Development Separation in Lambda-Calculus

Hongwei Xi [1,2]

Computer Science Department
Boston University
Boston, Massachusetts, USA

Abstract

We present a proof technique in λ-calculus that can facilitate inductive reasoning on λ-terms by separating certain β-developments from other β-reductions. We give proofs based on this technique for several fundamental theorems in λ-calculus such as the Church-Rosser theorem, the standardization theorem, the conservation theorem and the normalization theorem. The appealing features of these proofs lie in their inductive styles and perspicuities.

Key words: λ-calculus, development, parallel reduction

1 Introduction

Proofs based on structural induction have certain desirable features. They usually enhance comprehensibility, yield more on the meaning of the proven theorems, and can be formalized relatively easily. Unfortunately, many theorems in λ-calculus cannot be proven via direct structural induction on λ-terms. This is mainly due to the fact that β-reduction is not compositional, that is, a β-reduction sequence from $M[x := N]$ usually cannot be viewed as the composition of some reduction sequences from M and N since β-redexes may be generated by substitution. This naturally raises an issue of separating newly generated β-redexes from the residuals of existing ones. Labeling β-redexes is one common approach to addressing such an issue. Explicit labeling is often studied through the formation of some labeled λ-calculi while implicit labeling can be built on top of the notion of residuals of β-redexes.

A labeled λ-calculus is introduced in [Hyl76] and [Wad76] as a tool for examining the λ-models such as D_∞ and P_ω. Also, a more general labeled λ-calculus is considered in [Lév], where the notion of strongly equivalent β-reduction is introduced. Many fundamental theorems in λ-calculus, such as

[1] The work is partly funded by NSF grant CCR-0229480
[2] Email: hwxi@cs.bu.edu

This is a preliminary version. The final version will be published in
Electronic Notes in Theoretical Computer Science
URL: www.elsevier.nl/locate/entcs

the Church-Rosser theorem and the standardization theorem, can be readily proven through the use of labeled β-reductions. This is largely due to the fact that labeled λ-reduction enjoys strong normalization, that is, there is no infinite labeled β-reduction sequence. However, too much labeling also shadows, sometimes, the meaning of proofs in certain cases. For instance, a proof of Church-Rosser theorem in the context of labeled λ-calculus often does not yield much on how to construct a commutative diagram. This is partly caused by the involvement of the strong normalization theorem for labeled β-reduction.

The notion of parallel β-reduction is introduced through a proof of the Church-Rosser theorem due to Tait and Martin-Löf (Section 3.2 [Bar84]). Parallel β-reductions are *complete* developments, in which only the existing β-redexes and their residuals are reduced. Since parallel β-reduction can be defined through structural induction on λ-terms, this makes it amenable to constructing inductive proofs. On the other hand, tracking residuals can, sometimes, help understand the proven theorems, but parallel β-reduction makes this difficult to do. Also, parallel β-reduction may have problems when reasoning on the length of β-reduction sequences is of the concern. For instance, the theorem on finiteness of developments is not easy to be formulated in terms of parallel β-reduction.

In this paper, we introduce a proof technique in λ-calculus that can facilitate inductive reasoning on λ-terms by separating certain developments from other β-reductions. After establishing a lemma on *development separation*, we are to give proofs for several fundamental theorems in λ-calculus such as the Church-Rosser theorem, the standardization theorem, the conservation theorem and the normalization theorem. The appealing features of these proofs lie in their inductive styles and perspicuities. Lastly, we also mention some closely related work.

We give preliminaries in Section 2 and then prove some basic properties on developments in Section 3. We present a proof of the Church-Rosser theorem in Section 4 and a proof of the standardization theorem in Section 5. We then include proofs of the conservation theorem and the normalization theorem in Section 6. Lastly, we mention some related work and conclude.

2 Preliminaries

We give a brief description on the notations and terminologies used in this paper. Most details that are not included here can be found in [Bar84].

Definition 2.1 [λ-terms] The set Λ of λ-terms is defined inductively as follows.

- (Variable) There are infinitely many variables $x, y, z, \ldots$ in Λ; variables are the only subterms of themselves.
- (Abstraction) If $M \in \Lambda$ then $\lambda x.M \in \Lambda$; N is a subterm of $\lambda x.M$ if N is

$\lambda x.M$ or N is a subterm of M,

- (Application) If $M_1, M_2 \in \Lambda$ then $M_1(M_2) \in \Lambda$; N is a subterm of $M_1(M_2)$ if N is $M_1(M_2)$ or N is a subterm of M_i for $i \in \{1, 2\}$.

We use $\mathbf{FV}(M)$ for the set of free variables in M, which is defined as usual. The set Λ_I consists of all λ-terms $M \in \Lambda$ such that $x \in \mathbf{FV}(M_0)$ whenever $\lambda x.M_0$ is a subterm of M. We write $M[x := N]$ for the result of substituting N for x in M, which is properly defined in [Bar84].

Definition 2.2 [β-redex and β-reduction] A β-redex R is a λ-term of the form $(\lambda x.M)(N)$, and the contractum of R is $M[x := N]$. A β-redex $(\lambda x.M)(N)$ is also a β_I-redex if $x \in \mathbf{FV}(M)$ holds. We write $M_1 \to_\beta M_2$ for a β-reduction step in which M_2 is obtained from replacing some β-redex in M_1 with its contractum and say that a λ-term M is in β-normal form if M contains no β-redexes.

Let $\to_\beta^n$ stand form n steps of β-reduction and $\to_\beta^*$ stands for some (possibly zero) steps of β-reduction. There may exist several different β-redexes in a λ-term M, and we say that a β-redex R_1 in M is to the left of another β-redex R_2 in M if the first symbol of R_1 is to the left of the first symbol of R_2.

Definition 2.3 [β-reduction sequence] Given a β-redex R in M, we write $M \xrightarrow{R}_\beta N$ for the β-reduction step in which β-redex R gets contracted. A (possibly) infinite β-reduction sequence is written as follows:

$$M_1 \xrightarrow{R_1}_\beta M_2 \xrightarrow{R_2}_\beta \cdots$$

Also, we write $[R_1] + [R_2] + \cdots + [R_n]$ for a β-reduction sequence of the following form:

$$M_1 \xrightarrow{R_1}_\beta M_2 \xrightarrow{R_2}_\beta \cdots \xrightarrow{R_n}_\beta M_{n+1}$$

We use $\sigma, \tau, \ldots$ for finite β-reduction sequences.

Notations We use $\emptyset$ for the empty β-reduction sequence, and $\sigma : M \to_\beta^* N$ or $M \xrightarrow{\sigma}_\beta^* N$ for a β-reduction sequence from M to N, and $\sigma(M)$ for the λ-term to which M is reduced by σ, and $|\sigma|$ for the length of σ, that is, the number of β-reduction steps involved in σ. Given $\sigma_1 : M_1 \to_\beta^* M_2$ and $\sigma_2 : M_2 :\to_\beta^* M_3$, we write $\sigma_1 + \sigma_2$ for the concatenation of σ_1 and σ_2, that is, $\sigma_1 + \sigma_2 : M_1 \xrightarrow{\sigma_1}_\beta^* M_2 \xrightarrow{\sigma_2}_\beta^* M_3$.

We need some conventions to make our notations more convenient; given a β-reduction sequence $\sigma : M \to_\beta^* M_2$ and a context $C[]$, $C[\sigma]$ is for the β-reduction sequence from $C[M_1]$ to $C[M_2]$ induced by σ in the evident manner, and we may often use σ for $C[\sigma]$ if the context $C[]$ can be readily inferred. An immediate consequence of this convention is that $\sigma_1 + \sigma_2$ can be regarded as $C_1[\sigma_1] + C_2[\sigma_2]$ for some proper contexts $C_1[]$ and $C_2[]$. Also $\sigma[x := N]$

stands for the β-reduction sequence obtained by substituting N for each free occurrences of x in σ.

We now introduce the concept of *residuals* of β-redexes, which is rigorously defined in [Hue94]. Let $\mathcal{R}$ be a set of β-redexes in a λ-term M_1 and $M_1 \xrightarrow{R}_\beta M_2$ holds for some β-redex $R = (\lambda x.M)(N)$. This reduction step affects the β-redexes R' in $\mathcal{R}$ in the following ways:

- R' is R. Then R' has no residual in M_2.
- R' is in N. Then all copies of R' in $M[x := N]$ are residuals of R' in M_2.
- R' is in M. Then the λ-term $R'[x := N]$ in M_2 is the residual of R'.
- R' contains R. Then the residual of R' in M_2 is the λ-term obtained from replacing R in R' with its contractum.
- R' and R are disjoint. Then R' is not affected and is its own residual.

The residual relation is transitive. We denote by $\mathcal{R}/[R]$ the set of residuals of all β-redexes in $\mathcal{R}$ after R is contracted. In general, $\mathcal{R}/\sigma$ is defined to be $\mathcal{R}/[R_1]/\ldots/[R_n]$, where $\sigma = [R_1] + \ldots + [R_n]$.

Definition 2.4 [Involvement] Given $M \to_\beta^* M_1$ and $\sigma : M_1 \to_\beta M_2 \to_\beta \cdots$, a β-redex R in M is involved in σ if a residual of R is contracted in σ.

3 Developments

Definition 3.1 [Developments] Given a λ-term M and a set $\mathcal{R}$ of β-redexes in M, a β-reduction sequence σ from M is a development of $\mathcal{R}$ if for every R contracted in σ, $R \in \mathcal{R}$ is a residual of some β-redex in $\mathcal{R}$.

Clearly, if $[R] + \sigma$ is a development of $\mathcal{R}$, then σ is a development of $\mathcal{R}/[R]$. An alternative way of defining developments is through labeling as is done below.

Definition 3.2 [λ_0-terms] The set Λ_0 of λ_0-terms is defined as follows.

- (Variable) There are infinitely many variables $x, y, z, \ldots$ in Λ_0.
- (Abstraction) If $M \in \Lambda_0$ then $\lambda x.M \in \Lambda_0$
- (Application) If $M_1, M_2 \in \Lambda_0$ then $M_1(M_2) \in \Lambda_0$.
- (β_0-redex) if $M_1, M_2 \in \Lambda_0$ then $(\lambda_0 x.M_1)(M_2) \in \Lambda_0$.

Intuitively, given a β-redex $R = (\lambda x.M)(N)$, we can mark R to obtain a β_0-redex $R_0 = (\lambda_0 x.M)(N)$. Given a λ-term and a set $\mathcal{R}$ of β-redexes in M, $M_{\mathcal{R}}$ is the λ_0-term obtained from marking all β-redexes in $\mathcal{R}$. We may use M for $M_{\mathcal{R}}$ if there is no risk of confusion.

Definition 3.3 [β_0-reduction] Given a β_0-redex $R_0 = (\lambda_0 x.M)(N)$, the contractum of R_0 is $M[x := N]$. We write $M \to_{\beta_0} N$ for a β-reduction step that replaces some β_0-redex in M with its contractum to obtain N.

Note that β_0-terms are closed under β_0-reductions. We present another definition of developments as follows.

Definition 3.4 [Developments] Given a λ-term M and a set $\mathcal{R}$ of β-redexes in M, a development of $\mathcal{R}$ from M is obtained from erasing all the labels, that is, turning all occurrences of λ_0 into λ, in a β_0-reduction sequence starting from $M_{\mathcal{R}}$.

Definition 3.5 [Canonical and Standard Developments] Let $\sigma = [R_1] + \ldots + [R_n]$ be a development of $\mathcal{R}$. If R_j are not residuals of any β-redex containing R_i for all $1 \leq i < j \leq n$, then σ is a canonical development. Furthermore, if R_j are not residuals of any β-redex to the left of R_i for all $1 \leq i < j \leq n$, then σ is a standard development.

Clearly, a standard development is canonical but not necessarily vice versa. Let $\sigma = [R_1] + \ldots + [R_n]$ be a canonical development. If σ is not standard, then there exists $1 \leq i < n$ such that R_{i+1} is the residual of some β-redex to the left of R_i; since σ is canonical, R cannot contain R_i; this implies that R is disjoint from R_i; then the conflict can be resolved if the reduction order of R_i and R_{i+1} is reversed; in this way, a canonical development can be permuted into a standard one.

Notations Given a λ-term M, $M[x \ldots, x]_x$ is a representation of M in which all the free occurrences of x in M are enumerated from left to right in $[x, \ldots, x]$. We write $M[N_1, \ldots, N_n]_x$ for the λ-term obtained by substituting N_i for the ith free occurrence of x in M for $i = 1, \ldots, n$.

Given a β-reduction sequence σ from M, we can simply construct a corresponding β-reduction sequence σ' from $M[N_1, \ldots, N_n]_x$ by treating N_i as if there were the variable x. It is easy to observe that $\sigma'(M[N_1, \ldots, N_n])$ is of the form $\sigma(M)[N'_1, \ldots, N'_{n'}]$, where every N'_j is some N_i for $1 \leq i \leq n$. This observation will be used in the proof of the next lemma.

Lemma 3.6 (Development Separation) *Let $M = (\lambda x.M_1)(M_2)$, $\mathcal{R}_1$ be a set of β-redexes in M_1, $\mathcal{R}_2$ be a set of β-redexes in M_2 and $\mathcal{R} = \{M\} \cup \mathcal{R}_1 \cup \mathcal{R}_2$. For every development σ of $\mathcal{R}$ in which M is involved, $\sigma(M)$ is of the following form:*

$$\sigma_1(M_1)[\sigma_{2,1}(M_2), \ldots, \sigma_{2,n}(M_2)]_x,$$

where σ_1 is a development of $\mathcal{R}_1$ from M_1 and $\sigma_{2,i}$ are developments of $\mathcal{R}_2$ from M_2 for $i = 1 \leq i \leq n$.

Proof. Since M is involved in σ, σ must be of the following form:

$$(\lambda x.M_1)(M_2) \xrightarrow{\tau_1}{}^{*}_{\beta} (\lambda x.M'_1)(M'_2) \to_\beta M'_1[x := M'_2] \xrightarrow{\tau_2}{}^{*}_{\beta} \sigma(M)$$

Clearly, we may assume $\tau_1 = \tau_{1,1} + \tau_{1,2}$, where $M_1 \xrightarrow{\tau_{1,1}}{}^{*}_{\beta} M'_1$ is a development of $\mathcal{R}_1$ and $M_2 \xrightarrow{\tau_{1,2}}_\beta^* M_2'$ is a development of $\mathcal{R}_2$. Now let us verify by induction on the length of τ_2 that $\sigma(M) = \tau_2(M_1'[x := M_2'])$ is of the given form.

- $\tau_2 = \emptyset$. Then $\sigma(M) = \tau_{1,1}(M_1)[\tau_{1,2}(M_2), \ldots, \tau_{1,2}(M_2)]_x$ is of the given form.
- $\tau_2 = \tau_2' + [R]$. By induction hypothesis, $\tau_2'(M_1'[x := M_2'])$ is of the following form:

$$\sigma_1'(M_1)[\sigma_{2,1}'(M_2), \ldots, \sigma_{2,n'}'(M_2)]_x$$

where σ_1' is a development of $\mathcal{R}_1$ from M_1 and $\sigma_{2,1}', \ldots, \sigma_{2,n'}'$ are developments of $\mathcal{R}_2$ from M_2. Now we have two subcases as follows.

· R is a residual of some β-redex in $\mathcal{R}_2$. Then R is in some $\sigma_{2,i}'(M_2)$, and therefore $\tau_2(M_1'[x := M_2'])$ is of the form:

$$\sigma_1'(M_1)[\sigma_{2,1}'(M_2), \ldots, (\sigma_{2,i}' + [R])(M_2), \ldots, \sigma_{2,n'}'(M_2)]_x$$

· R is a residual of some β-redex R_1 in $\mathcal{R}_1$. Then there exists a residual R_1' of R_1 in $\sigma_1'(M_1)$ such that $R = R_1'[N_1', \ldots, N_k']_x$, where every N_i' is some $\sigma_{2,j}'(M_2)$. Hence, with the previous observation, $\tau_2(M_1'[x := M_2'])$ is of the form:

$$(\sigma_1' + [R_1'])(M_1)[N_1', \ldots, N_{k'}']_x$$

where each N_i' is some $\sigma_{2,j}'(M_2)$.

Therefore, $\sigma(M) = \tau(M_1'[x := M_2'])$ is of the given form.

This is a constructive proof. Hence, we can use **sep**(σ) for the following β-reduction sequence:

$$[M] + \sigma_1[x := M_2] + \sigma_{2,1} + \ldots + \sigma_{2,n}$$

It can be readily verified that for each $R \in \mathcal{R}$, $\mathcal{R}$ is involved in σ if it is involved in **sep**(σ). □

The idea of development separation can also be found in [Hin78]. To illustrate this point, we present a proof of the *finiteness of developments* (FD) in Hindley's style, though some minor changes are made here. We define the size of a λ-term as follows.

$$|x| = 1 \qquad |\lambda x.M| = |M| \qquad |M_1(M_2)| = |M_1| + |M_2|$$

Lemma 3.7 *For each development σ from M, we have $|\sigma(M)| \leq 2^{|M|}$.*

Proof. With Lemma 3.6, the proof immediately follows from structural induction on M. □

Theorem 3.8 (Finiteness of Developments) *Given a λ-term M, we have $|\sigma| < 2^{|M|}$ for every development σ from M.*

Proof. We proceed by structural induction on M.

- $M = x$. Then this case is trivial.

- $M = \lambda x.M_0$. Then this case follows from induction hypothesis on M_0 immediately.
- $M = M_1(M_2)$ and M is not a β-redex. Then we may assume that $\sigma = \sigma_1 + \sigma_2$, where σ_1 and σ_2 are developments from M_1 and M_2, respectively. By induction hypotheses on M_1 and M_2, we have

$$|\sigma| = |\sigma_1| + |\sigma_2| < 2^{|M_1|} + 2^{|M_2|} \leq 2^{|M_1|+|M_2|} = 2^{|M|}$$

- $M = (\lambda x.M_1)(M_2)$. If M as a β-redex is not involved in σ, the case is the same as the previous one. Let us assume that M is involved in σ. By Lemma 3.6, $\sigma(M)$ is of the following form:

$$\sigma_1(M_1)[\sigma_{2,1}(M_2), \ldots, \sigma_{2,n}(M_2)]_x$$

  where σ_1 is a development from M_1 and $\sigma_{2,i}$ are developments from M_2 for $1 \leq i \leq n$. By induction hypothesis on M_2, $|\sigma_{2,i}(M_2)| \leq 2^{|M_2|}$ for $1 \leq i \leq n$. By Lemma 3.7, there are at most 2^{M_1} occurrences of x in σ_1. It can then be readily verified that

$$|\sigma| \leq |\sigma_1| + 2^{|M_1|}|\sigma_2| \leq 2^{M_1} - 1 + 2^{|M_1|}(2^{|M_2|} - 1) < 2^{|M|}$$

All the cases are completed now. □

A proof of FD due to Hyland [Hyl73] can yield the same bound. Also, a proof due to de Vrijer [dV85] gives an exact bound for the lengths of developments from a given λ-term. As is stated in [Bar84], there exists a real number $\alpha > 0$ such that one can find a sequence of λ-terms $M_1, M_2, \ldots$ with $\mu_0(M_i) \geq 2^{\alpha|M_i|}$ for $i \geq 1$ and $\lim_n |M_n| = \infty$, where $\mu_0(M)$ measures the length of a longest development from M.

Next we show that every development σ can be transformed into a standard development $\mathbf{std}(\sigma)$ such that if a β-redex is involved in $\mathbf{std}(\sigma)$ then it is involved in σ.

Lemma 3.9 (Standardization of Developments) *For every development $\sigma : M \rightarrow^*_\beta N$ of $\mathcal{R}$, there exists a standard development $\mathbf{std}(\sigma) : M \rightarrow^*_\beta N$ such that for every $R \in \mathcal{R}$, R is involved in σ if it is involved in $\mathbf{std}(\sigma)$.*

Proof. Since a canonical development can be readily permuted into a standard development, it suffices to show that there exists a canonical development $\mathbf{cad}(\sigma) : M \rightarrow^*_\beta N$ of $\mathcal{R}$ such that for each β-redex $R \in \mathcal{R}$, R is involved in σ if it is involved in $\mathbf{cad}(\sigma)$. Let us proceed by structural induction on M.

- M is a variable. Then $\sigma = \emptyset$ is canonical.
- $M = \lambda x.M_0$. Then this case simply follows from the induction hypothesis on M_0.
- $M = M_1(M_2)$, where M is not a β-redex. Then we can assume that σ is of the form $\sigma_1 + \sigma_2$, where σ_i are developments from M_i for $i = 1, 2$. By induction hypothesis, $\mathbf{cad}(\sigma_i)$ are defined for $i = 1, 2$. Let $\mathbf{cad}(\sigma)$ be $\mathbf{cad}(\sigma_1) + \mathbf{cad}(\sigma_2)$, which is a canonical development from M by definition.

Assume that $R \in \mathcal{R}$ is involved in $\mathbf{cad}(\sigma)$. Then R is involved in $\mathbf{cad}(\sigma_p)$ for $p = 1$ or $p = 2$. By induction hypothesis on σ_p, R is involved in σ_p and thus it is involved in σ.

- $M = (\lambda x.M_1)(M_2)$. If M is not involved in σ, then this case is the same as the previous one. We now assume that M is involved in σ. By Lemma 3.6, $\mathbf{sep}(\sigma)$ is of the following form:

$$[M] + \sigma_1[x := N] + \sigma_{2,1} + \ldots + \sigma_{2,n}$$

where σ_1 is a development from M_1 and $\sigma_{2,i}$ are developments from M_2 for $1 \leq i \leq n$. By induction hypothesis, we can defined $\mathbf{cad}(\sigma)$ as follows:

$$\mathbf{cad}(\sigma) = [M] + \mathbf{cad}(\sigma_1)[x := N] + \mathbf{cad}(\sigma_{2,1}) + \ldots + \mathbf{cad}(\sigma_{2,n})$$

Clearly, $\mathbf{cad}(\sigma)$ is canonical since both $\mathbf{cad}(\sigma_1)$ and $\mathbf{cad}(\sigma_{2,i})$ are canonical for $1 \leq i \leq n$. Assume that $R \in \mathcal{R}$ is involved in $\mathbf{cad}(\sigma)$. Then it can be readily verified that R is involved in $\mathbf{sep}(\sigma)$. Hence, R is also involved in σ.

For each development σ, we can permute $\mathbf{cad}(\sigma)$ into a standard development $\mathbf{std}(\sigma)$. □

Given a development σ of $\mathcal{R}$, if $\mathcal{R}/\sigma = \emptyset$, then σ is a *complete* development of $\mathcal{R}$. One step of parallel β-reduction (Section 3.2 [Bar84]) can be regarded as a complete development (of some $\mathcal{R}$).

4 Church-Rosser Theorem

The Church-Rosser theorem (CR) was first proven in [CR36], and many other proofs have been published since then. One approach to proving CR is to first prove a so-called *strip* lemma and then carry out induction on the length of β-reduction sequences. The theorem FD is often employed in a proof of the strip lemma, which may make reasoning less perspicuous since many β-redexes are unnecessarily reduced when FD is applied. In the following proof of CR, we spare the use of FD, trying to bring out a clearer picture.

Lemma 4.1 (CR of Developments) *Given a pair of developments $\langle\sigma_1, \sigma_2\rangle$ from M, we can construct another pair of developments $\mathbf{cr}(\langle\sigma_1, \sigma_2\rangle) = \langle\tau_1, \tau_2\rangle$ such that $(\sigma_1 + \tau_1)(M) = (\sigma_2 + \tau_2)(M)$.*

Proof. Let us define $\mathbf{cr}(\langle\sigma_1, \sigma_2\rangle)$ by structural induction on M.

- M is a variable. Then $\sigma_1 = \sigma_2 = \emptyset$. Let $\mathbf{cr}(\langle\sigma_1, \sigma_2\rangle) = \langle\emptyset, \emptyset\rangle$.
- $M = \lambda x.M_0$. This case follows from induction hypothesis straightforwardly.
- $M = M_1(M_2)$, where M is not a β-redex. Then we can assume that $\sigma_i = \sigma_{i,1} + \sigma_{i,2}$ for $i = 1, 2$, where $\sigma_{i,1}$ and $\sigma_{i,2}$ are developments from M_1 and M_2, respectively. Let $\langle\tau_{i,1}, \tau_{i,2}\rangle = \mathbf{cr}(\langle\sigma_{i,1}, \sigma_{i,2}\rangle)$ for $i = 1, 2$. Clearly, $\mathbf{cr}(\langle\sigma_1, \sigma_2\rangle)$ can be defined as follows:

$$\mathbf{cr}(\langle\sigma_1, \sigma_2\rangle) = \langle\tau_{1,1} + \tau_{1,2}, \tau_{2,1} + \tau_{2,2}\rangle$$

- $M = (\lambda x.M_1)(M_2)$. We may assume that M is involved in σ_p for some $p \in \{1,2\}$ as otherwise the case is the same as the previous one. If M is not involved in σ_q for $q \neq p \in \{1,2\}$, then we can replace σ_q with $\sigma_q + [\sigma_q(M)]$. By Lemma 3.6, $\sigma_i(M)$ are of the following forms:

$$\sigma_{i,1}(M_1)[\sigma_{i,2}^1(M_2), \ldots, \sigma_{i,2}^{n_i}(M_2)]_x$$

for $i = 1, 2$, where $\sigma_{i,1}$ are developments from M_1 and $\sigma_{i,2}^1, \ldots, \sigma_{i,2}^{n_i}$ are developments from M_2. By induction hypothesis, we can assume $\langle \tau_{1,1}, \tau_{2,1} \rangle = \mathbf{cr}(\langle \sigma_{1,1}, \sigma_{2,1} \rangle)$. Thus, we can construct $\tau_{i,1}^*$ corresponding to $\tau_{i,1}$, reducing $\sigma_i(M)$ to the following forms

$$(\sigma_{i,1} + \tau_{i,1})(M_1)[M_{i,2}^1, \ldots, M_{i,2}^n]_x$$

for $i = 1, 2$ and some n, where each $M_{i,2}^j$ $(j = 1, \ldots, n)$ is $\sigma_{i,2}^k(M_2)$ for some $k = k(i, j)$. By induction hypothesis, we can assume:

$$\langle \tau_{1,2}^j, \tau_{2,2}^j \rangle = \mathbf{cr}(\langle \sigma_{1,2}^{k(1,j)}, \sigma_{2,2}^{k(2,j)} \rangle)$$

for $i = 1, 2$ and $j = 1, \ldots, n$. Let $\mathbf{cr}(\langle \sigma_1, \sigma_2 \rangle)$ be defined as follows:

$$\langle \tau_{1,1}^* + \tau_{1,2}^1 + \ldots + \tau_{1,2}^n, \tau_{2,1}^* + \tau_{2,2}^1 + \ldots + \tau_{2,2}^n \rangle$$

It can be readily verified that this definition suffices.

We conclude the proof as all cases are completed. □

Theorem 4.2 (CR) *Given two β-reduction sequences $\sigma_1 : M \to_\beta^* M_1$ and $\sigma_2 : M \to_\beta^* M_2$, there exist τ_1 and τ_2 such that $(\sigma_1 + \tau_1)(M) = (\sigma_2 + \tau_2)(M)$.*

Proof. The theorem follows immediately from Lemma 4.1 since $\to_\beta^*$ is a transitive closure of developments. □

This proof of CR is closely related to one in [Bar84] due to Tait and Martin-Löf, where the notion of parallel β-reduction is introduced. In both cases, the need for FD is spared and some structural induction on λ-terms is employed. With Lemma 3.6, our proof exhibits an illustrating picture on why CR holds in λ-calculus, which seems to be somewhat hidden in the proof due to Tait and Martin-Löf.

5 Standardization Theorem

The standardization theorem was first proven in [CF58], stating that every β-reduction sequence can be standardized in the sense given by the following definition:

Definition 5.1 [Standardization of β-reduction sequences] Given a β-reduction sequence σ of the following form:

$$M_1 \xrightarrow{R_1}_\beta M_2 \xrightarrow{R_2}_\beta \cdots \xrightarrow{R_n}_\beta M_{n+1}$$

we say that σ is standard if for all $1 \leq i < j \leq n$, R_j is not a residual of some β-redex to the left of R_i. We say that $\sigma_s : M \rightarrow^*_\beta N$ standardizes σ if σ_s is a standard β-reduction sequence and for every R in M that is involved in σ_s, R is also involved in σ.

We now prove that for every β-reduction sequence σ, there exists σ_s that standardizes σ.

Lemma 5.2 *Given $\sigma = \sigma_1 + \sigma_2$, where σ_1 is a standard development of $\mathcal{R}$ and σ_2 is a standard β-reduction sequence, we can construct a β-reduction sequence $\mathbf{std}_2(\sigma_1, \sigma_2)$ which standardizes σ.*

Proof. By Lemma 3.9, the function **std** is defined on all developments. Let us define $\mathbf{std}_2(\sigma_1, \sigma_2)$ and prove that $\mathbf{std}_2(\sigma_1, \sigma_2)$ standardizes $\sigma_1 + \sigma_2$ by induction on $\langle |\sigma_2|, |\sigma_1| \rangle$, lexicographically ordered. Clearly, for σ_1, σ_2, $\mathbf{std}(\sigma_1, \emptyset)$ and $\mathbf{std}(\emptyset, \sigma_2)$ can be defined as σ_1 and σ_2, respectively. We now assume $\sigma_1 = [R_1] + \sigma'_1$ and $\sigma_2 = [R_2] + \sigma'_2$, and we have two cases.

- R_2 is a residual of some β-redex in $\mathcal{R}$ that is to the left of R_1. Hence, $\sigma_1 + [R_2]$ is a development. We define $\mathbf{std}_2(\sigma_1, \sigma_2)$ as follows:

$$\mathbf{std}_2(\sigma_1, \sigma_2) = \mathbf{std}_2(\mathbf{std}(\sigma_1 + [R_2]), \sigma'_2)$$

  Assume that $R \in \mathcal{R}$ is involved in $\mathbf{std}_2(\sigma_1, \sigma_2)$. Then by induction hypothesis, R is involved in $\mathbf{std}(\sigma_1 + [R_2]) + \sigma'_2$. This implies that R is involved in $\sigma_1 + [R_2] + \sigma'_2 = \sigma_1 + \sigma_2 = \sigma$.

- R_2 is not a residual of any β-redex in $\mathcal{R}$ that is to the left of R_1. Then we define $\mathbf{std}_2(\sigma_1, \sigma_2)$ as follows:

$$\mathbf{std}_2(\sigma_1, \sigma_2) = [R_1] + \mathbf{std}_2(\sigma'_1, \sigma_2)$$

Assume that R is a β-redex to the left of R_1. Then R is not involved in σ_1 as σ_1 is standard. Then it can be readily verified that R has no residual that is to the right of R_2. Note that R_2 is not a residual of any β-redex in $\mathcal{R}$. This implies that R is not involved in σ_2. It is now straightforward to see that $\mathbf{std}(\sigma_1, \sigma_2)$ is standard. □

Theorem 5.3 (Standardization of β-reduction sequences) *For every β-reduction sequence σ, we can construct a β-reduction sequence $\mathbf{std}_1(\sigma)$ that standardizes σ.*

Proof. Let us define $\mathbf{std}_1$ as follows:

$$\mathbf{std}_1(\emptyset) = \emptyset \qquad \mathbf{std}_1([R] + \sigma) = \mathbf{std}_2([R], \mathbf{std}_1(\sigma))$$

By Lemma 5.2, $\mathbf{std}_1(\sigma)$ standardizes σ. □

The key idea of this proof is to repeatedly shift the leftmost involved β-redex in a β-reduction sequence to the front. Though this is also the idea in a proof presented in [Klo80], we use a different strategy to prove the termination of the process. With Lemma 3.9, our proof not only obviates the need for FD but also presents a sharp inductive argument on why the shifting process terminates. Another advantage of our proof is that it can be readily modified to generate a bound for the length of the standardized β-reduction sequence based on the length of the original β-reduction sequence [Xi99].

6 Conservation and Normalization Theorems

In this section, we present inductive proofs for the conservation theorem and the normalization theorem in λ-calculus.

Definition 6.1 Given a λ-term M, M is strongly β-normalizing if there exists no infinite β-reduction sequence from M.

If M is strongly normalizing, let $\mu(M)$ be the least natural number such that $|\sigma| \leq \mu(M)$ holds for each β-reduction sequence from M. Otherwise, let $\mu(M) = \infty$.

Lemma 6.2 *Assume $M \xrightarrow{R}_\beta M'$, where $R = (\lambda x.N_1)(N_2)$ is the leftmost β-redex in M. Then $\mu(M) \leq \mu(M') + \mu(N_2)$ holds.*

Proof. Please see the proof of Lemma 13.2.5(i) in [Bar84]. □

Lemma 6.3 *Assume $\sigma : M \rightarrow^*_\beta M'$ is a standard development of $\mathcal{R}$, which contains only β_I-redexes. Then $\mu(M') < \infty$ implies $\mu(M) < \infty$.*

Proof. Let us proceed by induction on $\langle \mu(M'), |\sigma| \rangle$, lexicographically ordered. If M is in β-normal form, then we are done as $M' = M$. We now assume that $M \xrightarrow{R_l}_\beta M_l$, where $R_l = (\lambda x.N_1)(N_2)$ is the leftmost β-redex in M.

- R_l is involved in σ. Since σ is standard, we have $\sigma : M \xrightarrow{R_l}_\beta M_l \xrightarrow{\sigma'}^*_\beta M'$ for some σ'. By induction hypothesis, $\mu(M_l) < \infty$ holds since $|\sigma'| < |\sigma|$. Note that R is a β_I-redex in this case. Hence $\mu(N_2) < \infty$ as N_2 is a subterm of M_l. By Lemma 6.2, we have $\mu(M) < \infty$.
- R_l is not involved in σ. Hence, R_l has a residual $R'_l = (\lambda x.N'_1)(N'_2)$ in M', which also happens to be the leftmost β-redex in M'. Clearly σ is of the form $\sigma_1 + \sigma_2 + \sigma_3$, where $\sigma_1 : N_1 \rightarrow^*_\beta N'_1$ and $\sigma_2 : N_2 \rightarrow^*_\beta N'_2$ are standard developments and σ_3 is also a standard development. Since $|\sigma_2| \leq |\sigma|$ holds and $\mu(N'_2) < \mu(M')$, we have $\mu(N_2) < \infty$ by induction hypothesis. Assume $M' \xrightarrow{R'_l}_\beta M'_l$. Then $\sigma + [R'_l]$ is a development of $\mathcal{R} \cup \{R_l\}$. Therefore, $\mathbf{std}(\sigma + [R'_l]) = R_l + \sigma'$ for some standard development of $\mathcal{R}/[R_l]$. It can be immediately verified that $\mathcal{R}/[R_l]$ is a set of β_I-redexes. Since $\sigma' : M_l \rightarrow^*_\beta M'_l$ and $\mu(M'_l) < \mu(M)$, we have $\mu(M_l) < \infty$ by induction hypothesis. This yields $\mu(M) < \infty$ by Lemma 6.2.

□

Theorem 6.4 (Conservation) *Assume $M \xrightarrow{R}_\beta M'$ for some β_I-redex R. Then $\mu(M') < \infty$ implies $\mu(M) < \infty$.*

Proof. This follows from Lemma 6.3 since $M \xrightarrow{R}_\beta M'$ is obviously a standard development of a β_I-redex. □

The normalization theorem in λ-calculus follows from the standardization theorem immediately. However, in some settings such as the call-by-value λ-calculus λ_v [Plo75], it seems rather involved to establish a version of standardization theorem. This makes it desirable to prove the normalization theorem in the following style.

Given a λ-term M, let $\Lambda(M)$ be the longest leftmost β-reduction sequence from M, which may be of infinite length.

Lemma 6.5 *Assume that* $\sigma : M \rightarrow^*_\beta M'$ *is a standard development. If* $|\Lambda(M')| < \infty$*, then* $|\Lambda(M')| \leq |\Lambda(M)| < \infty$ *holds.*

Proof. The proof proceeds by induction on $\langle |\Lambda(M')|, |\sigma| \rangle$, lexicographically ordered. If M is in β-normal form, then $M' = M$ and we are done. We now assume $M \xrightarrow{R_l}_\beta M_l$, where R_l is the leftmost β-redex in M. Then $\Lambda(M) = [R_l] + \Lambda(M_l)$. We have two cases as follows.

- R_l is involved in σ. Since σ is standard, σ is of the form $M \xrightarrow{R_l}_\beta M_l \xrightarrow{\sigma'}^*_\beta M'$ for some standard development σ'. Since $|\sigma'| < |\sigma|$ holds, we have $|\Lambda(M')| \leq |\Lambda(M_l)| < \infty$ by induction hypothesis. Hence $|\Lambda(M')| \leq |\Lambda(M)| < \infty$ holds.
- R_l is not involved in σ. Then R_l has a residual R'_l in M', which also happens to be the leftmost β-redex in M'. Then $\sigma + [R'_l]$ is a development of $\mathcal{R} \cup \{R_l\}$. Hence $\mathbf{std}(\sigma + [R'_l]) = R_l + \sigma'$ for some $\sigma' : M_l \rightarrow^*_\beta M'_l$, which is a standard development of $\mathcal{R}/[R_l]$. Assume $M' \xrightarrow{R'_l}_\beta M'_l$. Then $|\Lambda(M'_l)| < |\Lambda(M')|$ holds. By induction hypothesis, we have $|\Lambda(M'_l)| \leq |\Lambda(M_l)| < \infty$. This yields that $|\Lambda(M')| = 1 + |\Lambda(M'_l)| \leq 1 + |\Lambda(M_l)| = |\Lambda(M)| < \infty$.

□

Theorem 6.6 (Normalization) *If* M *can be reduced to a normal form, then* $|\Lambda(M)| < \infty$ *holds.*

Proof. With Lemma 6.5, the theorem follows from straightforward induction on the length of σ. □

7 Conclusion and Related Work

We have demonstrated some interesting uses of the *development separation* lemma (Lemma 3.6), proving by structural induction on λ-terms that developments are Church-Rosser and can be standardized. The Church-Rosser theorem in λ-calculus follows immediately. Also, we have employed the technique of development separation in establishing structurally inductive proofs for the standardization theorem, the conservation theorem and the normalization theorem.

When compared to the three proofs of the Church-Rosser theorem in [Bar84], our proof combines the brevity of the first proof (Section 3.2 [Bar84]) and the perspicuity of the second proof (Section 11.1 [Bar84]). Several proofs of the standardization theorem can be found in [Bar84,Tak95], and our proof of the standardization theorem bears some resemblance to the ones due to Klop, where the main strategy is to shift the leftmost β-redex to the front of a β-reduction sequence, though a different strategy is adopted in our case to establish the termination of this process.

Parallel β-reductions are complete developments. Therefore, it is not surprising that the work in [Tak95] can also be done in our setting. On the other hand, Takahashi's method can clearly be used to establish various lemmas in this paper (after they are properly formulated in terms of parallel β-reductions). This can probably described as *separating parallel β-reductions from other β-reductions*.

The technique of separating developments from other β-reductions can also be applied to the call-by-value λ-calculus λ_v, simplifying many proofs in [Plo75]. A λ-calculus λ^v_{hd} is proposed in [Xi97], aiming at providing theoretical background for performing evaluations under λ-abstraction in functional programming languages. The notion of development separation plays a key rôle in establishing several fundamental theorems in λ^v_{hd}.

References

[Bar84] Hendrik Pieter Barendregt. *The Lambda Calculus, Its Syntax and Semantics*, volume 103 of *Studies in Logic and the Foundations of Mathematics*. North-Holland, Amsterdam, revised edition edition, 1984. ISBN 0-444-87508-5. xv+621 pp.

[CF58] Haskell Brookes Curry and Robert Feys. *Combinatory Logic, Volume I*. Studies in Logic and the Foundations of Mathematics. North-Holland, Amsterdam, 1958. xvi+417 pp. Second printing 1968.

[CR36] Alonzo Church and J. B. Rosser. Some Properties of Conversion. *Transactions of the American Mathematical Society*, 39:472–482, 1936.

[dV85] R. de Vrije. A direct proof of the finite developments theorem. *Journal of Symbolic Logic*, 50:339–343, 1985.

[Hin78] J. Roger Hindley. Reductions of residuals are finite. *Transactions of the American Mathematical Society*, 240:345–361, 1978.

[Hue94] Gérard Huet. Residual Theory in λ-Calculus: A Formal Development. *Journal of Functional Programming*, 4:371–394, 1994.

[Hyl73] J. Martin E. Hyland. A simple proof of the Church-Rosser theorem, 1973. Manuscript.

[Hyl76] J. Martin E. Hyland. A syntactic characterisation of the equality in some models of the λ-calculus. *Journal of London Mathematical Society*, 2(12): 361–370, 1976.

[Klo80] J.W. Klop. *Combinatory Reduction Systems*. Ph. D. dissertation, CWI, Amsterdam, 1980. Mathmatical center tracts, No. 127.

[Lév] Jean-Jacques Lévy. An algebraic interpretation of $\lambda - \beta - K$-calculus and a labelled λ-calculus. volume 37 of *Lecture Notes in Computer Science*, pages 147–165. Springer-Verlag.

[Plo75] Gordon D. Plotkin. Call-by-name, call-by-value and the λ-calculus. *Theoretical Computer Science*, 1(2):125–159, December 1975.

[Tak95] M. Takahashi. Parallel Reduction. *Information and Computation*, 118: 120–127, 1995.

[Wad76] Christopher P. Wadsworth. The Relation Between Computational and Denotational Properties for Scott's D_∞-Models of the Lambda-Calculus. *SIAM Journal on Computing*, 5(3):488–521, September 1976.

[Xi97] Hongwei Xi. Evaluation under lambda-abstraction. In *9th International Symposium on Programming Languages, Implementations, Logics, and Programs*, volume 1292 of *Lecture Notes in Computer Science*, pages 259–273. Southampton, September 1997.

[Xi99] Hongwei Xi. Upper bounds for standardisation and an application. *Journal of Symbolic Logic*, 64(1):291–303, March 1999.