Embeddings and Infinite reduction paths in Untyped λ-calculus

Morten Heine Sørensen*

Department of Computer Science, University of Copenhagen (DIKU)
and
Faculty of Mathematics & Informatics, Catholic University of Nijmegen

Abstract

The term $\mathbf{\Omega} \equiv \omega\,\omega$ with $\omega \equiv \lambda x.x\,x$ is the simplest term in untyped λ-calculus with an infinite reduction path: $\mathbf{\Omega} \to \mathbf{\Omega} \to \ldots$ In every step of the reduction path the term reduces to itself.

This characterization does not apply to all infinite reduction paths. In this paper we develop two characterizations which are true for all infinite reduction paths. The first states that in any infinite reduction path, some term must be embedded (in a certain sense) in a subsequent term. The second states that $\mathbf{\Omega}$ is embedded (in a certain sense) in every term of every infinite reduction path.

1. Introduction

1.1. Problem. In the untyped λ-calculus there are terms which have an infinite reduction path. The most obvious example is $\mathbf{\Omega} \equiv \omega\,\omega$, where $\omega \equiv \lambda x.x\,x$. It has an infinite reduction path where in every step the term reduces to itself:

$$\sigma : \quad \mathbf{\Omega} \to \mathbf{\Omega} \to \ldots$$

Not all infinite reduction paths have this form. Indeed, Lercher [16] shows that the terms which reduce to themselves in one step are exactly $C[\mathbf{\Omega}]$ where C ranges over all contexts.

For instance, the term $\Psi \equiv \psi\,\psi$, where $\psi \equiv \lambda x.x\,x\,y$, has the following infinite reduction path:

$$\tau : \quad \Psi \to \Psi\,y \to \Psi\,y\,y \to \ldots$$

In every step the redex Ψ appears as a subterm, and the context of the redex is extended with an application $\bullet\,y$.

*Send correspondence to: Morten Heine Sørensen, Department of Computer Science, University of Copenhagen, Universitetsparken 1, DK-2100 Copenhagen Ø, Denmark. Electronic mail: rambo@diku.dk

As a slightly more complicated example consider the term $v\ y\ v$, where $v \equiv \lambda a x.x\ (a\ y)\ x$. This term has the following infinite reduction path:

$$\rho : \quad v\,y\,v \to (\lambda x.x\,(y\,y)\,x)\,v \to v\,(y\,y)\,v \to (\lambda x.x\,(y\,y\,y)\,x)\,v \to v\,(y\,y\,y)\,v \to \ldots$$

The path ρ is similar to τ, but the extra application $\bullet\, y$ is not added to the context; it is added *inside* the redex.[1]

Although these three reduction paths have their differences they have two things in common:

(i) it happens infinitely often that a term is followed (after a number of steps) by another term which arises from the former by addition of a number of subterms;

(ii) every term arises from $\mathbf{\Omega}$ by addition of a number of subterms. This paper formalizes these two observations and shows that they are correct for all infinite reduction paths.

1.2. OVERVIEW. Section 2 reviews some notions that are used in various parts of the paper. Section 3 introduces two binary relations on the set of λ-terms: *strong embedding* ($\trianglelefteq$) and *weak embedding* ($\sqsubseteq$). As the terminology suggests, the former is contained in the latter, but not vice versa. The remainder of the paper falls in two parts dealing with the two observations above.

The first part studies a translation of the so-called Kruskal Tree Theorem into λ-calculus. Section 4 shows that for any infinite sequence of λ-terms, there is an infinite subsequence where every term is weakly embedded in its successor, and Section 5 applies this result to arbitrary infinite reduction paths.

The second part studies infinite β-reduction paths in connection with certain subcalculi. Section 6 shows that all λA-terms[2] are strongly normalizing, and obtains as a corollary that ω is strongly embedded in every term of every infinite reduction path. Section 7 introduces some tools that are used in Section 8 to introduce the class of λE-terms[3] and show that all λE-terms are strongly normalizing. As a corollary, $\mathbf{\Omega}$ is strongly embedded in every term of every infinite reduction path. Section 9 shows that every term which has both an infinite and a finite reduction path has both $\mathbf{\Omega}$ and $\lambda x.y$ strongly embedded in it (for some variable y other than x).

Section 10 reviews open problems.

1.3. RELATED WORK. The result in Section 4 is a translation into λ-calculus of Kruskal's Tree Theorem due to Higman [8] and in a more general formulation to Kruskal [14]. The result has a beautiful proof due to Nash-Williams [18]. Our proof follows the idea of Nash-William's proof but with some modifications to account for λ-terms instead of finite trees. This involves both simplifications

[1] In effect the context in reduction of Ψ has been turned into an argument by CBV CPS translation [20] of Ψ followed by some trivial manipulations.

[2] Those terms in which for every λ-abstraction $\lambda x.M$, x has at most one free occurrence in M.

[3] Those terms without two disjoint subterms $\lambda x.M$ and $\lambda y.N$ where x occurs free more than once in M and y occurs free more than once in N.

and complications. More references and information about Kruskal's Theorem and related results can be found in [15]. The result has previously been applied to termination tests of term rewrite systems, see Dershowitz' survey [6].

Hindley [9] Shows that all λA-terms are strongly normalizing in order to show that all λA-terms can be typed in the simply typed λ-calculus (à la Curry [2]).

2. Preliminaries

This section reviews some notions that are used in various parts of the paper. We begin with some definitions related to *orders*.

2.1. DEFINITION. Let S be a set with a binary relation R.

(i) R is a *partial order* on S iff R is reflexive, transitive, and antisymmetric.

(ii) R is a *quasi order* on S iff R is reflexive and transitive.

2.2. DEFINITION. Let S be a set.

(i) A partial order R on S is a *well-order* iff for every non-empty subset T of S there is a $t \in T$ such that for all $t' \in T$, tRt'.

(ii) A quasi order R on S is a *well-quasi order* iff for every sequence $s_1, s_2, \ldots$ of elements in S there are $i, j \in \mathbb{N}$ such that $i < j$ and $s_i R s_j$.

Next we review some concepts related to *reduction*. We assume familarity with the concept of a *notion of reduction*, e.g. β and $\beta\eta$, its *compatible closure* $\rightarrow_R$ and its *compatible, reflexive, transitive closure* $\twoheadrightarrow_R$.

2.3. DEFINITION. Let $\mathbf{R}$ be a notion of reduction. An *R-reduction path* is a finite or infinite sequence $M_0, M_1, \ldots$ such that $M_0 \rightarrow_R M_1 \rightarrow_R \ldots$. We assign the name σ to the sequence using the notation

$$\sigma : M_0 \rightarrow_R M_1 \rightarrow_R \ldots$$

We say that M_0 *has* the reduction path σ, and that σ *starts in* M_0. If the sequence is finite $M_0, M_1, \ldots, M_n$ we say that σ *ends in* M_n and that σ *has length* n.

2.4. DEFINITION. Let $\mathbf{R}$ be a notion of reduction.

(i) $\infty_R(M) \Leftrightarrow M$ has an infinite R-reduction path.

(ii) $\mathrm{NF}_R(M) \Leftrightarrow M$ has no R-reduction path of length 1 or more.

(iii) $\mathrm{SN}_R(M) \Leftrightarrow$ All R-reduction paths starting from M are finite.

(iv) $\mathrm{WN}_R(M) \Leftrightarrow$ There is a finite R-reduction path starting from M.

2.5. NOTATION. When the index R is β we will omit it. This happens in Sections 6–9. Also, instead of e.g. $\mathrm{SN}(M)$ we often write $M \in \mathrm{SN}$.

2.6. TERMINOLOGY. Let $\mathbf{R}$ be a notion of reduction.

(i) Elements of NF_R are called *R-normal forms.*

(ii) Elements of SN_R are said to be *R-strongly normalizing.*

(iii) Elements of WN_R are said to be *R-weakly normalizing.*

(iv) If $M \twoheadrightarrow_R N$ and $N \in \mathrm{NF}_R$ then we say that M *has* normal form N.

We assume familiarity with substitution.

2.7. CONVENTION. We identify terms differing only in the names of bound variables. We assume that in any mathematical context (definition, proof, etc.) involving some terms $M_1, \ldots, M_n$ all bound variables in these terms are different from the free variables (see [1, Sec. 2.1]).

2.8. DEFINITION. The *size* $\|M\|$ of a M is defined as follows.

$$\begin{array}{llcl}
(1) & \|x\| & = & 1 \\
(2) & \|\lambda x.P\| & = & 1 + \|P\| \\
(3) & \|P\,Q\| & = & 1 + \|P\| + \|Q\|
\end{array}$$

2.9. DEFINITION. The number of free occurrences of the variable x in M, $\|x\|_{(}, M)$, is defined as follows.

$$\begin{array}{llcl}
(1) & \|x\|_{(}, x) & = & 1 \\
(2) & \|x\|_{(}, y) & = & 0 \quad \text{if } x \not\equiv y \\
(3) & \|x\|_{(}, \lambda y.P) & = & \|x\|_{(}, P) \\
(4) & \|x\|_{(}, P\,Q) & = & \|x\|_{(}, P) + \|x\|_{(}, Q)
\end{array}$$

2.10. LEMMA.

(i) $\|M\{x := N\}\| = \|M\| + \|x\|_{(}, M) \cdot (\|N\| - 1)$

(ii) $\|x\|_{(}, M\{y := N\}) = \|x\|_{(}, M) + \|x\|_{(}, N) \cdot \|y\|_{(}, M)$

PROOF.

(i) Induction on M.

(ii) Induction on M. $\square$

2.11. DEFINITION.

(i) $\lambda x.M$ is *duplicating* if $\|x\|_{(}, M) > 1$, and *non-duplicating* otherwise.

(ii) $\lambda x.M$ is *erasing* if $\|x\|_{(}, M) = 0$, and *non-erasing* otherwise.

3. Embedding relations on Λ

This section introduces three binary relations on Λ and establishes their most fundamental properties and connections.

3.1. DEFINITION. The *subterm* relation $\subseteq$ on Λ is defined as follows.

$$
\begin{array}{llll}
(1) & P \equiv Q & \Rightarrow & P \subseteq Q \\
(2) & P \subseteq Q & \Rightarrow & P \subseteq \lambda x.Q \\
(3) & P \subseteq Q & \Rightarrow & P \subseteq Q\,Z \\
(4) & P \subseteq Q & \Rightarrow & P \subseteq Z\,Q
\end{array}
$$

$M \subset N \Leftrightarrow M \subseteq N$ and $M \not\equiv N$.

3.2. EXAMPLE.

(i) $x \subseteq \lambda y.y$ (since $x \subseteq \lambda x.x$ and we identify α-equivalent terms in the meta-language).

(ii) $\lambda x.x \subset \lambda z.(\lambda x.x)\,z$.

(iii) $\lambda x.x \not\subseteq \lambda x.x\,x$.

3.3. PROPOSITION.

(i) $M \subset N \Rightarrow \|M\| < \|N\|$.

(ii) $\subseteq$ *is a partial order.*

PROOF.

(i) By induction on the derivation of $M \subseteq N$ show that $M \not\equiv N$ implies $\|M\| < \|N\|$.

(ii) *Reflexivity:* By rule (4). *Antisymmetry:* If $M \subseteq N$ and $N \subseteq M$ but $M \not\equiv N$ then $M \subset N$ and $N \subset M$, and therefore by (i) $\|M\| < \|N\|$ and $\|N\| < \|M\|$, a contradiction. *Transitivity:* By induction on the derivation of $M_2 \subseteq M_3$ prove that $M_1 \subseteq M_2$ and $M_2 \subseteq M_3$ implies $M_1 \subseteq M_3$.$\square$

3.4. REMARK. $M \subseteq N$ formalizes the idea that M as a whole occurs somewhere inside N.

3.5. DEFINITION. The *strong homeomorphic embedding* relation $\trianglelefteq$ on Λ is defined as follows.

$$
\begin{array}{llll}
(1) & x \trianglelefteq x & & \\
(2) & P \trianglelefteq Q & \Rightarrow & P \trianglelefteq \lambda x.Q \\
(3) & P \trianglelefteq Q & \Rightarrow & P \trianglelefteq Q\,Z \\
(4) & P \trianglelefteq Q & \Rightarrow & P \trianglelefteq Z\,Q \\
(5) & P \trianglelefteq Q & \Rightarrow & \lambda x.P \trianglelefteq \lambda x.Q \\
(6) & P_1 \trianglelefteq Q_1, P_2 \trianglelefteq Q_2 & \Rightarrow & P_1\,P_2 \trianglelefteq Q_1\,Q_2
\end{array}
$$

$M \triangleleft N \Leftrightarrow M \trianglelefteq N$ and $M \not\equiv N$.

3.6. Example.

(i) $\lambda x.x \trianglelefteq \lambda z.(\lambda x.x)\, z$.

(ii) $\lambda x.x \trianglelefteq \lambda x.x\, x$.

(iii) $(\lambda x.x\, x)\, (\lambda x.x\, x) \trianglelefteq (\lambda x.x\, x\, y)\, (\lambda x.x\, x\, y)$.

(iv) $\lambda x.x\, x\, y \not\trianglelefteq \lambda x.x\, x$.

(v) $\lambda xy.x \not\trianglelefteq \lambda xy.y$.

3.7. Proposition.

(i) $M \triangleleft N \Rightarrow \|M\| < \|N\|$.

(ii) $\trianglelefteq$ *is a partial order.*

Proof.

(i) By induction on the derivation $M \trianglelefteq N$ show that $M \not\equiv N$ implies $\|M\| < \|N\|$.

(ii) *Reflexivity:* Prove by induction on M that $M \trianglelefteq M$. *Antisymmetry:* Like in the proof of Proposition 3.3(ii), using (i) instead of Proposition 3.3(i). *Transitivity:* By induction on the derivation of $M_1 \trianglelefteq M_2$ and $M_2 \trianglelefteq M_3$ prove that $M_1 \trianglelefteq M_2$ and $M_2 \trianglelefteq M_3$ implies $M_1 \trianglelefteq M_3$, splitting into cases according to the last rule applied to show $M_2 \trianglelefteq M_3$: Rule (1)-(3), rule (4), rule (5), and rule (6).□

3.8. Remark. $M \trianglelefteq N$ formalizes the idea that all the pieces of M are present in N in the same order. Indeed, $M \trianglelefteq N$ iff M arises from N by deleting a number of operators, arguments, and abstractors.

3.9. Definition. The *weak homeomorphic embedding* relation $\sqsubseteq$ on Λ is defined as follows.

$$\begin{array}{lll} (1) & x \sqsubseteq y & \\ (2) & P \sqsubseteq Q & \Rightarrow \quad P \sqsubseteq \lambda x.Q \\ (3) & P \sqsubseteq Q & \Rightarrow \quad P \sqsubseteq Q\, Z \\ (4) & P \sqsubseteq Q & \Rightarrow \quad P \sqsubseteq Z\, Q \\ (5) & P \sqsubseteq Q & \Rightarrow \quad \lambda x.P \sqsubseteq \lambda y.Q \\ (6) & P_1 \sqsubseteq Q_1, P_2 \sqsubseteq Q_2 & \Rightarrow \quad P_1\, P_2 \sqsubseteq Q_1\, Q_2 \end{array}$$

$M \sqsubset N \Leftrightarrow M \sqsubseteq N$ and $M \not\equiv N$.

3.10. Example.

(i) $(\lambda x.x\, x)\, (\lambda x.x\, x) \sqsubseteq (\lambda x.x\, x\, y)\, (\lambda x.x\, x\, y)$

(ii) $\lambda xy.x \sqsubseteq \lambda xy.y$.

(iii) $\lambda xy.y \sqsubseteq \lambda xy.x$.

(iv) $\lambda x.x\,x\,y \not\sqsubseteq \lambda x.x\,x$.

3.11. PROPOSITION.

(i) $M \sqsubseteq N \Rightarrow \|M\| \leq \|N\|$.

(ii) $\sqsubseteq$ *is a quasi order.*

PROOF.

(i) By induction on the derivation of $M \sqsubseteq N$.

(ii) *Reflexivity:* Prove by induction on M that $M \trianglelefteq M$. *Transitivity:* As in the proof of Proposition 3.7(ii).□

3.12. REMARK.

(i) A strict version of Proposition 3.11(i) similar to Propositions 3.3(i) and 3.7(i) does not hold as Example 3.10(ii) shows. Example 3.10(ii)-(iii) show that $\sqsubseteq$ is not antisymmetric.

(ii) $M \sqsubseteq N$ has a similar meaning to $M \trianglelefteq N$, except that whether $M \sqsubseteq N$ holds or not, only depends on the *skeleton* of M and N (see Proposition 3.15).

3.13. DEFINITION (Barendregt [1, Ex. 13.6.8]). The *skeleton* of a λ-term is defined by the map $|\bullet|$ as follows.

$$\begin{array}{lll} (1) & |x| & = & \square \\ (2) & |\lambda x.M| & = & \lambda\square.|M| \\ (3) & |M\,N| & = & |M|\,|N| \end{array}$$

3.14. EXAMPLE. $|\lambda y.y\,(\lambda x.x\,y\,z)| = \lambda\square.\square\,(\lambda\square.\square\,\square\,\square)$.

3.15. PROPOSITION. *Suppose* $M, M', N, N' \in \Lambda$ *are such that*

$$|M| = |M'| \text{ and } |N| = |N'|$$

Then

$$M \sqsubseteq N \Leftrightarrow M' \sqsubseteq N'$$

PROOF. It suffices to show that $M \sqsubseteq N$ implies both $M' \sqsubseteq N$ and $M \sqsubseteq N'$. Both properties are proved by induction on the derivation of $M \sqsubseteq N$.□

3.16. REMARK. Example 3.6(v) shows that this proposition does not hold for $\trianglelefteq$.

3.17. PROPOSITION.

(i) $M \subseteq N \Rightarrow M \trianglelefteq N$

(ii) $M \trianglelefteq N \Rightarrow M \sqsubseteq N$

(iii) $M \subseteq N \not\Leftarrow M \trianglelefteq N$

(iv) $M \trianglelefteq N \not\Leftarrow M \sqsubseteq N$

PROOF.

(i) Induction on the derivation of $M \subseteq N$.

(ii) Induction on the derivation of $M \trianglelefteq N$.

(iii) Examples 3.2(iii) and 3.6(ii).

(iv) Examples 3.6(v) and 3.10(ii). □

4. Kruskal's Tree Theorem on Λ

This section shows that for any infinite sequence of λ-terms there is a term and a subsequent term such that the former is weakly embedded in the latter.

The first lemma can often replace applications of the Axiom of Choice[4] when proving assertions concerning infinite sequences of λ-terms. The intuition of the proof is that we can enumerate all λ-terms and choose from any non-empty subset of Λ the term with the smallest index.

4.1. LEMMA. *There exists a (choice) function*

$$f : \mathcal{P}(\Lambda)\backslash\{\emptyset\} \to \Lambda$$

such that for all $\Delta \in \mathcal{P}(\Lambda)\backslash\{\emptyset\}$

$$f(\Delta) \in \Delta$$

PROOF. It suffices to show that there is a well-order on Λ (see [7, p69]).

Let $i : \Lambda \to I\!N$ be an injective map (a Gödel numbering of Λ, see e.g. [2, 2.2.13]) and define on Λ the following partial order:

$$M \preceq N \quad \Leftrightarrow \quad i(M) \leq i(N)$$

Now let Δ be a non-empty subset of Λ, and consider the set

$$\{i(M)|M \in \Delta\}$$

As a non-empty set of numbers this set has a smallest member $n_0 \in I\!N$. Then there is an $M_0 \in \Delta$ with $i(M_0) = n_0$. Clearly, for all M in Δ, $i(M_0) \leq i(M)$, i.e. $M_0 \preceq M$.□

4.2. DEFINITION.

(i) The set of infinite sequences of λ-terms is denoted $\Sigma_\infty(\Lambda)$.

[4] This section requires a slender acquaintance with the axioms of set theory as presented in [7].

(ii) $(M_0, M_1, \ldots) \in \Sigma_\infty(\Lambda)$ is *good* if there are $i, j \in I\!N$ with $i < j$ and $M_i \sqsubseteq M_j$.

(iii) If $\sigma \in \Sigma_\infty(\Lambda)$ is not good, it is *bad*.

(iv) $(M_0, M_1, \ldots) \in \Sigma_\infty(\Lambda)$ is *ascending* if for all $i \in I\!N$: $M_i \sqsubseteq M_{i+1}$.

In proofs of Kruskal's Tree Theorem the analogue of the following lemma is usually justified simply by reference to the Axiom of Choice, because the details involve a standard set theoretical technique. Below these details are included to show that Lemma 4.1 is sufficient to avoid using the Axiom of Choice. For readers acquainted with the ways of the Axiom of Choice this will already be apparent.

4.3. LEMMA. *If $\sigma = (N_0, N_1, \ldots) \in \Sigma_\infty(\Lambda)$ is bad then there is another bad sequence $\sigma' = (M_0, M_1, \ldots) \in \Sigma_\infty(\Lambda)$ such that for all $n \in I\!N$ there is no bad sequence whose first $n+1$ elements are (in this order) $M_0, M_1, \ldots, M_{n-1}, M'$ where M' is some term with $\|M'\| < \|M_n\|$.*

PROOF. Let f be the function from Lemma 4.1, and let $\mathcal{S}$ be the collection of all subsets Δ of Λ such that for some bad $(L_0, L_1, \ldots) \in \Sigma_\infty(\Lambda)$ and some $n \in I\!N$

$$\Delta = \cup_{i=0}^{n-1}\{L_i\}$$

Note that since σ is bad and $\emptyset = \cup_{i=0}^{-1}\{N_i\}$, it holds that $\emptyset \in \mathcal{S}$.

For every $\Delta \in \mathcal{S}$ define

$$\begin{array}{rcl} \Delta' & = & \{M \in \Lambda \mid \Delta \cup \{M\} \in \mathcal{S} \ \wedge\ \forall N \in \Delta : N \not\sqsubseteq M\} \\ \Delta'' & = & \{M \in \Delta' \mid \forall N \in \Delta' : \|M\| \leq \|N\|\} \end{array}$$

We claim that

$$(1) \qquad \Delta'' \neq \emptyset \text{ and } \forall M \in \Delta'' : \Delta \cup \{M\} \in \mathcal{S}$$

The second part follows by the definition of Δ' and the fact that $\Delta'' \subseteq \Delta'$.

For the first part we first show that $\Delta' \neq \emptyset$. Since $\Delta = \cup_{i=0}^{n-1}\{L_i\}$ for some $n \in I\!N$ and bad $(L_0, L_1, \ldots) \in \Sigma_\infty(\Lambda)$, it holds that $\Delta \cup \{L_n\} = \cup_{i=0}^{n}\{L_i\}$ and so $\Delta \cup \{L_n\} \in \mathcal{S}$. Moreover, since $(L_0, L_1, \ldots)$ is bad, $L_i \not\sqsubseteq L_n$ for all $i \in I\!N$ with $0 \leq i < n$. Thus $L_n \in \Delta'$.

To see that Δ'' is non-empty if Δ' is, consider the set $\{\|M\| \mid M \in \Delta'\}$. As a non-empty set of numbers it has a smallest element $n_0 \in I\!N$. Then for some $M_0 \in \Delta'$, $\|M_0\| = n_0$, and clearly $M_0 \in \Delta''$.

According to (1) now define $g : \mathcal{S} \to \mathcal{S}$ by

$$g(\Delta) = \Delta \cup \{f(\Delta'')\}$$

and $u : I\!N \to \mathcal{S}$ recursively by

$$\begin{array}{rcl} u(0) & = & \emptyset \\ u(n+1) & = & g(u(n)) \end{array}$$

and finally $\sigma' \in \Sigma_\infty(\Lambda)$ by

$$\sigma' = (f(u(0)''), f(u(1)''), \ldots)$$

To see that σ' is bad prove by induction on n that

$$(2) \quad u(n) = \cup_{k=0}^{n-1}\{f(u(k)'')\}$$

and by straight-forward calculation of $u(n)''$ that

$$(3) \quad \forall N \in u(n)\ \forall M \in u(n)'' : N \not\sqsubseteq M$$

If $i < j$ for some $i, j \in \mathbb{N}$ then $f(u(i)'') \in u(j)$ by (2), and since $f(u(j)'') \in u(j)''$, (3) implies that $f(u(i)'') \not\sqsubseteq f(u(j)'')$, as required.

To see that σ' is minimal in the sense of the lemma, note that

$$(4) \quad \forall M \in u(n)''\ \forall N \in u(n)' : \|M\| \leq \|N\|$$

and show by induction on n that, if

$$\sigma' = (f(u(0)''), f(u(1)''), \ldots, f(u(n-1)''), K_0, K_1, \ldots)$$

is bad for some $K_0, K_1, \ldots$ then $K_0 \in u(n)'$. Since $f(u(n)'') \in u(n)''$, (4) implies $\|f(u(n)'')\| \leq \|K_0\|$ as required. □

The following lemma is well-known [8, 18].

4.4. Lemma. *Let $\Sigma \subseteq \Sigma_\infty(\Lambda)$ be closed under subsequence formation.*

$\forall \sigma \in \Sigma : \sigma$ *is good* $\Leftrightarrow \forall \sigma \in \Sigma : \sigma$ *has an infinite ascending subsequence*

Proof.

$\Leftarrow$:Obvious.

$\Rightarrow$: Suppose all $\sigma \in \Sigma$ are good, and consider a given $\sigma = (M_0, M_1, \ldots)$. Call M_i *terminal* if there is no $j \in \mathbb{N}$ such that $M_i \sqsubseteq M_j$. There are only finitely many terminals; otherwise they would form a bad sequence. Thus there is an $N \in \mathbb{N}$ such that for all $n > N$ there is an $m > n$ with $M_n \sqsubseteq M_m$. From this one easily defines by recursion an infinite ascending subsequence.□

4.5. Theorem (Higman [8], Kruskal [14]). *$\sqsubseteq$ is a well-quasi order on Λ.*

Proof (Nash-Williams [18]). Suppose the theorem were false, i.e. there were a bad sequence. By Lemma 4.3 there is a bad sequence

$$\sigma = (M_0, M_1, \ldots)$$

such that for all $n \in \mathbb{N}$ there is no bad sequence whose first $n+1$ elements are (in this order) $M_0, M_1, \ldots, M_{n-1}, M'$ where M' is some term with $\|M'\| < \|M_n\|$.

Since $x \sqsubseteq y$ for all variables x, y and σ is bad, σ must have an infinite subsequence

$$(M_{i_0}, M_{i_1}, \ldots)$$

such that all elements are abstractions or all elements are applications. We show that both cases lead to a contradiction.

Case 1. $M_{i_j} \equiv \lambda x_j.N_j$ for all $j \in I\!N$. Consider the sequence

$$\sigma' = (N_0, N_1, \ldots)$$

Subcase 1.1. σ' is bad. Then consider

$$\sigma'' = (M_0, M_1, \ldots M_{i_0-1}, N_0, N_1, \ldots)$$

Since σ and σ' are bad, $M_i \not\sqsubseteq M_j$ (for $0 \leq i < j \leq i_0 - 1$) and $N_n \not\sqsubseteq N_m$ (for $0 \leq n < m$). Moreover, if $M_i \sqsubseteq N_m$ (where $0 \leq i \leq i_0 - 1, 0 \leq m$) then $M_i \sqsubseteq \lambda x_m.N_m \equiv M_{i_m}$, contradicting badness of σ. Thus $M_i \not\sqsubseteq N_m$. So σ'' is bad, but this contradicts minimality of σ.

Subcase 1.2. σ' is good. Then $N_n \sqsubseteq N_m$ for some $n < m$, but then

$$M_{i_n} \equiv \lambda x_n.N_n \sqsubseteq \lambda x_m.N_m \equiv M_{i_m}$$

contradicting badness of σ.

Case 2. $M_{i_j} \equiv N_j\, K_j$ for all $j \in I\!N$. Consider the *sets*

$$\{N_0, N_1, \ldots\} \qquad \{K_0, K_1, \ldots\}$$

Subcase 2.1. One of the sets, say $\{N_0, N_1, \ldots\}$, has a bad sequence. Then in particular it has a bad sequence

$$N_{j_0}, N_{j_1}, \ldots$$

where $j_0 < j_k$ for all $k > 0$ (from the former bad sequence simply remove the finitely many elements with an index smaller than the index of the first element). As in Subcase 1.1

$$(M_0, M_1, \ldots M_{i_{j_0}-1}, N_{j_0}, N_{j_1}, \ldots)$$

is then a bad sequence contradicting mimimality of σ.

Subcase 2.2. All sequences in each of the two sets are good. By Lemma 4.4 the sequence $(N_0, N_1, \ldots)$ has an infinite ascending sequence $(N_{j_0}, N_{j_1}, \ldots)$. The corresponding $(K_{j_0}, K_{j_1}, \ldots)$ being an infinite sequence in $\{K_0, K_1, \ldots\}$ is good, so there are $n < m$ such that $K_{j_n} \sqsubseteq K_{j_m}$. Since $(N_{j_0}, N_{j_1}, \ldots)$ is ascending and $\sqsubseteq$ is transitive, $N_{j_n} \sqsubseteq N_{j_m}$. But then also

$$M_{i_{j_n}} \equiv K_{j_n}\, N_{j_n} \sqsubseteq K_{j_m}\, N_{j_M} \equiv M_{i_{j_m}}$$

contradicting badness of σ.$\Box$

4.6. Remark. The reader familiar with [18] may note that although the structure of the present proof is similar to that in [18], both the applications of the Axiom of Choice in [18] (one in the main result and one in a lemma) have been avoided. The first one has been avoided by taking into account that a choice function for Λ can be constructed without the Axiom of Choice, and the secod application has been avoided by taking into account that the term formation operations in Λ (abstraction and application) have fixed arity 1 and 2, respectively.

5. Application to infinite reduction paths

Suppose

$$\sigma : M_0 \to_R M_1 \to_R \dots$$

is an infinite reduction path. It clearly cannot be the case that

$$\|M_0\| > \|M_1\| > \dots$$

so there must be $i < j$ such that $\|M_i\| \leq \|M_j\|$.

However this only concerns the *size* of the terms in σ. Recall from the introduction that some kind of repetition of *structure* seemed to occur in the example infinite reduction paths. One might guess that for every infinite reduction path there must be a term and a subsequent term such that the former is identical to the latter or a subterm of the latter (as suggested by the first two examples in the introduction). However, the third example in the introduction shows that this is not the case. As shown below, one needs to use $\sqsubseteq$ instead of $\subseteq$.

5.1. DEFINITION. Let $\mathbf{R}$ be a notion of reduction. For $M, N \in \Lambda$ a reduction $M \twoheadrightarrow_R N$ is an *R-cycle* if $M \equiv N$, an *strong R-self-embedding* if $M \trianglelefteq N$, and an *weak R-self-embedding* if $M \sqsubseteq N$.

5.2. REMARK. Cycles are studied by Klop [12]. Related notions are studied by Sekimoto and Hirokawa [10, 22], Zilli [23, 24], and Böhm and Micali [4]. Dershowitz [6] and Plaisted [19] study a notion of self-embedding in term rewrite systems.

5.3. COROLLARY. *Let* $\mathbf{R}$ *be a notion of reduction. Suppose for* $M_i \in \Lambda$

$$\sigma : M_0 \to_R M_1 \to_R \dots$$

Then there are $i_0 < j_0 < i_1 < j_1 < \dots$ *such that for all* $k \in \mathbb{N}$

$$M_{i_k} \twoheadrightarrow_R M_{j_k}$$

is a weak R-self-embedding.

PROOF. By Lemma 4.4 and Theorem 4.5.□

5.4. REMARK. There are terms M with N such that $M \twoheadrightarrow_\beta N$ is a weak self-embedding, and yet M has no infinite reduction paths. For example,

$$(\lambda x.x\,x)\;(\lambda x.y\,x\,x) \to (\lambda x.y\,x\,x)\;(\lambda x.y\,x\,x) \to y\;(\lambda x.y\,x\,x)\;(\lambda x.y\,x\,x)$$

is a weak self-embedding but the first term has no infinite reduction paths (the last term is in normal form). By the same example, there are terms with strong self-embeddings and no infinite reduction paths.

5.5. REMARK. The generality of the preceding corollary, holding for any notion of reduction, suggests that something better can be obtained specifically for β-reduction. More precisely, by instantiating the notion of reduction in the corollary to β one might expect it to hold with a smaller relation than $\sqsubseteq$. On the other hand, from the examples in the introduction it seems that a lot of $\sqsubseteq$ is needed already for β. A clarification of this aspect remains future work.

6. Strong embedding and λA- terms

In this section we introduce λA-terms and show that they are strongly normalizing. As a corollary, ω is strongly embedded in every term which is not strongly normalizing. From now on the notion of reduction is β.

6.1. DEFINITION. The set of λA-terms, Λ_A is defined as follows.

$$\begin{array}{llcl} (1) & x \in \Lambda_A & & \\ (2) & P \in \Lambda_A, \|x\|_{(}, P) \leq 1 & \Rightarrow & \lambda x.P \in \Lambda_A \\ (3) & P, Q \in \Lambda_A & \Rightarrow & P\,Q \in \Lambda_A \end{array}$$

6.2. REMARK. The terminology λA is from [11] Hindley [9] calls elements of Λ_A *linear* λ-terms, crediting the terminology to Rezus. Because of the correspondence between combinatory terms with the combinators $\mathbf{B}, \mathbf{C}, \mathbf{K}$ and λA-terms, Komori [13] calls elements of Λ_A $BCK\lambda$-terms.

The essential syntactical property of Λ_A is that it is the set of λ-terms for which no term has a subterm which is a duplicating abstraction. This explains the following result.

6.3. LEMMA. *$\forall M \in \Lambda : M \notin \Lambda_A \Rightarrow \omega \trianglelefteq M$.*

PROOF. By induction on M. In the case of an abstraction use the fact that $occ(x, M) \geq 2$ implies $x\,x \trianglelefteq M$.□

6.4. LEMMA.

(i) $M \in \Lambda_A, N \in \Lambda_A \Rightarrow M\{x := N\} \in \Lambda_A$.

(ii) $M \in \Lambda_A, M \to N \Rightarrow \|x\|_{(}, M) \geq \|x\|_{(}, N)$

(iii) $M \in \Lambda_A, M \to N \Rightarrow N \in \Lambda_A$.

PROOF.

(i) Induction on the derivation of $M \in \Lambda_A$.

(ii) Induction on the derivation of $M \to N$, using Lemma 2.10(ii).

(iii) By induction on the derivation of $M \to N$, using (i)-(ii). □

The last result above shows that Λ_A is closed under reduction. The intuition behind this is as follows. Recall that all abstractions in λA-terms are non-duplicating. Abstractions in the contractum of a term are descendants of abstractions in the original term, and to turn a non-duplicating abstraction into a duplicating one, one needs to duplicate a variable in the body of the former abstraction, but this would itself require a duplicating abstraction.

6.5. LEMMA. *$M \in \Lambda_A, M \to N \Rightarrow \|M\| > \|N\|$.*

PROOF. By induction on the derivation of $M \to N$, using Lemma 2.10(i). □

6.6. PROPOSITION (Hindley [9]). $M \in \Lambda_A \Rightarrow M \in \mathrm{SN}$.

PROOF. Suppose $M \in \Lambda_A$ and $\infty(M)$, i.e. M has an infinite reduction path

$$\sigma : M \equiv M_0 \to M_1 \to \dots$$

By induction on i using Lemma 6.4(iii) one shows that for all $i \in I\!N$, $M_i \in \Lambda_A$. By Lemma 6.5 this gives an infinite sequence

$$\|M_0\| > \|M_1\| > \dots$$

which is clearly a contradiction. Thus $M \in \mathrm{SN}$.□

6.7. COROLLARY. $\infty(M) \Rightarrow \omega \trianglelefteq M$.

PROOF. If $\omega \not\trianglelefteq M$ then $M \in \Lambda_A$ by Lemma 6.3 and then $M \in \mathrm{SN}$ by Proposition 6.6.□

7. A perpetual strategy

This section interrupts the development of the preceding section in order to establish tools to strengthen Corollary 6.7.

7.1. DEFINITION.

(i) A *one-step R-reduction strategy* is a map $F : \Lambda \to \Lambda$ such that $M \notin \mathrm{NF} \Rightarrow M \to_R F(M)$, and $M \in \mathrm{NF} \Rightarrow F(M) = M$.

(ii) F is *R-perpetual* iff $\infty_R(M) \Rightarrow \infty_R(F(M))$.

A well-known one-step β-reduction strategy is F_l which reduces the leftmost redex (see [1]).

7.2. DEFINITION. Define $F : \Lambda \to \Lambda$ as follows. If $M \in \mathrm{SN}$ then $F(M) = F_l(M)$; otherwise,

$$\begin{array}{lcll}
F(x\, P_1 \dots P_n) & = & x\, P_1 \dots P_{i-1}\, F(P_i)\, P_{i+1} \dots P_n & \text{If } P_1, \dots, P_{i-1} \in \mathrm{SN}, P_i \notin \mathrm{SN} \\
F(\lambda x.P) & = & \lambda x.F(P) & \\
F((\lambda x.P_0)\, P_1 \dots P_n) & = & P_0\{x := P_1\}\, P_2 \dots P_n & \text{If } P_0, \dots, P_n \in \mathrm{SN} \\
F((\lambda x.P_0)\, P_1 \dots P_n) & = & (\lambda x.F(P_0))\, P_1 \dots P_n & \text{If } P_0 \notin \mathrm{SN} \\
F((\lambda x.P_0)\, P_1 \dots P_n) & = & (\lambda x.P_0)\, P_1 \dots P_{i-1} F(P_i)\, P_{i+1} \dots P_n & \text{If } P_0, \dots, P_{i-1} \in \mathrm{SN}, P_i \notin \mathrm{SN}
\end{array}$$

7.3. REMARK. For every $M \in \Lambda$, either $M \in \mathrm{SN}$ or $M \notin \mathrm{SN}$. In the latter case either $M \equiv x\, P_1 \dots P_n$ where $n \geq 1$ and $P_i \notin \mathrm{SN}$ for some i, or $M \equiv \lambda x.P$, or $M \equiv (\lambda x.P_0)\, P_1 \dots P_n$ where $n \geq 1$. It follows that F is defined on all λ-terms.

7.4. PROPOSITION. *F is a perpetual one-step reduction strategy.*

PROOF. We are to prove that for all $M \in \Lambda$

$$\begin{array}{llll} (1) & M \notin \mathrm{NF} & \Rightarrow & M \to F(M) \\ (2) & M \in \mathrm{NF} & \Rightarrow & F(M) = M \\ (3) & \infty(M) & \Rightarrow & \infty(F(M)) \end{array}$$

If $M \in \mathrm{SN}$ then (1)-(2) hold because F_l is a one-step reduction strategy, and (3) holds trivially. It therefore suffices to show that, if $\infty(M)$ then $M \to F(M)$ and $\infty(F(M))$. We proceed by induction on the size of M, splitting into cases according to Remark 7.3.

Case 1. $M \equiv x\ P_1 \dots P_n, n \geq 1$. Let $i = \mu j : \infty(P_j)$. By induction hypothesis $P_i \to F(P_i)$ and $\infty(F(P_i))$. Then

$$M \equiv x\ P_1 \dots P_n \to x\ P_1 \dots P_{i-1}\ F(P_i)\ P_{i+1} \dots P_n = F(M)$$

and clearly $\infty(F(P_i))$.

Case 2. $M \equiv \lambda x.P$. Similar to *Case 1.*

Case 3. $M \equiv (\lambda x.P_0)\ P_1 \dots P_n, n \geq 1$. We consider three cases.

Subcase 3.1. $P_0, \dots, P_n \in \mathrm{SN}$. Then

$$M \to P_0\{x := P_1\}\ P_2 \dots P_n = F(M)$$

Since $\infty(M)$ but $P_0, \dots P_n \in \mathrm{SN}$, there must be an infinite reduction path from M of form

$$M \twoheadrightarrow_\beta (\lambda x.P_0')\ P_1'\ P_2' \dots P_n' \to P_0'\{P_1' := \}P_2' \dots P_n' \to \dots$$

but then there also is an infinite reduction path from $F(M)$:

$$F(M) = P_0\{x := P_1\}\ P_2 \dots P_n \twoheadrightarrow_\beta P_0'\{x := P_1'\}\ P_2' \dots P_n' \to \dots$$

Subcase 3.2. $P_0 \notin \mathrm{SN}$. Similar to Case 1.

Subcase 3.3. $P_0, \dots, P_{i-1} \in \mathrm{SN}, P_i \notin \mathrm{SN}$. Similar to Case 1. □

7.5. REMARK. The proof that F is perpetual is somewhat simpler than other related proofs of perpetuality, e.g. [1, 13.4.16] and [3]. This is mainly because F propagates to subterms with infinite reduction paths as often as possible, implying that the induction hypothesis applies immediately.

The essential property of F is that if F goes under a λ to contract a redex in a term with an infinite reduction, then all redexes in the remainder of the infinite reduction will appear under that λ. The next definition and lemma are intended to formalize and prove this.

7.6. DEFINITION. Define $V : \Lambda \to \Lambda$ as follows. If $M \in \mathrm{SN}$ then $V(M) = M$; otherwise,

$$\begin{array}{lcll} V(x\ P_1 \dots P_n) & = & V(P_i) & \text{If } P_1, \dots, P_{i-1} \in \mathrm{SN}, P_i \notin \mathrm{SN} \\ V(\lambda x.P) & = & V(P) & \\ V((\lambda x.P_0)\ P_1 \dots P_n) & = & (\lambda x.P_0)\ P_1 \dots P_n & \text{If } P_0, \dots, P_n \in \mathrm{SN} \\ V((\lambda x.P_0)\ P_1 \dots P_n) & = & V(P_0) & \text{If } P_0 \notin \mathrm{SN} \\ V((\lambda x.P_0)\ P_1 \dots P_n) & = & V(P_i) & \text{If } P_0, \dots, P_{i-1} \in \mathrm{SN}, P_i \notin \mathrm{SN} \end{array}$$

7.7. Remark. For all $M \in \Lambda$, $V(M) \subseteq M$ is well-defined.

7.8. Lemma. *If* $\infty(M)$ *then*

$$V(M) = (\lambda x.Q_0)\, Q_1 \ldots Q_n$$

for some $Q_0 \ldots Q_n$, $n \geq 1$, *with*

$$V(F(M)) \subseteq Q_0\{x := Q_1\}\, Q_2 \ldots Q_n$$

Proof. By induction on the size of M, splitting into cases according to Remark 7.3.

Case 1. $M \equiv x\ P_1 \ldots P_n, n \geq 1$. Let $i = \mu j : \infty(P_j)$. By induction hypothesis $V(P_i) = (\lambda x.Q_0)\, Q_1 \ldots Q_n$ for some $Q_0 \ldots Q_n$, $n \geq 1$, with $V(F(P_i)) \subseteq Q_0\{x := Q_1\}\, Q_2 \ldots Q_n$. Then

$$V(M) = V(P_i) = (\lambda x.Q_0)\, Q_1 \ldots Q_n$$

By Proposition 7.4 it holds that $\infty(F(P_i))$, so

$$V(F(M)) = V(xP_1 \ldots P_{i-1}F(P_i)P_{i+1} \ldots P_n) = V(F(P_i)) \subseteq Q_0\{x := Q_1\}Q_2 \ldots Q_n$$

Case 2. $M \equiv \lambda x.P$. Similar to *Case 1.*

Case 3. $M \equiv (\lambda x.P_0)\ P_1 \ldots P_n, n \geq 1$. We consider three subcases.

Subcase 3.1. $P_0, \ldots, P_n \in \mathrm{SN}$. Then

$$V(M) = M = (\lambda x.P_0)\ P_1 \ldots P_n$$

and

$$V(F(M)) = V(P_0\{x := P_1\}\ P_2 \ldots P_n) \subseteq P_0\{x := P_1\}\ P_2 \ldots P_n$$

Subcase 3.2. $P_0 \notin \mathrm{SN}$. Similar to Case 1.

Subcase 3.3. $P_0, \ldots, P_{i-1} \in \mathrm{SN}, M_i \notin \mathrm{SN}$. Similar to Case 1. □

This concludes our study of F.

8. Strong embedding and λE-terms

This section extends Proposition 6.6 to a somewhat larger set. As a corollary, every term that is not strongly normalizing has $\mathbf{\Omega}$ strongly embedded in it.

8.1. Definition. Define the set of *λE-terms*, Λ_E, as follows.

$$\begin{array}{lll} (1) & x \in \Lambda_E & \\ (2) & M \in \Lambda_E & \Rightarrow \quad \lambda x.M \in \Lambda_E \\ (3) & M \in \Lambda_E, N \in \Lambda_A & \Rightarrow \quad M\ N \in \Lambda_E \\ (4) & M \in \Lambda_A, N \in \Lambda_E & \Rightarrow \quad M\ N \in \Lambda_E \end{array}$$

The essential syntactical property of Λ_E is that it is the set of λ-terms for which no term has a two disjoint subterms that are duplicating abstractions. This explains the following result.

8.2. LEMMA. $\forall M \in \Lambda : M \notin \Lambda_E \Rightarrow \mathbf{\Omega} \trianglelefteq M$

PROOF. Induction on M using Lemma 6.3.□

8.3. LEMMA.

(i) $\Lambda_A \subseteq \Lambda_E$.

(ii) $M \in \Lambda_A, N \in \Lambda_E, \|x\|_(, M) \leq 1 \Rightarrow M\{x := N\} \in \Lambda_E$.

(iii) $M \in \Lambda_E, N \in \Lambda_A \Rightarrow M\{x := N\} \in \Lambda_E$.

PROOF.

(i) Induction on the derivation of $M \in \Lambda_A$ prove $M \in \Lambda_E$.

(ii) Induction on the derivation of $M \in \Lambda_A$.

(iii) induction on the derivation of $M \in \Lambda_E$, using (i) and Lemma 6.4(i).□

The next result shows that Λ_E is closed under reduction. The intuition behind this result is as follows. Recall that λE-terms do not have disjoint duplicating abstractions. Any abstraction in the reduct of a term is a descendant of a term in the original term. If some non-duplicating abstraction in the original term becomes duplicating in the reduct then there must be a duplicating abstraction in the body of the original abstraction and this forces the new duplicating abstraction to contain or be contained in all other duplicating abstractions in the reduct. Moreover, any non-disjoint duplicating abstractions in the original term remain non-disjoint in the reduct. Thus, if a term does not contain disjoint duplicating abstractions, neither does its reduct.

8.4. LEMMA. *If $M \in \Lambda_E$ and $M \to N$ then $N \in \Lambda_E$.*

PROOF. By induction on the derivation of $M \to N$, using Lemmas 8.3(ii) and 8.3(iii). □

The analogue of Proposition 6.6 for Λ_E requires the following definition and lemma. The map $\|\bullet\|_\omega$ counts the number of duplicating abstractions in a term.

8.5. DEFINITION. Define the map $\| \bullet \|_\omega : \Lambda \to I\!N$ as follows.

$$\begin{array}{lll}
(1) \quad \|x\|_\omega & = & 0 \\[2ex]
(2) \quad \|\lambda x.M\|_\omega & = & \left\{ \begin{array}{ll} \|M\|_\omega + 1 & \text{If } \|x\|_(, M) > 1 \\ \|M\|_\omega & \text{Otherwise} \end{array} \right. \\[2ex]
(3) \quad \|M\ N\|_\omega & = & \|M\|_\omega + \|N\|_\omega
\end{array}$$

8.6. LEMMA.

(i) $M \in \Lambda_E, N \subseteq M \Rightarrow N \in \Lambda_E$.

(ii) $\|P\{x := Q\}\|_\omega = \|P\|_\omega + \|x\|_{(}, P) \cdot \|Q\|_\omega$.

(iii) $N \subseteq M \Rightarrow \|N\|_\omega \leq \|M\|_\omega$

(iv) $M \in \Lambda_A \Rightarrow \|M\|_\omega = 0$.

PROOF.

(i) By induction on the derivation of $N \subseteq M$, using Lemma 8.3(i).

(ii) By induction on P.

(iii) By induction on the derivation of $N \subseteq M$.

(iv) By induction on the derivation of $M \in \Lambda_A$.□

The following result states that all λE-terms are strongly normalizing. The intuition behind this is as follows.

Firstly, whenever a redex with non-duplicating operator is contracted, no new duplicating abstractions are created. Moreover, the size of the reduct is strictly smaller than the size of the original term. Thus the reduction step decreases the lexicographically ordered pair consisting of the number of duplicating abstractions in the term and the size of the term.

Secondly, whenever a redex with duplicating operator is contracted, one duplicating abstraction is removed, and any new duplicating abstractions have to come either from proliferation of duplicating abstractions in the argument of the redex or from duplication of variables in the body of some abstraction.

The first case is impossible, since it requires two disjoint duplicating abstractions: the operator and one inside the argument.

In the second case, any new duplicating abstractions must have their λ outside the redex. When such a redex is contracted in a term with an infinite reduction the strategy F will continue to contract redexes under this λ. Thus, if we only consider the part of the term where redexes will be contracted by F then every reduction step in a λE-term with an infinite reduction path strictly decreases the lexicographically ordered pair consisting of the number of duplicating abstractions in the term and the size of the term, but then the term must actually be strongly normalizing.

8.7. PROPOSITION. $M \in \Lambda_E \Rightarrow M \in \mathrm{SN}$.

PROOF. Suppose $M \in \Lambda_E$ and $\infty(M)$. By induction on i using Proposition 7.4, there is an infinite reduction path

$$\sigma : M \equiv M_0 \to M_1 \to \dots$$

such that for all $i \in I\!N$, $F(M_i) = M_{i+1}$. By induction on i using Lemma 8.4 one easily shows that for all $i \in I\!N$, $M_i \in \Lambda_E$.

We now claim that for all $i \in I\!N$ (using the lexicographical ordering)

$$(1) \quad (\|V(M_i)\|_\omega, \|V(M_i)\|) > (\|V(M_{i+1})\|_\omega, \|V(M_{i+1})\|)$$

To prove this, first note that by Lemma 7.8 and Remark 7.7.

$$\begin{array}{llcll} (2) & V(M_i) & = & (\lambda x.Q_0)\, Q_1 \ldots Q_n & \subseteq M_i \quad (n \geq 1) \\ (3) & V(M_{i+1}) & \subseteq & Q_0\{x := Q_1\}\, Q_2 \ldots Q_n & \quad\quad\;\; (n \geq 1) \end{array}$$

for some $Q_0, \ldots, Q_n$. Since $(\lambda x.Q_0)\, Q_1 \subseteq M_i \in \Lambda_E$, Lemma 8.6(i) implies $(\lambda x.Q_0)\, Q_1 \in \Lambda_E$. Now consider the following two cases.

Case 1. $\|x\|_{(}, Q_0) > 1$. Then $\lambda x.Q_0 \in \Lambda_E \backslash \Lambda_A$. Then, by definition of Λ_E, $Q_1 \in \Lambda_A$, and so by Lemma 8.6(iv) $\|Q_1\|_\omega = 0$. By (2)-(3) and Lemma 8.6(ii)-(iii) we then have

$$\begin{array}{lcl} \|V(M_{i+1})\|_\omega & \leq & \|Q_0\{x := Q_1\}\, Q_2 \ldots Q_n\|_\omega \\ & = & \|Q_0\|_\omega + \|x\|_{(}, Q_0) \cdot \|Q_1\|_\omega + \|Q_2\|_\omega + \ldots + \|Q_n\|_\omega \\ & = & \|Q_0\|_\omega + \|Q_2\|_\omega + \ldots + \|Q_n\|_\omega \\ & < & \|Q_0\|_\omega + \ldots + \|Q_n\|_\omega + 1 \\ & = & \|(\lambda x.Q_0)\, Q_1 \ldots Q_n\|_\omega \\ & = & \|V(M_i)\|_\omega \end{array}$$

Case 2. $\|x\|_{(}, Q_0) \leq 1$. Then by (2)-(3) and Lemma 8.6(ii)-(iii) we have

$$\begin{array}{lcl} \|V(M_{i+1})\|_\omega & \leq & \|Q_0\{x := Q_1\}\, Q_2 \ldots Q_n\|_\omega \\ & = & \|Q_0\|_\omega + \|x\|_{(}, Q_0) \cdot \|Q_1\|_\omega + \|Q_2\|_\omega + \ldots + \|Q_n\|_\omega \\ & \leq & \|Q_0\|_\omega + \ldots + \|Q_n\|_\omega \\ & = & \|(\lambda x.Q_0)\, Q_1 \ldots Q_n\|_\omega \\ & = & \|V(M_i)\|_\omega \end{array}$$

and by (2)-(3), Proposition 3.3(i), and Lemma 2.10(i) we have

$$\begin{array}{lcl} \|V(P_{i+1})\| & \leq & \|Q_0\{x := Q_1\}\, Q_2 \ldots Q_n\| \\ & = & \|Q_0\| + \|x\|_{(}, Q_0) \cdot (\|Q_1\| - 1) + \|Q_2\| + \ldots + \|Q_n\| + n - 1 \\ & \leq & \|Q_0\| + \ldots + \|Q_n\| + n - 2 \\ & < & \|Q_0\| + \ldots + \|Q_n\| + n + 1 \\ & = & \|(\lambda x.Q_0)\, Q_1 \ldots Q_n\| \\ & = & \|V(P_i)\| \end{array}$$

as required, concluding the proof of (1).

However, (1) implies that we have an infinite sequence

$$(\|V(M)\|_\omega, \|V(M)\|) = (\|V(M_0)\|_\omega, \|V(M_0)\|) > (\|V(M_1)\|_\omega, \|V(M_1)\|) > \ldots$$

This is clearly a contradiction. Thus $M \in \mathrm{SN}$.□

8.8. Corollary.

(i) $\infty(M) \Rightarrow \mathbf{\Omega} \trianglelefteq M$.

(ii) $\infty(M) \Rightarrow \|\mathbf{\Omega}\| \leq \|M\|$.

(iii) $(\forall N \trianglelefteq M : N \in \mathrm{WN}) \Rightarrow M \in \mathrm{SN}$.

Proof.

(i) By Lemma 8.2 and Proposition 8.7.

(ii) By (i) and Proposition 3.7(i).

(iii) We show the contrapositive: $\infty(M) \Rightarrow \exists N \trianglelefteq M : \neg N \in \mathrm{WN}$. If $\infty(M)$ then by (i) $\boldsymbol{\Omega} \trianglelefteq M$, and $\boldsymbol{\Omega} \notin \mathrm{WN}$.□

8.9. Remark. $\infty(M) \not\Leftarrow \boldsymbol{\Omega} \trianglelefteq M$ as the example $(\lambda x.y\ x\ x)\ (\lambda x.y\ x\ x)$ shows. This example also shows that $\Leftarrow$ does not hold in any of the assertions in Corollary 8.8.

9. Strong embedding and λI-terms

9.1. Definition. The set of λI-terms, Λ_I is defined as follows.

$$\begin{array}{llcl} (1) & x \in \Lambda_I & & \\ (2) & P \in \Lambda_I, \|x\|_{(}, P) \geq 1 & \Rightarrow & \lambda x.P \in \Lambda_I \\ (3) & P, Q \in \Lambda_I & \Rightarrow & P\,Q \in \Lambda_I \end{array}$$

The essential syntactical property of Λ_I is that it is the set of λ-terms for which no term has a subterm which is an erasing abstraction. This explains the following result.

9.2. Lemma. $\forall M \in \Lambda : M \notin \Lambda_I \Rightarrow \lambda x.y \trianglelefteq M$ *for some variable* $y \not\equiv x$.

Proof. Induction on M. In the case of an abstraction use the following property: for all $M \in \Lambda$ there is a variable y such that $y \trianglelefteq M$. □

9.3. Proposition. *If* $\infty(M)$ *and* $M \in \mathrm{WN}$ *then* $\lambda x.y \trianglelefteq M$ *for some variable* $y \not\equiv x$.

Proof. If $\lambda x.y \not\trianglelefteq M$ for all variables $y \not\equiv x$ then $M \in \Lambda_I$ by Lemma 9.2. Since $M \in \mathrm{WN}$ the Conservation Theorem for λI [1, 9.1.5] implies that $M \in \mathrm{SN}$, but this contradicts $\infty(M)$. □

10. Open problems

10.1. Problem. Is the following assertion true? Suppose $M_1, M_2, \ldots$ is a sequence of λ-terms such that the set of all variables occurring free in at least one M_i is finite. Then there are $i, j \in \mathbb{N}$ such that $i < j$ and $M_i \trianglelefteq M_j$.

10.2. Problem. Can results for self-embeddings in term rewrite systems be translated into interesting results for λ-calculus and combinatory logic? For instance, Plaisted [19] shows that it is undecidable whether a term rewrite system is self-embedding, but the proof does not work, right away, for λ-calculus or combinatory logic.

10.3. Problem. The so-called *pumping lemma* in the theory of context-free languages states, essentially, that if a grammar describes an infinite language of strings, then there must be a certain form of cyclicity in the grammar. Are there any interesting connections between this result and Kruskal's Tree Theorem?

10.4. Problem. There is an amazing series of papers by N. Robertson and P.D. Seymour in the Journal of Combinatorial Theory, series (B), concerning K. Wagner's conjecture that for every sequence $G_1, G_2, \ldots$ of (undirected) graphs there are $i, j \in \mathbb{N}$ with $i < j$ such that G_i is isomorphic to a *minor* of G_j (where a graph is a minor of another if the first can be obtained from a sub-graph of the second by edge-contraction). For some classes of graphs the conjecture has been proven, see [21, 17].

The translation of Kruskal's Theorem to λ-calculus comes, essentially, by viewing λ-terms as trees; for instance nameless λ-terms [5] can be viewed as trees where the leaf nodes contain numbers. Replacing in these trees the numbers in leaf nodes by arcs to the corresponding λ instead, λ-terms are certain (directed) graphs. Does Wagner's conjecture give an interesting characterization of infinite sequences of λ-terms, when the latter are viewed as undirected graphs?

10.5. Problem. Can one constructively define a relation $\preceq$ on Λ such that

(1) $(\forall N \preceq M : N \in \mathrm{WN}) \Rightarrow M \in \mathrm{SN}$
(2) For any PTS λS, if M has a type in λS and $N \preceq M$ then N has a type in λS.

The strong embedding relation does not satisfy (2). For instance, in the simply typed λ-calculus (à la Curry [2, Ch.4]) $\lambda x.x\ (x\ \lambda y.y)$ has type $((\alpha \to \alpha) \to (\alpha \to \alpha)) \to (\alpha \to \alpha)$, but $\lambda x.x\ x \trianglelefteq \lambda x.x\ (x\ \lambda y.y)$ has type no type.

A previous attempt has been refuted by Gonthier. The relation was:

$$M \preceq N \quad \Leftrightarrow \quad \begin{array}{l} M \to N \\ \text{or } M \text{ arises from } N \text{ by replacing a subterm by a variable} \\ \text{or } M \subseteq N \end{array}$$

The counter-example is the term $w\,w$ where $w \equiv \lambda x.(\lambda yz.z)\,x\,u\,x\,x$. All $N \preceq w\,w$ are weakly normalizing, but $w\,w$ itself is not strongly normalizing.

10.6. Problem. Is the following true? Every M in Λ_E can be typed in second-order typed λ-calculus (à la Curry [2, Ch.4]).

References

[1] H.P. Barendregt. *The Lambda Calculus: Its Syntax and Semantics.* North-Holland, second, revised edition, 1984.

[2] H.P. Barendregt. Lambda calculi with types. In S. Abramsky, D.M. Gabbay, and T.S.E. Maibaum, editors, *Handbook of Logic in Computer Science*, volume II, pages 117–309. Oxford University Press, 1992.

[3] J.A. Bergstra and J.W. Klop. Strong normalization and perpetual reductions in the lambda calculus. *Journal of Information Processing and Cybernetics*, 18:403–417, 1982.

[4] C. Böhm and S. Micali. Minimal forms in λ-calculus computations. *Journal of Symbolic Logic*, 45(1):165–171, 1980.

[5] N.G. de Bruijn. Lambda calculus notation with nameless dummies, a tool for automatic formula manipulation. *Indagitiones Mathematica*, 34:381–392, 1972.

[6] N. Dershowitz. Termination of rewriting. *Journal of Symbolic Computation*, 3, 1987.

[7] P. Halmos. *Naive Set Theory*. Springer-Verlag, undergraduate texts in mathematics, 1974 edition, 1960.

[8] G. Higman. Ordering by divisibility in abstract algebras. *Proceedings of the London Mathematical Society*, (3) 2:326–336, 1952.

[9] J.R. Hindley. BCK-combinators and linear λ-terms have types. *Theoretical Computer Science*, 64:97–105, 1989.

[10] S. Hirokawa. Some properties of one-step recurrent terms in lambda calculus. *Mathematica Japonica*, 6:927–935, 1987.

[11] B. Jacobs. Semantics of lambda-I and of other substructure calculi. In M. Bezem and J.F. Groote, editors, *Typed Lambda Calculus and Applications*, volume 664 of *Lecture Notes in Computer Science*, pages 195–208. Springer-Verlag, 1993.

[12] J.W. Klop. *Combinatory Reduction Systems*. PhD thesis, Utrecht University, 1980. CWI Tract, Amsterdam.

[13] Y. Komori. BCK-algebras and lambda calculus. In *Proceedings of the 10th Symposium on Semigroups*, pages 5–11. Josai University, Sakado, 1987.

[14] J.B. Kruskal. Well-quasi-ordering, the tree theorem, and Vazsonyi's conjecture. *Transactions of the American Mathematical Society*, 95:210–225, 1954.

[15] J.B. Kruskal. The theory of well-quasi-ordering: A frequently discovered concept. *Journal of Combinatorial Theory, Series (A)*, 13:297–305, 1972.

[16] B. Lercher. Lambda-calculus terms that reduce to themselves. *Notre Dame Journal of Formal Logic*, XVII(2):291–292, 1976.

[17] W. Mader. Wohlquasigeordnete klassen endlicher graphen. *Journal of Combinatorial Theory, Series (B)*, 12:105–122, 1972.

[18] C.St.J.A. Nash-Williams. On well-quasi-ordering finite trees. *Proceedings of the Cambridge Mathematical Society*, 59:833–835, 1963.

[19] D.A. Plaisted. The undecidability of self-embedding for term-rewriting systems. *Information Processing Letters*, 20:61–64, 1985.

[20] G. Plotkin. Call-by-name, call-by-value and the λ-calculus. *Theoretical Computer Science*, 1:125–159, 1975.

[21] N. Robertson and P.D. Seymour. Graph minors. IV tree-width and well-quasi-ordering. *Journal of Combinatorial Theory, Series (B)*, 48:227–254, 1990.

[22] S. Sekimoto and S. Hirokawa. One-step recurrent terms in $\lambda - \beta$-calculus. *Theoretical Computer Science*, 56:223–231, 1988.

[23] M.V. Zilli. Head-recurrent terms in combinatory logic: a generalization of the notion of head normal form. In G. Ausiello and C. Böhm, editors, *Automata, Languages, and Programming*, volume 62 of *Lecture Notes in Computer Science*, pages 477–493. Springer-Verlag, 1975.

[24] M.V. Zilli. Reduction graphs in the untyped lambda calculus. *Theoretical Computer Science*, 29:251–275, 1984.