Proceedings of the second annual IEEE symposium on logic in computer science*, pages 194–204, 1987.
15. J.R. Hindley and J.P. Seldin. *Introduction to combinators and λ-calculus.* London Mathematical Society Student Texts. Cambridge University Press, 1986.
16. J.W. Klop. *Combinatory Reduction Systems.* PhD thesis, CWI, Amsterdam, Mathematical Centre Tracts Nr. 127, 1980.
17. R.P. Nederpelt. *Strong normalization in a typed lambda calculus with lambda structured types.* PhD thesis, Technische hogeschool Eindhoven, 1973.
18. A. Scedrov. Normalization revisited. In J.W. Gray and A. Scedrov, editors, *Proceedings of the AMS research conference*, pages 357–369. American Mathematical Society, 1987.
19. D.T. van Daalen. *The language theory of Automath.* PhD thesis, Technische hogeschool Eindhoven, 1980.

A Postponement of β_K-Contractions

Lemma 5. *Let $M, N, O \in \Lambda$ be such that $M \to_{\beta_K} N$. Then $M[x{:=}O] \to_{\beta_K} N[x{:=}O]$.*

Proof. By induction on the derivation of $M \to_{\beta_K} N$. □

Lemma 6. *Let $M, N, O \in \Lambda$ be such that $M \to_{\beta_K} N$. Then $O[x{:=}M] \twoheadrightarrow_{\beta_K} O[x{:=}N]$.*

Proof. By induction on the structure of O. □

Proof of Lemma 3.

We treat the case $R = \beta_I$, the other case being similar. The proof is by induction on the derivation of $M \to_{\beta_K} N$.

1. $M \equiv (\lambda x.\, M_1\, M_2)$, $N \equiv M_1$, and $x \notin \mathrm{FV}(M_1)$
   Take $P \equiv (\lambda x.\, O\, M_2)$.

2. $M \equiv \lambda x.\, M_1$, $N \equiv \lambda x.\, N_1$ and $M_1 \to_{\beta_K} N_1$.
   O must be of the form $\lambda x.\, O_1$, with $N_1 \to_{\beta_I} O_1$. Apply the induction hypothesis.
3. $M \equiv (M_1\, M_2)$, $N \equiv (N_1\, M_2)$ and $M_1 \to_{\beta_K} N_1$.
   There are three subcases according to the way $N \to_{\beta_I} O$.
   (a) $O \equiv (O_1\, M_2)$ and $N_1 \to_{\beta_I} O_1$.
   Apply the induction hypothesis.
   (b) $O \equiv (N_1\, O_2)$ and $M_2 \to_{\beta_I} O_2$.
   Take $P \equiv (M_1\, O_2)$.
   (c) $N_1 \equiv \lambda x.\, N_{11}$ and $O \equiv N_{11}[x{:=}M_2]$.
   There are two subcases according to the form of M_1.
      i. $M_1 \equiv \lambda x.\, M_{11}$ and $M_{11} \to_{\beta_K} N_{11}$.
      Take $P \equiv M_{11}[x{:=}M_2]$. Indeed, $M_{11}[x{:=}M_2] \to_{\beta_K} N_{11}[x{:=}M_2]$, by Lemma A.
      ii. $M_1 \equiv (\lambda y.\, \lambda x.\, N_{11}\, M_{11})$ and $y \notin \mathrm{FV}(\lambda x.\, N_{11})$.
      Take $P \equiv (\lambda y.\, N_{11}[x{:=}M_2]\, M_{11})$. Indeed we have:

$$\begin{aligned}
((\lambda y.\, \lambda x.\, N_{11}\, M_{11})\, M_2) &\to_{\beta_S} (\lambda y.\, (\lambda x.\, N_{11}\, M_2)\, M_{11}) \\
&\to_{\beta_I} (\lambda y.\, N_{11}[x{:=}M_2]\, M_{11}) \\
&\to_{\beta_K} N_{11}[x{:=}M_2].
\end{aligned}$$

4. $M \equiv (M_1\, M_2)$, $N \equiv (M_1\, N_2)$ and $M_2 \to_{\beta_K} N_2$.
   This case is similar to case (3), using Lemma A instead of Lemma A. □

B The Church-Rosser Property for β_I^+ and β_S^+

The Church-Rosser Property for β_S^+

The contraction of a β_S^+-redex does not multiply the other β_S^+-redexes. For this reason, the relation of β_S^+-contraction satisfies the Church-Rosser property.

Lemma 7. *Let $M, N, O \in \Lambda^{\mathbb{N}}$ be such that $M \to_{\beta_S^+} N$ and $M \to_{\beta_S^+} O$. Then there exists $P \in \Lambda^{\mathbb{N}}$ such that $N \to_{\beta_S^+} P$ and $O \to_{\beta_S^+} P$.*

Proof. By induction on the derivation of $M \to_{\beta_S^+} N$, distinguishing subcases according to the way $M \to_{\beta_S^+} O$. □

Lemma 8. *Let $M, N, O \in \Lambda^{\mathbb{N}}$ be such that $M \twoheadrightarrow_{\beta_S^+} N$ and $M \twoheadrightarrow_{\beta_S^+} O$. Then there exists $P \in \Lambda^{\mathbb{N}}$ such that $N \twoheadrightarrow_{\beta_S^+} P$ and $O \twoheadrightarrow_{\beta_S^+} P$.*

Proof. By a diagram chase, using Lemma B. □

The Church-Rosser Property for β_I^+

To prove that β_I^+ is Church-Rosser, we use the Tait–Martin-Löf method. We first define the binary relation $\underset{1}{\twoheadrightarrow}$ on $\Lambda^{\mathbb{N}}$.

Definition 21. *The binary relation $\underset{1}{\twoheadrightarrow}$ is defined on $\Lambda^{\mathbb{N}}$ by the following system.*

$$i. \quad M \underset{1}{\twoheadrightarrow} M,$$

$$ii.\quad \frac{M \underset{1}{\twoheadrightarrow} N}{\lambda x.\, M^n \underset{1}{\twoheadrightarrow} \lambda x.\, N^n}, \qquad iii.\quad \frac{M \underset{1}{\twoheadrightarrow} O \qquad N \underset{1}{\twoheadrightarrow} P}{(M\, N)^n \underset{1}{\twoheadrightarrow} (O\, P)^n},$$

$$iv.\quad \frac{M \underset{1}{\twoheadrightarrow} O \qquad N \underset{1}{\twoheadrightarrow} P}{(\lambda x.\, M^m\, N)^n \underset{1}{\twoheadrightarrow} (O[x{:=}P])^{+(m+n+1)}} \quad \text{if } x \in \mathrm{FV}(M).$$

Lemma 9. *Let $M, N \in \Lambda^{\mathbb{N}}$ be such that $M \underset{1}{\twoheadrightarrow} N$. Then $M^{+n} \underset{1}{\twoheadrightarrow} N^{+n}$.*

Proof. By induction on the derivation of $M \underset{1}{\twoheadrightarrow} N$. □

Lemma 10. *Let $M, N, O \in \Lambda^{\mathbb{N}}$ be such that $M \underset{1}{\twoheadrightarrow} N$. Then $O[x{:=}M] \underset{1}{\twoheadrightarrow} O[x{:=}N]$.*

Proof. By induction on the structure of O, using Lemma B when $O \equiv (x)^n$. □

Lemma 11. *Let $M, N, O, P \in \Lambda^{\mathbb{N}}$ be such that $M \underset{1}{\twoheadrightarrow} N$ and $O \underset{1}{\twoheadrightarrow} P$. Then $M[x{:=}O] \underset{1}{\twoheadrightarrow} N[x{:=}P]$.*

Proof. By induction on the derivation of $M \underset{1}{\twoheadrightarrow} N$, using Lemma B for the case $M \underset{1}{\twoheadrightarrow} M$. □

Lemma 12. *Let $M, N, O \in \Lambda^{\mathbb{N}}$ be such that $M \underset{1}{\twoheadrightarrow} N$ and $M \underset{1}{\twoheadrightarrow} O$. Then there exists $P \in \Lambda^{\mathbb{N}}$ such that $N \underset{1}{\twoheadrightarrow} P$ and $O \underset{1}{\twoheadrightarrow} P$.*

Proof. By induction on the derivation of $M \underset{1}{\twoheadrightarrow} N$.

1. $M \equiv N$.
   Take $P \equiv O$.
2. $M \equiv \lambda x.\, M_1^n$, $N \equiv \lambda x.\, N_1^n$, and $M_1 \underset{1}{\twoheadrightarrow} N_1$.
   We must have $O \equiv \lambda x.\, O_1^n$, with $M_1 \underset{1}{\twoheadrightarrow} O_1$. Apply the induction hypothesis.
3. $M \equiv (M_1\, M_2)^n$, $N \equiv (N_1\, N_2)^n$, $M_1 \underset{1}{\twoheadrightarrow} N_1$, and $M_2 \underset{1}{\twoheadrightarrow} N_2$.
   There are two subcases according to the way $M \underset{1}{\twoheadrightarrow} O$.
   (a) $O \equiv (O_1\, O_2)^n$, $M_1 \underset{1}{\twoheadrightarrow} O_1$, and $M_2 \underset{1}{\twoheadrightarrow} O_2$.
   Apply the induction hypothesis.
   (b) $M_1 \equiv \lambda x.\, M_{11}^m$, $O \equiv (O_{11}[x{:=}O_2])^{+(m+n+1)}$, $M_{11} \underset{1}{\twoheadrightarrow} O_{11}$, and $M_2 \underset{1}{\twoheadrightarrow} O_2$.
   Then we must have $N_1 \equiv \lambda x.\, N_{11}^m$, with $M_{11} \underset{1}{\twoheadrightarrow} N_{11}$. Therefore, by induction hypothesis, there exists $P_{11}, P_2 \in \Lambda^{\mathbb{N}}$ such that $N_{11} \underset{1}{\twoheadrightarrow} P_{11}$, $O_{11} \underset{1}{\twoheadrightarrow} P_{11}$, $N_2 \underset{1}{\twoheadrightarrow} P_2$ and $O_2 \underset{1}{\twoheadrightarrow} P_2$. Hence, by Lemmas B and B, we may take $P \equiv (P_{11}[x{:=}P_2])^{+(m+n+1)}$.
4. $M \equiv (\lambda x.\, M_1^m\, M_2)^n$, $N \equiv (N_1[x{:=}N_2])^{+(m+n+1)}$, $M_1 \underset{1}{\twoheadrightarrow} N_1$, and $M_2 \underset{1}{\twoheadrightarrow} N_2$.
   There are two subcases according to the way $M \underset{1}{\twoheadrightarrow} O$.
   (a) $O \equiv (\lambda x.\, O_1^m\, O_2)^n$, $M_1 \underset{1}{\twoheadrightarrow} O_1$, and $M_2 \underset{1}{\twoheadrightarrow} O_2$.
   This case is symmetric to Case (3.b).
   (b) $N \equiv (O_1[x{:=}O_2])^{+(m+n+1)}$, $M_1 \underset{1}{\twoheadrightarrow} O_1$, and $M_2 \underset{1}{\twoheadrightarrow} O_2$.
   Apply the induction hypothesis and use Lemmas B and B.

□

Lemma 13. *Let $M, N, O \in \Lambda^{\mathbb{N}}$ be such that $M \twoheadrightarrow_{\beta_I^+} N$ and $M \twoheadrightarrow_{\beta_I^+} O$. Then there exists $P \in \Lambda^{\mathbb{N}}$ such that $N \twoheadrightarrow_{\beta_I^+} P$ and $O \twoheadrightarrow_{\beta_I^+} P$.*

Proof. Because $\twoheadrightarrow_{\beta_I^+}$ is the transitive closure of $\underset{1}{\twoheadrightarrow}$. □

Commutation of β_I^+ and β_S^+

Lemma 14. *Let $M, N \in \Lambda^{\mathbb{N}}$ be such that $M \to_{\beta_S^+} N$. Then $M^{+n} \to_{\beta_S^+} N^{+n}$.*

Proof. By induction on the derivation of $M \to_{\beta_S^+} N$. □

Lemma 15. *Let $M, N, O \in \Lambda^{\mathbb{N}}$ be such that $M \to_{\beta_S^+} N$. Then $M[x{:=}O] \to_{\beta_S^+} N[x{:=}O]$.*

Proof. By induction on the derivation of $M \to_{\beta_S^+} N$. □

Lemma 16. *Let $M, N, O \in \Lambda^{\mathbb{N}}$ be such that $M \to_{\beta_S^+} N$. Then $O[x{:=}M] \to_{\beta_S^+} O[x{:=}N]$.*

Proof. By induction on the structure of O, using Lemma B for the case $O \equiv (x)^n$. □

To establish the commutation of β_I^+- and β_S^+-reductions, we prove what some authors call the trapezium property, by analogy with the diamond property.

Lemma 17. *Let $M, N, O \in \Lambda^{\mathbb{N}}$ be such that $M \to_{\beta_I^+} N$ and $M \to_{\beta_S^+} O$. Then there exists $P \in \Lambda^{\mathbb{N}}$ such that $N \twoheadrightarrow_{\beta_S^+} P$ and $O \to_{\beta_I^+} P$.*

Proof. By induction on the derivation of $M \to_{\beta_I^+} N$.

1. $M \equiv (\lambda x.\, M_1^m\, M_2)^n$ and $N \equiv (M_1[x{:=}M_2])^{+(m+n+1)}$.
   There are two subcases according to the way $M \to_{\beta_S^+} O$.
   (a) $O \equiv (\lambda x.\, O_1^m\, M_2)^n$ and $M_1 \to_{\beta_S^+} O_1$.
   By lemma B, we may take $P \equiv (O_1[x{:=}M_2])^{+(m+n+1)}$
   (b) $O \equiv (\lambda x.\, M_1^m\, O_2)^n$ and $M_2 \to_{\beta_S^+} O_2$.
   By lemma B, we may take $P \equiv (M_1[x{:=}O_2])^{+(m+n+1)}$
2. $M \equiv \lambda x.\, M_1^n$, $N \equiv \lambda x.\, N_1^n$ and $M_1 \to_{\beta_I^+} N_1$.
   O must be of the form $\lambda x.\, O_1^n$, with $M_1 \to_{\beta_S^+} O_1$. Apply the induction hypothesis.
3. $M \equiv (M_1\, M_2)^o$, $N \equiv (N_1\, M_2)^o$ and $M_1 \to_{\beta_I^+} N_1$.
   There are three subcases according to the way $M \to_{\beta_S^+} O$.
   (a) $O \equiv (O_1\, M_2)^o$ and $M_1 \to_{\beta_S^+} O_1$.
   Apply the induction hypothesis.
   (b) $O \equiv (M_1\, O_2)^o$ and $M_2 \to_{\beta_S^+} O_2$.
   Take $P \equiv (N_1\, O_2)^o$.
   (c) $M_1 \equiv (\lambda x.\, M_{11}^m M_{12})^n$, $x \notin \mathrm{FV}(M_{11})$, and $O \equiv (\lambda x.\, (M_{11} M_2) M_{12})^{m+n+o+1}$.
   We must have $N_1 \equiv (\lambda x.\, N_{11}^m\, N_{12})^n$ with $M_{11} \to_{\beta_I^+} N_{11}$ and $M_{12} \equiv N_{12}$, or with $M_{11} \equiv N_{11}$ and $M_{12} \to_{\beta_I^+} N_{12}$. Take $P \equiv (\lambda x.\, (N_{11}\, M_2)\, N_{12})^{m+n+o+1}$.
4. $M \equiv (M_1\, M_2)^o$, $N \equiv (M_1\, N_2)^o$ and $M_2 \to_{\beta_I^+} N_2$.
   This case is similar to Case (3).

□

Lemma 18. *Let $M, N, O \in \Lambda^{\mathbb{N}}$ be such that $M \twoheadrightarrow_{\beta_I^+} N$ and $M \twoheadrightarrow_{\beta_S^+} O$. Then there exists $P \in \Lambda^{\mathbb{N}}$ such that $N \twoheadrightarrow_{\beta_S^+} P$ and $O \twoheadrightarrow_{\beta_I^+} P$.*

Proof. By a diagram chase, using Lemma B. □

Proof of Theorem 4.

By Lemma B, Lemma B, and Lemma B. □