1

Properties of Infinite Reduction Paths in Untyped λ-Calculus

Morten Heine Sørensen

1.1 Introduction

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

$$\mathbf{\Omega} \to \mathbf{\Omega} \to \dots$$

There are terms that do have an infinite reduction path, but where the path does not have this simple form.[1] For instance, $\Psi \equiv \psi\, \psi$, where $\psi \equiv \lambda x.x\, x\, y$, has the infinite reduction path:

$$\Psi \to \Psi\, y \to \Psi\, y\, y \to \dots$$

In every step the redex Ψ appears as a subterm, and the context of the redex is extended with an application $\bullet\, y$. As a more complicated example consider the term $v\, y\, v$, where $v \equiv \lambda ax.x\, (a\, y)\, x$:

$$v\,y\,v \to (\lambda x.x\,(y\,y)\,x)\,v \to v\,(y\,y)\,v \to (\lambda x.x\,(y\,y\,y)\,x)\,v \to v\,(y\,y\,y)\,v \to \dots$$

This path is similar to the preceding one, but the extra application $\bullet\, y$ is added *inside* the redex.

Although these three reduction paths have their differences they have two common properties. First, in all three paths every term has $\mathbf{\Omega}$ as a substring. Second, in all three paths it happens infinitely often that a term is followed (after a number of steps) by another term of which the former is a substring.

It is natural to ask whether these properties are shared by all infinite reduction paths. The main concern of this paper is to formalize the two

[1]Lercher (1976) shows that $M \to M$ iff $M \equiv C[\mathbf{\Omega}]$ for some context C.

observations and to show that, indeed, they are correct for all infinite reduction paths. More specifically, the contributions of the paper are:

1. Formalization and proof of the first observation. Corollaries include a technique to deduce strong normalization from weak normalization, and some results that previously relied on tedious case analyses. For instance, $\mathbf{\Omega}$ is the only term with an infinite reduction path among the 2622 closed terms of size 9 or less.
2. Formalization and proof of the second observation. The result is an analog of Kruskal's Tree Theorem (Higman 1952, Kruskal 1954, Nash-Williams 1963) which has many applications in functional and logic programming; we sketch an application in λ-calculus.

Section 2 introduces labeled terms and embeddings. Sections 3-6 are concerned with (1), and Sections 7-8 are concerned with (2). We end this section with some preliminary definitions; see also Barendregt (1984).

Λ is the set of λ-terms. We use vector notation $\vec{P}$ for a sequence of terms $P_1 \ldots P_n$, e.g., $Q\,\vec{P}$ for $Q\,P_1 \ldots P_n$, and $\vec{P} \in S$ for $P_1, \ldots, P_n \in S$. We assume familiarity with substitution, the variable convention, and notions of reduction $\mathbf{R}$ on Λ, in particular β-reduction. We denote by $\rightarrow_R, \twoheadrightarrow_R$, and $\twoheadrightarrow_R^+$ the compatible closure of $\mathbf{R}$, the compatible, reflexive, transitive closure of $\mathbf{R}$, and the compatible, transitive closure of $\mathbf{R}$, respectively. Both the subterm and subset relation are denoted by $\subseteq$.

Definition 1 Let $\mathbf{R}$ be a notion of reduction. An *R-reduction path* from M_0 is a finite or infinite sequence $M_0 \rightarrow_R M_1 \rightarrow_R \ldots$. If the sequence is finite it *ends* in the last term M_n.

Definition 2 Let $\mathbf{R}$ be a notion of reduction.

1. $\infty_R(M) \Leftrightarrow M$ has an infinite R-reduction path.
2. $\mathrm{NF}_R(M) \Leftrightarrow M$ has no R-reduction path of length 1 or more.
3. $\mathrm{SN}_R(M) \Leftrightarrow$ All R-reduction paths from M are finite.
4. $\mathrm{WN}_R(M) \Leftrightarrow M$ has an R-reduction path ending in N with $\mathrm{NF}_R(N)$.

Instead of, e.g., $\mathrm{SN}_R(M)$ we often write $M \in \mathrm{SN}_R$. Elements of $\mathrm{NF}_R, \mathrm{SN}_R, \mathrm{WN}_R$ are *R-normal forms*, *R-strongly normalizing*, and *R-weakly normalizing*, respectively. When the index R is β we omit it.

Definition 3

1. $\|M\|$ is the size of M (the number of variables, abstractions, and applications).
2. $\|M\|_x$ is the number of free occurrences of x in M.
3. $\lambda x.M$ is *duplicating* if $\|M\|_x > 1$, and *non-duplicating* otherwise.
4. $\|M\|_\omega$ is the number of duplicating abstractions in M.

1.2 Labeled terms and an embedding relation

To formalize what it means that one term is a substring of another we have to deal with certain pathological situations. It seems fair to say that $\lambda x.x\,x$ is a substring of $\lambda x.x\,(\lambda y.x)$, but is the former term also a substring of $\lambda x.x\,(\lambda x.x)$? To deal with such questions without relying on subtle consequences of variable conventions we introduce labeled terms.

Definition 4 Let R be a binary relation on a set S.

1. R is a *quasi-order* (qo) on S iff R is transitive and reflexive.
2. R is a *well-quasi order* (wqo) on S iff R is a qo on S and for every infinite sequence $s_1, s_2, \ldots$ of elements in S there are $i, j \in I\!N$ such that $i < j$ and $s_i R s_j$.

Example 1 The identity relation $=$ is a wqo on $\{1, \ldots, n\}$, for any n.

Definition 5 Let $\mathcal{L}$ be a countable set with a wqo $\preceq$.

1. Define labeled terms $\Lambda^{\mathcal{L}}$ as follows.

$$\begin{array}{lll} (1)\quad l \in \mathcal{L} & \Rightarrow & x^l \in \Lambda^{\mathcal{L}} \\ (2)\quad l \in \mathcal{L}, P \in \Lambda^{\mathcal{L}} & \Rightarrow & (\lambda x.P)^l \in \Lambda^{\mathcal{L}} \\ (3)\quad l \in \mathcal{L}, P, Q \in \Lambda^{\mathcal{L}} & \Rightarrow & (P\,Q)^l \in \Lambda^{\mathcal{L}} \end{array}$$

2. β-reduction on $\Lambda^{\mathcal{L}}$ is defined by: $((\lambda x.P)^l\,Q)^k\,\beta\,P\{x := Q\}$, where

$$\begin{array}{lll} (1)\quad x^l\{x := P\} & \equiv & P \\ (2)\quad y^l\{x := P\} & \equiv & y^l \\ (3)\quad (\lambda y.Q)^l\{x := P\} & \equiv & (\lambda y.Q\{x := P\})^l \\ (3)\quad (R\,Q)^l\{x := P\} & \equiv & (R\{x := P\}\,Q\{x := P\})^l \end{array}$$

3. For $M' \in \Lambda^{\mathcal{L}}$, $|M'| \in \Lambda$ is the result of erasing all labels. If $|M'| \equiv M$ then M' is a *labeling* of M. If all subterms of M' have different labels then M' is an *initial* labeling. If all subterms except variables have different labels, and variables have the same label as their binding abstraction, then M' is a *structural* labeling.

Example 2 We write $\lambda x^l.M$ and $M \bullet^l N$ rather than $(\lambda x.P)^l$ and $(PQ)^l$. Let $\mathcal{L} = \{1, \ldots, 9\}$ with wqo $=$. Then in $\Lambda^{\mathcal{L}}$:

1. $(\lambda x^1.x^2 \bullet^3 x^4) \bullet^5 (\lambda x^6.x^7 \bullet^8 x^9)$ is an initial labeling of $\mathbf{\Omega}$.
2. $(\lambda x^1.x^1 \bullet^2 x^1) \bullet^3 (\lambda x^4.x^4 \bullet^5 x^4)$ is a structural labeling of $\mathbf{\Omega}$.
3. $\lambda x^1.x^1 \bullet^2 x^3$ is neither a structural nor an intial labeling of ω.

Definition 6 Let $\mathcal{L}$ be countable with wqo $\preceq$. Define *embedding* $\sqsubseteq_{\mathcal{L}}$ on $\Lambda^{\mathcal{L}}$:

$$
\begin{array}{lll}
(1) & P \sqsubseteq_{\mathcal{L}} Q & \Rightarrow & P \sqsubseteq_{\mathcal{L}} (\lambda x.Q)^l \\
(2) & P \sqsubseteq_{\mathcal{L}} Q & \Rightarrow & P \sqsubseteq_{\mathcal{L}} (Q\,Z)^l \\
(3) & P \sqsubseteq_{\mathcal{L}} Q & \Rightarrow & P \sqsubseteq_{\mathcal{L}} (Z\,Q)^l \\
(4) & k \preceq l & \Rightarrow & x^k \sqsubseteq_{\mathcal{L}} y^l \\
(5) & k \preceq l, P \sqsubseteq_{\mathcal{L}} Q & \Rightarrow & (\lambda x.P)^k \sqsubseteq_{\mathcal{L}} (\lambda y.Q)^l \\
(6) & k \preceq l, P_1 \sqsubseteq_{\mathcal{L}} Q_1, P_2 \sqsubseteq_{\mathcal{L}} Q_2 & \Rightarrow & (P_1\,P_2)^k \sqsubseteq_{\mathcal{L}} (Q_1\,Q_2)^l
\end{array}
$$

Example 3 Let $\mathcal{L} = \{1, \ldots, 9\}$ with wqo $=$. Then in $\Lambda^{\mathcal{L}}$:

1. $\lambda x^1.x^1 \sqsubseteq_{\mathcal{L}} \lambda x^1.x^1 \bullet^2 x^1 \sqsubseteq_{\mathcal{L}} (\lambda x^1.x^1 \bullet^2 x^1) \bullet^3 (\lambda x^4.x^4 \bullet^5 x^4)$.
2. $\lambda x^1.\lambda y^2.x^1 \not\sqsubseteq_{\mathcal{L}} \lambda x^1.\lambda y^2.y^2$.
3. $(\lambda x^1.x^2 \bullet^3 x^4) \bullet^5 (\lambda x^6.x^7 \bullet^8 x^9) \not\sqsubseteq_{\mathcal{L}} (\lambda x^6.x^7 \bullet^8 x^9) \bullet^3 (\lambda x^6.x^7 \bullet^8 x^9)$.
4. $(\lambda x^6.x^7 \bullet^8 x^9) \bullet^3 (\lambda x^6.x^7 \bullet^8 x^9) \not\sqsubseteq_{\mathcal{L}} (\lambda x^6.x^7 \bullet^8 x^9) \bullet^8 (\lambda x^6.x^7 \bullet^8 x^9)$

The following is easily established.

Proposition 1 *Let $\mathcal{L}$ be countable with wqo $\preceq$.*

1. *$M \sqsubseteq_{\mathcal{L}} N$ implies $\|M\| \leq \|N\|$.*
2. *$\sqsubseteq_{\mathcal{L}}$ is a quasi order.*

1.3 Terms in which $\mathbf{\Omega}$ is not a substring

To prove that $\infty(M)$ implies that $\mathbf{\Omega}$ is a substring of M we define a notion of substring on unlabeled terms. Then we define the set of all terms in which $\mathbf{\Omega}$ is not a substring, and show that all such terms are strongly normalizing. This occupies Sections 3-5. Section 6 gives applications.

Definition 7 For $M, N \in \Lambda$, let $n = \max\{\|M\|, \|N\|\}$, and $\mathcal{L} = \{1, \ldots, n\}$ with wqo $=$. Then define $M \trianglelefteq N$ iff $M' \sqsubseteq_{\mathcal{L}} N'$ for some structural labelings M', N' of M and N.

Example 4 $\lambda x.x\,x \trianglelefteq \lambda x.x\,(\lambda y.x)$, but $\lambda x.x\,x \not\trianglelefteq \lambda x.x\,(\lambda x.x)$.

Definition 8 [Komori 1987,Hindley 1989,Jacobs 1993] Define Λ_ω by:

$$
\begin{array}{llll}
(1) & x \in \Lambda_\omega & & \\
(2) & P \in \Lambda_\omega, \|P\|_x \leq 1 & \Rightarrow & \lambda x.P \in \Lambda_\omega \\
(3) & P, Q \in \Lambda_\omega & \Rightarrow & P\,Q \in \Lambda_\omega
\end{array}
$$

Remark 1 The following equivalences are easily established:

$$\|M\|_\omega = 0 \Leftrightarrow M \in \Lambda_\omega \Leftrightarrow \omega \not\trianglelefteq M$$

One easily shows that Λ_ω is closed under reduction. The intuition is that if $M \in \Lambda_\omega$ and $N \notin \Lambda_\omega$ then M has no duplicating abstractions while N has at least one. Thus the reduction $M \to N$ must duplicate a variable in the body of some abstraction, but this would require a duplicating abstraction in M.

It is also easy to prove that reduction in Λ_ω decreases term size, since every step removes an application and an abstraction. With the preceding property this implies that every term in Λ_ω is strongly normalizing.

Definition 9 Define the set $\Lambda_{\mathbf{\Omega}}$ as follows.

$$\begin{array}{lllll}
(1) & x \in \Lambda_{\mathbf{\Omega}} & & \\
(2) & M \in \Lambda_{\mathbf{\Omega}} & \Rightarrow & \lambda x.M \in \Lambda_{\mathbf{\Omega}} \\
(3) & M \in \Lambda_{\mathbf{\Omega}}, N \in \Lambda_\omega & \Rightarrow & M\ N \in \Lambda_{\mathbf{\Omega}} \\
(4) & M \in \Lambda_\omega, N \in \Lambda_{\mathbf{\Omega}} & \Rightarrow & M\ N \in \Lambda_{\mathbf{\Omega}}
\end{array}$$

Remark 2 It is easy to show $\Lambda_\omega \subseteq \Lambda_{\mathbf{\Omega}}$ and the following equivalence:

$$M \in \Lambda_{\mathbf{\Omega}} \Leftrightarrow \mathbf{\Omega} \not\trianglelefteq M$$

Next we show that $\Lambda_{\mathbf{\Omega}}$ is closed under reduction. The intuition is as follows. If $M \in \Lambda_{\mathbf{\Omega}}$ and $N \notin \Lambda_{\mathbf{\Omega}}$ then M has no disjoint duplicating abstractions, while N has at least two. If $M \to N$ then non-disjoint duplicating abstractions in M are also non-disjoint in N. Therefore, the two disjoint duplicating abstractions in N must arise from M either by duplication into disjoint positions of a single duplicating abstraction, or by duplication of a variable in the body of a non-duplicating abstraction which is disjoint with a duplicating abstraction. Both cases are impossible because they entail that M has two disjoint duplicating abstractions.

Lemma 2 $M \in \Lambda_{\mathbf{\Omega}}, M \to N \;\Rightarrow\; N \in \Lambda_{\mathbf{\Omega}}$.

Proof. First prove by induction on the derivation of $M \in \Lambda_\omega$ that

(1) $$M \in \Lambda_\omega, N \in \Lambda_\omega \;\Rightarrow\; M\{x := N\} \in \Lambda_\omega$$

(2) $$M \in \Lambda_\omega, N \in \Lambda_{\mathbf{\Omega}}, \|M\|_x \leq 1 \;\Rightarrow\; M\{x := N\} \in \Lambda_{\mathbf{\Omega}}$$

Show by induction on the derivation of $M \in \Lambda_{\mathbf{\Omega}}$, using (1) and $\Lambda_\omega \subseteq \Lambda_{\mathbf{\Omega}}$

(3) $$M \in \Lambda_{\mathbf{\Omega}}, N \in \Lambda_\omega \;\Rightarrow\; M\{x := N\} \in \Lambda_{\mathbf{\Omega}}$$

Then proceed by induction on the derivation of $M \to N$ using (2-3). $\square$

As for Λ_ω, the idea for proving that all terms in $\Lambda_{\mathbf{\Omega}}$ are strongly normalizing is to find a decreasing measure, but $\| \bullet \|$ does not work. Instead we consider the lexicographically ordered measure $\langle \| \bullet \|_\omega, \| \bullet \| \rangle$.

Suppose $M \to N$ by contraction of the redex $\Delta \equiv (\lambda x.P)\ Q$. If $\lambda x.P$ is non-duplicating, contraction of Δ creates no new duplicating abstractions. Moreover, the size of N is strictly smaller than the size of M, so the reduction step decreases the measure.

If $\lambda x.P$ is duplicating, the reduction step removes one duplicating abstraction, and any new duplicating abstractions have to come either from proliferation of duplicating abstractions in Q or from duplication of variables in the body of some abstraction. The first case is impossible, since it implies that M has two disjoint duplicating abstractions. In the second case, new duplicating abstractions *may* be created, but they must have their λ to the left of Δ.

Recall that *standard* reductions $M_0 \to M_1 \to \ldots$ are such that whenever a redex Δ is contracted in M_i all abstractions to the left of Δ are marked, and a redex with marked abstraction is not allowed to be contracted in M_{i+1}. A classical result due to Curry and Feys (1958) states that if $M \twoheadrightarrow N$ then there is a standard reduction from M to N. Bergstra and Klop (1982) show that if some term has an infinite reduction then it also has a standard infinite reduction.

The idea then is as follows. Suppose some $M \in \Lambda_{\Omega}$ has an infinite reduction and hence an infinite standard reduction. Then the measure $\langle \|\bullet\|_\omega, \|\bullet\| \rangle$ is decreasing on this reduction path if we change $\|\bullet\|_\omega$ to only count non-marked abstractions, and thus we arrive at a contradiction. This is developed in detail in Sections 4-5.

1.4 A perpetual strategy computing standard reductions

Definition 10 (Barendregt et al. 1976) Let $F : \Lambda \to \Lambda$ be a map.

1. F is a *one-step reduction strategy* iff
$$\begin{array}{lcl} M \notin \mathrm{NF} & \Rightarrow & M \to F(M) \\ M \in \mathrm{NF} & \Rightarrow & M \equiv F(M) \end{array}$$
2. F is *perpetual* iff $\infty(M) \Rightarrow \infty(F(M))$.

A well-known strategy is F_l which reduces the leftmost redex (see Barendregt 1984). Although it has not previously appeared in the literature, to the best of our knowledge, the following strategy seems to be the simplest one computing infinite standard reductions paths.

Definition 11 Define $\underline{\bullet} : \Lambda \to \Lambda$ as follows.[2] If $M \in \mathrm{SN}$ then $\underline{M} = F_l(M)$; otherwise,

$$\begin{array}{lcll} \underline{x\,\vec{P}\,Q\,\vec{R}} & = & x\,\vec{P}\,\underline{Q}\,\vec{R} & \text{If } \vec{P} \in \mathrm{SN}, Q \notin \mathrm{SN} \\ \underline{\lambda x.P} & = & \lambda x.\underline{P} & \\ \underline{(\lambda x.P)\,Q\,\vec{R}} & = & P\{x := Q\}\,\vec{R} & \text{If } P, Q, \vec{R} \in \mathrm{SN} \\ \underline{(\lambda x.P)\,Q\,\vec{R}} & = & (\lambda x.\underline{P})\,Q\,\vec{R} & \text{If } P \notin \mathrm{SN} \\ \underline{(\lambda x.P)\,\vec{Q}\,R\,\vec{S}} & = & (\lambda x.P)\,\vec{Q}\,\underline{R}\,\vec{S} & \text{If } P, \vec{Q} \in \mathrm{SN}, R \notin \mathrm{SN} \end{array}$$

Remark 3 For every $M \in \Lambda$, either $M \in \mathrm{SN}$ or $M \notin \mathrm{SN}$. In the latter case either $M \equiv x\,P_1 \ldots P_n$ where $n \geq 1$ and $P_i \notin \mathrm{SN}$ for some i, or $M \equiv \lambda x.P$, or $M \equiv (\lambda x.P_0)\,P_1 \ldots P_n$ where $n \geq 1$. It follows that $\underline{\bullet}$ is defined on all λ-terms.

[2] The definition $\underline{M} = F_l(M)$ for $M \in \mathrm{SN}$ is an inessential technicality to make $\underline{\bullet}$ total on Λ and thereby a reduction strategy.

Proposition 3 *F is a perpetual one-step reduction strategy.*

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

$$\begin{array}{llcl} (1) & M \in \mathrm{NF} & \Rightarrow & M \equiv \underline{M} \\ (2) & M \notin \mathrm{NF} & \Rightarrow & M \to \underline{M} \\ (3) & \infty(M) & \Rightarrow & \infty(\underline{M}) \end{array}$$

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 $\infty(M)$ implies both $M \to \underline{M}$ and $\infty(\underline{M})$. We proceed by induction on the size of M, splitting into cases according to Remark 3.

1. $M \equiv x\, \vec{P}\, Q\, \vec{R}$ where $\vec{P} \in \mathrm{SN}, Q \notin \mathrm{SN}$. By induction hypothesis $Q \to \underline{Q}$ and $\infty(\underline{Q})$, so $M \equiv x\, \vec{P}\, Q\, \vec{R} \to x\, \vec{P}\, \underline{Q}\, \vec{R} \equiv \underline{M}$ and $\infty(\underline{M})$.
2. $M \equiv \lambda x.P$. Similar to Case 1.
3. $M \equiv (\lambda x.P)\, Q\, \vec{R}$ where $P, Q, \vec{R} \in \mathrm{SN}$. Then $M \to P\{x := Q\}\, \vec{R} \equiv \underline{M}$. Since $\infty(M)$ but $P, Q, \vec{R} \in \mathrm{SN}$, there must be an infinite reduction path from M of form
$$M \twoheadrightarrow (\lambda x.P')\, Q'\, \vec{R'} \to P'\{x := Q'\}\, \vec{R'} \to \ldots$$
but then there also is an infinite reduction path from $\underline{M}$:
$$\underline{M} \equiv P\{x := Q\}\, \vec{R} \twoheadrightarrow P'\{x := Q'\}\, \vec{R'} \to \ldots$$
4. $M \equiv (\lambda x.P)\, Q\, \vec{R}$ where $P \notin \mathrm{SN}$. Similar to Case 1.
5. $M \equiv (\lambda x.P)\ \vec{Q}\, R\, \vec{S}$ where $P, \vec{Q} \in \mathrm{SN}, R \notin \mathrm{SN}$. Similar to Case 1.

$\square$

The proof that $\underline{\bullet}$ is perpetual is somewhat simpler than related proofs of perpetuality due to Barendregt et al. (1976) and Bergstra and Klop (1982). This is because $\underline{\bullet}$ propagates to subterms with infinite reduction paths as often as possible, in which cases the induction hypothesis applies immediately. On the other hand, unlike the strategy in (Barendregt et al. 1976) F is not computable, but this has no implications for the applications in this paper.

Definition 12 Define $V : \Lambda \to \Lambda$ as follows. If $M \in \mathrm{SN}$ then $V(M) = M$; otherwise,

$$\begin{array}{lcll} V(x\, \vec{P}\, Q\, \vec{R}) & = & V(Q) & \text{If } \vec{P} \in \mathrm{SN}, Q \notin \mathrm{SN} \\ V(\lambda x.P) & = & V(P) & \\ V((\lambda x.P)\, Q\, \vec{R}) & = & (\lambda x.P)\, Q \vec{R} & \text{If } P, Q, \vec{R} \in \mathrm{SN} \\ V((\lambda x.P)\, Q\, \vec{R}) & = & V(P) & \text{If } P \notin \mathrm{SN} \\ V((\lambda x.P)\, \vec{Q}\, R\, \vec{S}) & = & V(R) & \text{If } P, \vec{Q} \in \mathrm{SN}, R \notin \mathrm{SN} \end{array}$$

Remark 4 *For all $M \in \Lambda$, $V(M) \subseteq M$ is well-defined.*

Lemma 4 *For all $M \in \Lambda$, if $\infty(M)$ then $V(M) = (\lambda x.P)\, Q\, \vec{R}$ and $V(\underline{M}) \subseteq P\{x := Q\}\, \vec{R}$ for some $P, Q, \vec{R} \in \Lambda$.*

Proof. Induction on M using Proposition 3 and Remark 3. □

1.5 Strong normalization of terms in $\Lambda_{\boldsymbol{\Omega}}$

The following lemma is easily proved by induction.

Lemma 5 $\|P\{x := Q\}\|_\omega = \|P\|_\omega + \|P\|_x \cdot \|Q\|_\omega$

Proposition 6 $M \in \Lambda_{\boldsymbol{\Omega}} \Rightarrow M \in \mathrm{SN}$.

Proof. Suppose $M \in \Lambda_{\boldsymbol{\Omega}}$ and $\infty(M)$. By Proposition 3, there is an infinite reduction path

$$M_0 \to M_1 \to \dots$$

such that for all i, $\underline{M_i} = M_{i+1}$ and, by Lemma 2, $M_i \in \Lambda_{\boldsymbol{\Omega}}$. We now claim that for all i

$$\langle \|V(M_i)\|_\omega, \|V(M_i)\| \rangle > \langle \|V(M_{i+1})\|_\omega, \|V(M_{i+1})\| \rangle \tag{4}$$

To prove this, first note that by Lemma 4 and Remark 4:

$$V(M_i) \;=\; (\lambda x.P)\, Q\, \vec{R} \quad \subseteq M_i \tag{5}$$

$$V(M_{i+1}) \;\subseteq\; P\{x := Q\}\, \vec{R} \tag{6}$$

for some $P, Q, \vec{R}$. Since $(\lambda x.P)\, Q \subseteq M_i \in \Lambda_{\boldsymbol{\Omega}}$, also $(\lambda x.P)\, Q \in \Lambda_{\boldsymbol{\Omega}}$. We now prove (4) by the following two cases.

a. $\|P\|_x > 1$. Then $\lambda x.P \in \Lambda_{\boldsymbol{\Omega}} \backslash \Lambda_\omega$, so $Q \in \Lambda_\omega$, and hence $\|Q\|_\omega = 0$. By (6), Lemma 5, and (5):

$$\|V(M_{i+1})\|_\omega < \|V(M_i)\|_\omega$$

b. $\|P\|_x \leq 1$. Then by (6), Lemma 5, and (5):

$$\|V(M_{i+1})\|_\omega \leq \|V(M_i)\|_\omega$$

Moreover, by (6) and (5):

$$\|V(M_{i+1})\| < \|V(M_i)\|$$

However, (4) implies that we have an infinite sequence

$$\langle \|V(M_0)\|_\omega, \|V(M_0)\| \rangle > \langle \|V(M_1)\|_\omega, \|V(M_1)\| \rangle > \dots$$

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

1.6 Applications

An S-term in combinatory logic is a term built of only the S-combinator and application, e.g., $S(SS)SSSS$ and $SSS(SS)SS$. Barendregt et al. (1976) show that these two S-terms have infinite reduction paths.

Duboué has verified by computer that the remaining 130 other S-terms with 7 or fewer occurrences of S are strongly normalizing.

The following shows that the first observation in the introduction is correct. It also shows that onely one among the 2622 closed terms of size 9 or less has an infinite reduction sequence.[3]

Corollary 7

1. $\infty(M) \Rightarrow \mathbf{\Omega} \trianglelefteq M$.
2. $\infty(M) \Rightarrow \|\mathbf{\Omega}\| \leq \|M\|$.
3. $\infty(M), \|M\| \leq \|\mathbf{\Omega}\| \Rightarrow M \equiv \mathbf{\Omega}$.

Proof. (1): By Remark 2 and Proposition 6. (2): By (1) and Proposition 1(1). (3): By (1-2) using $O \trianglelefteq M, \|M\| \leq \|O\| \Rightarrow M \equiv O$. □

The next application gives a technique to reduce proofs that some term is strongly normalizing to proofs that terms are weakly normalizing. The latter is usually easier.

Corollary 8 $(\forall N \trianglelefteq M : N \in \mathrm{WN}) \Rightarrow M \in \mathrm{SN}$.

Proof. If $\infty(M)$ then by Corollary 7(1) $\mathbf{\Omega} \trianglelefteq M$, and $\mathbf{\Omega} \notin \mathrm{WN}$. □

The term $M \equiv (\lambda x.y\ x\ x)\ (\lambda x.y\ x\ x)$ shows that $\Leftarrow$ does not hold in Corollary 7(1-2), nor in Corollary 8.

1.7 Kruskal's Tree Theorem on $\Lambda^{\mathcal{L}}$

This section shows that for any infinite sequence of terms in $\Lambda^{\mathcal{L}}$ there is a term and a subsequent term such that the former is embedded in the latter; Section 8 gives applications. The result is analogous to Kruskal's Tree Theorem due to Higman (1952) and in a more general formulation to Kruskal (1954). The result has a beautiful proof due to Nash-Williams (1963). Our proof follows the idea of Nash-William's proof but with some simpifications.

Rather than choosing a definition of embedding of unlabeled terms we deal with $\sqsubseteq_{\mathcal{L}}$ on labeled terms, so in Sections 7-8 all terms are labeled.

Definition 13

1. The set of infinite sequences of terms from $\Lambda^{\mathcal{L}}$ is denoted $\Sigma^*(\Lambda^{\mathcal{L}})$.
2. $(M_0, M_1, \ldots) \in \Sigma^*(\Lambda^{\mathcal{L}})$ is *good* if there are $i, j \in I\!N$ with $i < j$ and $M_i \sqsubseteq_{\mathcal{L}} M_j$. If $\sigma \in \Sigma^*(\Lambda^{\mathcal{L}})$ is not good, it is *bad*.

[3] Mogensen gives a formula $f(n,m)$ for the number of λ-terms of size $n \geq 1$ with at most $m \geq 0$ free variables:

$$\begin{array}{lcl} f(1,m) & = & m \\ f(n+1,m) & = & f(n,m+1) + \sum_{i=1}^{n-1} f(i,v) \cdot f(n-i,v) \end{array}$$

3. $(M_0, M_1, \ldots) \in \Sigma^*(\Lambda^{\mathcal{L}})$ is *ascending* if for all $i \in I\!N$: $M_i \sqsubseteq_{\mathcal{L}} M_{i+1}$.
4. A bad $(M_0, M_1, \ldots) \in \Sigma^*(\Lambda^{\mathcal{L}})$ is *mimimal* if for all $n \geq 0$ there is no bad sequence whose first $n+1$ elements are $M_0, M_1, \ldots, M_{n-1}, M'$ where M' is some term with $\|M'\| < \|M_n\|$.

In proofs of Kruskal's Theorem one uses an analog of the following lemma. In the usual proof of the lemma one uses AC (the axiom of choice) to construct a choice function which is then used to construct the desired minimal sequence. In our setting this choice function can easily be constructed explicitly since one only needs to choose from finite non-empty sets of λ-terms. Since the construction follows a standard set-theoretic technique, the details are usually left out and we will also do so; the technique is explained by Halmos (1960).

Lemma 9 *If* $\sigma \in \Sigma^*(\Lambda^{\mathcal{L}})$ *is bad, then there is a minimal bad* $\tau \in \Sigma^*(\Lambda^{\mathcal{L}})$.

The following lemma gives a small twist to a well-known result (Higman 1952, Nash-Williams 1963).

Lemma 10 *Let* $\Sigma \subseteq \Sigma^*(\Lambda^{\mathcal{L}})$ *be closed under formation of infinite subsequences. Then all* $\sigma \in \Sigma$ *are good iff every* $\sigma \in \Sigma$ *has an infinite ascending subsequence.*

Theorem 11 (Higman 1952, Kruskal 1954) $\sqsubseteq_{\mathcal{L}}$ *is a wqo on* $\Lambda^{\mathcal{L}}$.

Proof.(Nash-Williams 1963) Suppose the theorem were false, i.e., there were a bad sequence. By Lemma 9 there is a minimal bad sequence

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

Clearly σ has an infinite subsequence

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

such that all M_{i_j} are variables, all M_{i_j} are abstractions, or all M_{i_j} are applications. We show that all cases lead to a contradiction.

1. $M_{i_j} \equiv x_j^{l_j}$ for all $j \in I\!N$. Since $\preceq$ is a wqo on $\mathcal{L}$ there are $n < m$ with $l_n \preceq l_m$, so
$$M_{i_n} \equiv x_n^{l_n} \sqsubseteq_{\mathcal{L}} x_m^{l_m} \equiv M_{i_m}$$
contradicting badness of σ.
2. $M_{i_j} \equiv (\lambda x_j.N_j)^{l_j}$ for all $j \in I\!N$. Consider the set $\mathcal{N} = \{N_0, N_1, \ldots\}$.
Subcase 2.1. There is a bad sequence of elements from $\mathcal{N}$. Then in particular there is a bad
$$\sigma' = (N_{j_0}, N_{j_1}, \ldots)$$
where $j_0 < j_m$ for all $m > 0$ (from the former bad sequence simply remove the finitely many elements with an index smaller than the

index of the first element). Then consider

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

Since σ and σ' are bad, $M_i \not\sqsubseteq_{\mathcal{L}} M_j$ (for $0 \leq i < j \leq i_{j_0} - 1$) and $N_{j_n} \not\sqsubseteq_{\mathcal{L}} N_{j_m}$ (for $0 \leq n < m$). Moreover, if $M_i \sqsubseteq_{\mathcal{L}} N_{j_m}$ (where $0 \leq i \leq i_{j_0} - 1, 0 \leq m$) then $M_i \sqsubseteq_{\mathcal{L}} (\lambda x_{j_m}.N_{j_m})^{l_{j_m}} \equiv M_{i_{j_m}}$. Since $j_0 < j_m$ and hence $i < i_{j_0} < i_{j_m}$ this contradicts badness of σ. Thus $M_i \not\sqsubseteq_{\mathcal{L}} N_m$. So σ'' is bad, contradicting minimality of σ.

Subcase 2.2. All infinite sequences of elements from $\mathcal{N}$ are good. Then in particular $(N_1, N_2, \ldots)$ is good, so by Lemma 10 it has an infinite ascending subsequence $(N_{j_0}, N_{j_1}, \ldots)$. Since $\preceq$ is a wqo on $\mathcal{L}$ there are $n < m$ such that $l_{j_n} \preceq l_{j_m}$ and then

$$M_{i_{j_n}} \equiv (\lambda x_{j_n}.N_{j_n})^{l_{j_n}} \sqsubseteq_{\mathcal{L}} (\lambda x_{j_m}.N_{j_m})^{l_{j_m}} \equiv M_{i_{j_m}}$$

contradicting badness of σ.

3. $M_{i_j} \equiv (N_j\ K_j)^{l_j}$ for all $j \in \mathbb{N}$. Consider the sets $\{N_0, N_1, \ldots\}$ and $\{K_0, K_1, \ldots\}$.

   Subcase 3.1. One of the sets has an infinite bad sequence. Then proceed as in Subcase 2.1.

   Subcase 3.2. All sequences in each set are good. Then proceed as in Subcase 2.2 using Lemma 10 twice.

□

Our proof differs from Nash-Williams' (1963) classical proof in two respects. First, we have avoided what Dershowitz (1987) calls Higman's lemma. This is possible because λ-terms are constructed only from unary and binary operations (abstraction and application). Second, the application of AC can be avoided by explicitly constructing a choice function for $\Lambda^{\mathcal{L}}$. Some results about the mathematical means required to prove the general form of Kruskal's Theorem for finite trees are presented by Simpson (1985) and Gallier (1991).

1.8 Applications

Definition 14 Let **R** be a notion of reduction on $\Lambda^{\mathcal{L}}$. $M \twoheadrightarrow_R^+ N$ is an *R-cycle* if $M \equiv N$, and an *R-self-embedding* if $M \sqsubseteq_{\mathcal{L}} N$.

The following shows that the second observation in the introduction is generally correct.

Corollary 12 *If* $M_0 \rightarrow_R M_1 \rightarrow_R \ldots$ *then*

$$M_0 \twoheadrightarrow_R M_{i_1} \twoheadrightarrow_R^+ M_{j_1} \twoheadrightarrow_R M_{i_2} \twoheadrightarrow_R^+ M_{j_2} \twoheadrightarrow_R \ldots$$

where $M_{i_1} \twoheadrightarrow_R^+ M_{j_1}$, $M_{i_2} \twoheadrightarrow_R^+ M_{j_2}$, $\ldots$ *are self-embeddings.*

Proof. By Lemma 10 and Theorem 11. □

Proposition 13 *It is decidable whether for arbitrary $M \in \Lambda^{\mathcal{L}}$ there exist K, N such that $M \twoheadrightarrow K \twoheadrightarrow^{+} N$ where $K \twoheadrightarrow^{+} N$ is a self-embedding.*

Proof. Given M we want to decide whether there are K, N such that $M \twoheadrightarrow K \twoheadrightarrow^{+} N$ where $K \twoheadrightarrow^{+} N$ is a self-embedding. Let $\underline{\bullet}$ be the strategy from Section 4 (the use of $\underline{\bullet}$ is not essential, it merely simplifies the proof a bit). Compute the sequence

$$M_0 \to M_1 \to M_2 \to \ldots$$

where $M_0 \equiv M$ and $M_{i+1} \equiv \underline{M_i}$, until either $M_i \sqsubseteq_{\mathcal{L}} M_j$ for some $i < j$ or the sequence reaches a normal form. One of the two must eventually happen: if $\infty(M)$ then by perpetuality of $\underline{\bullet}$ and Corollary 12 the first event happens; if $\mathrm{SN}(M)$ then a normal form is eventually encountered unless the sequence is stopped with $M_i \sqsubseteq_{\mathcal{L}} M_j$ for some $i < j$ first.

In any case we know either that $M_i \sqsubseteq_{\mathcal{L}} M_j$ for some $i < j$ in which case we answer *yes*, or that $\mathrm{SN}(M)$ in which case we can systematically search through the finitely many different reduction paths to see whether one contains a self-embedding. $\square$

Plaisted (1985) defines a rewrite system S to be self-embedding if there is a rewrite sequence $t_1 \Rightarrow \ldots \Rightarrow t_n$ using only rules from S such that t_1 is embedded in t_n, and shows that it is undecidable whether a term rewrite system is self-embedding. This does not contradict Propostion 13 since Plaisted's definition ranges over all rewrite systems, whereas proposition 13 concerns the specific system with labeled terms and β-reduction.

It is undecidable whether, for any M, there is an infinite reduction path from M, so the preceding proposition shows that there are terms M, N, K such that $M \twoheadrightarrow K \twoheadrightarrow^{+} N$ where $K \twoheadrightarrow^{+} N$ is a self-embedding, and yet M has no infinite reduction paths. However, if one considers only M with initial labelings, such examples are not straightforward to find. For instance the reduction path from $(\lambda x.y\,x\,x)\,(\lambda x.y\,x\,x)$ with the initial labeling does not have any self-embeddings. See also Example 3(3-4).

Analogs of the preceding corollary and proposition have many applications in term rewrite systems and functional and logic programming, see the papers by Dershowitz (1987) and Bol et al. (1986). We briefly sketch a similar application in λ-calculus; a fuller exposition is not possible in the limited space available.

A λ-calculus *interpreter* is an algorithm that takes a term as input and returns the normal form of the term, if the term has one, and loops infinitely otherwise (ignoring questions concerning encoding). In functional and logic programming there are lines of work aimed at interpreters that detect some forms of non-termination, see the papers by

Holst and Hughes (1992) and Bol et al. (1986). Corollary 12 and Proposition 13 give a criterion that can be implemented in an interpreter for λ-calculus or for a higher-order functional language as follows. Give the input term an initial labeling. Then let the interpreter reduce on the term. As long as the interpreter does not encounter a term in which a previous term is embedded, the interpreter has *definitely not* entered an infinite loop; if the interpreter encounters a term in which a previous one is embedded, the interpreter *may* have entered a loop, and this information can be passed to the user.

1.9 Related work

Dershowitz (1987) and Plaisted (1985) study a notion of self-embedding in term rewrite systems. We do not know of any studies of Kruskal's Theorem in a λ-calculus setting. In the INRIA technical report with abstracts for the 2nd International Workshop on Termination (1995) Benaissa et al. announce an application of Nash-Williams' proof technique to λ-calculus with explicit substitution. Reduction cycles in combinatory logic are studied by Klop (1980) by different techniques.

Böhm et al. (1975) and Böhm and Dezani-Ciancaglini (1975) give, for any normal form M, a constructive definition of a set of normal forms N for which $M\ N$ has a normal form. Since any λ-term can be transformed to an equivalent term which is an applicative combination of normal forms, this can be used to generally approximate whether a term has a normal form or not. On the other hand, Λ_{Ω} directly characterizes a class of terms with arbitrary nesting of λ's and application which are all strongly normalizing.

Reduction strategies in general and perpetual strategies in particular were introduced by Barendregt et al. (1976), and used by these authors as well as by Bergstra and Klop (1982) to characterize redexes whose contraction preserve the possibility of infinite reductions. Some recent contributions to the theory of perpetual strategies are due to Khasidashvili (1994a,1994b), Severi and van Raamsdonk (1995), and Sørensen (1996). The later paper has no overlap with the present paper.

Acknowledgements. I am indebted to Henk Barendregt for posing some of the questions answered in this paper, for comments to an earlier draft, and for his hospitality. Thanks to the λ-group in Nijmegen and TOPPS group at DIKU for creating pleasent working environments, and to Fritz Henglein, Torben Mogensen, Herman Geuvers, and the anonymous referees for feedback. Last but not least thanks to Zurab Khasidashvili for comments and hospitality.

Bibliography

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

Barendregt, H.P., J. Bergstra, J.W. Klop, and H. Volken. 1976. Degrees, Reductions and Representability in the Lambda Calculus. Technical Report Preprint 22. University of Utrecht, Department of Mathematics.

Bergstra, J.A., and J.W. Klop. 1982. Strong Normalization and Perpetual Reductions in the Lambda Calculus. *Journal of Information Processing and Cybernetics* 18:403–417.

Böhm, C., M. Coppo, and M. Dezani-Ciancaglini. 1975. Termination Tests Inside λ-Calculus. Lecture Notes in Computer Science, Vol. 37, 95–110. Springer-Verlag.

Böhm, C., and M. Dezani-Ciancaglini. 1975. λ-Terms as Total or Partial Functions on Normal Forms. In *λ-Calculus and Computer Science Theory.*, ed. C. Böhm, Lecture Notes in Computer Science, Vol. 52, 96–121. Springer-Verlag.

Bol, R.N., K.R. Apt, and J.W. Klop. 1986. An Analysis of Loop Checking Mechanisms for Logic Programs. *Theoretical Computer Science* 86:35–79.

Curry, H.B., and R. Feys. 1958. *Combinatory Logic.* Amsterdam: North-Holland.

Dershowitz, N. 1987. Termination of Rewriting. *Journal of Symbolic Computation* 3.

Gallier, J. 1991. What's so Special about Kruskal's Theorem and the Ordinal Γ_0? *Annals of Pure and Applied Logic* 31:199–260.

Halmos, P. 1960. *Naive Set Theory.* Springer-Verlag. Undergraduate texts in mathematics, 1974 edition.

Higman, G. 1952. Ordering By Divisibility in Abstract Algebras. *Proceedings of the London Mathematical Society* (3) 2:326–336.

Hindley, J.R. 1989. BCK-Combinators and Linear λ-Terms have Types. *Theoretical Computer Science* 64:97–105.

Holst, C.K., and J. Hughes. 1992. A Loop-Detecting Interpreter for Lazy Programs. In *Proc. Glasgow 1991 Workshop on Functional Programming,* 164–176. Springer-Verlag. Workshops in Computing.

INRIA-Lorraine and Centre de Recherche en Informatique de Nancy. 1995. *Abstracts from the 2nd International Workshop on Termination.* May. Distributed electronically.

Jacobs, B. 1993. Semantics of Lambda-I and of other Substructure Calculi. In *Typed Lambda Calculi and Applications*, ed. M. Bezem and J.F. Groote, Lecture Notes in Computer Science, Vol. 664, 195–208. Springer-Verlag.

Khasidashvili, Z. 1994a. The Longest Perpetual Reductions in Orthogonal Expression Reduction Systems. In *Logical Foundations of Computer Science*, ed. A. Nerode and Yu. V. Matiyasevich, Lecture Notes in Computer Science, Vol. 813, 191–203. Springer-Verlag.

Khasidashvili, Z. 1994b. Perpetuality and Strong Normalization in Orthogonal Term Rewriting Systems. In *11th Annual Symposium on Theoretical Aspects of Computer Science*, ed. P. Enjalbert, E.W. Mayr, and K.W. Wagner, Lecture Notes in Computer Science, Vol. 775, 163–174. Springer-Verlag.

Klop, J.W. 1980. *Combinatory Reduction Systems.* Doctoral dissertation, Utrecht University. CWI Tract, Amsterdam.

Komori, Y. 1987. BCK-Algebras and Lambda Calculus. In *Proceedings of the 10th Symposium on Semigroups.* 5–11. Josai University, Sakado.

Kruskal, J.B. 1954. Well-Quasi-Ordering, the Tree Theorem, and Vazsonyi's Conjecture. *Transactions of the American Mathematical Society* 95:210–225.

Lercher, B. 1976. Lambda-Calculus Terms that Reduce to Themselves. *Notre Dame Journal of Formal Logic* XVII(2):291–292.

Nash-Williams, C.St.J.A. 1963. On Well-Quasi-Ordering Finite Trees. *Proceedings of the Cambridge Mathematical Society* 59:833–835.

Plaisted, D.A. 1985. The Undecidability of Self-Embedding for Term-Rewriting Systems. *Information Processing Letters* 20:61–64.

Simpson, S.G. 1985. Nichtbeweisbarkeit von Gewissen Kombinatorishen Eigenschaften Endlicher Bäume. *Archiv für Mathematische Logik und Grundlagenforschung* 25:45–65.

Sørensen, M.H. 1996. Effective Longest and Infinite Reduction Paths in Untyped λ-Calculi. In *Colloqium on Trees in Algebra and Programming.* Lecture Notes in Computer Science. Springer-Verlag. To appear.

van Raamsdonk, F., and P. Severi. 1995. On normalisation. Technical Report CS-R9545. CWI.