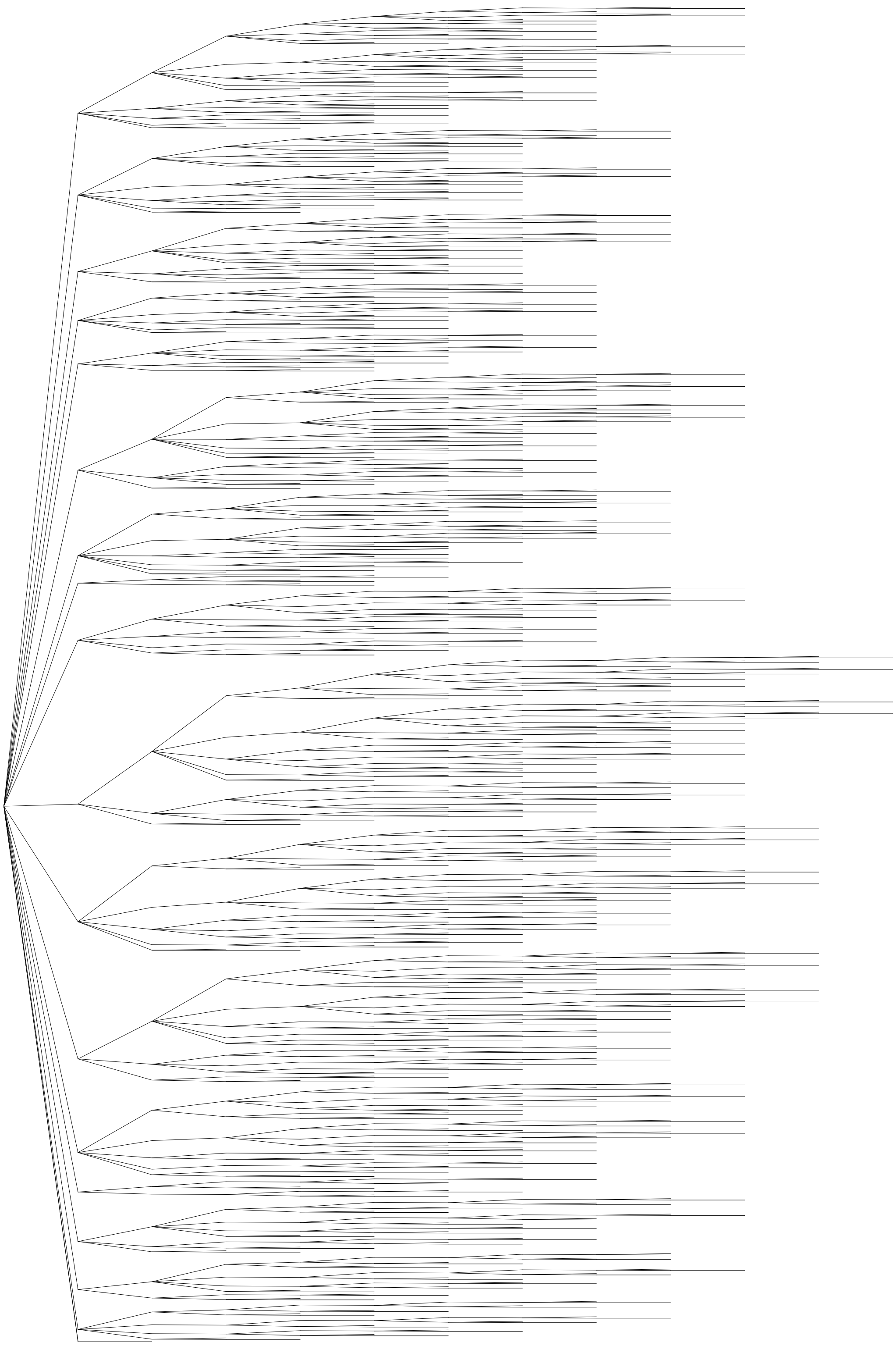
Part II

The Combinator S



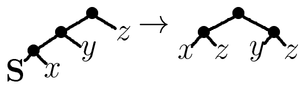
S (**S S S**) **S** (**S** (**S S**) **S**)
after 50 head reductions

Chapter 5

Introduction to CL(S)

5.1 Introduction

CL(**S**) is a combinatory term rewriting system with just one combinator



All combinatory TRS are left linear and non-overlapping. By proposition 2.5.5, they are confluent. Therefore normal forms are unique if they exist.

Moreover, the set $\mathcal{N}$ of normal forms in **CL**(**S**) is rational by proposition 4.6.1.

Furthermore, **CL**(**S**) is non-erasing, so normalization and termination coincide by proposition 2.6.5.

These properties are shared by all non-erasing **CL** systems.

But from now on, we will investigate properties that are special to **CL**(**S**). In the present chapter, we will first fix a convenient notation, and then present basic observations.

They will prepare for the subject treatment in the subsequent chapters.

5.2 Notations

Symbols

Symbols appear in different fonts. We use

- bold face, upper case, for combinators $(\mathbf{S}, \mathbf{K}, \ldots)$,
- italic face, lower case, for variables in the system $(x, y, \ldots)$,
- italic face, upper case, for meta-variables (that denote terms or sets of terms) $(X, Y, \ldots)$,
- calligraphic face, upper case, for fixed sets of terms $(\mathcal{P}, \mathcal{Q}, \ldots)$,
- sans serif face, upper case, for certain fixed terms $(\mathsf{T} = \mathsf{S} \bullet \mathsf{S},\ \mathsf{A} = \mathsf{T} \bullet \mathsf{S} = (\mathsf{S} \bullet \mathsf{S}) \bullet \mathsf{S})$.

We sometimes switch between one- and two-dimensional representation of terms: $\mathbf{S}\ \mathbf{S}\ \mathbf{S} = (\mathsf{S} \bullet \mathsf{S}) \bullet \mathsf{S}$.
Moreover, since the rest of this thesis exclusively deals with the **S** combinator, we often write it as $\circ$ from now on.

Sets of Terms

5.2.1. Definition. $\mathcal{N}$ denotes the set of all terms in *normal* form.

5.2.2. Proposition. $\mathcal{N} = \mathsf{S} \ \cup\ \mathsf{S} \bullet \mathcal{N} \ \cup\ (\mathsf{S} \bullet \mathcal{N}) \bullet \mathcal{N}$

5.2.3. Definition. $\mathcal{M}$ denotes the set of *all* terms.

5.2.4. Proposition. $\mathcal{M} = \mathsf{S} \ \cup\ \mathcal{M} \bullet \mathcal{M}$

Inside an expression, we often just write $*$ for $\mathcal{M}$.
We often write a term when meaning a set containing exactly this term.
Operations are understood to be extended from terms to sets of terms. For instance, $(\mathsf{S} \bullet (\mathsf{S} \cup \mathsf{T})) \bullet \mathsf{T}$ denotes the set $\left\{(\mathsf{S} \bullet \mathsf{S}) \bullet \mathsf{T},\ (\mathsf{S} \bullet \mathsf{T}) \bullet \mathsf{T}\right\}$.
If E and F are sets, $E \to F$ denotes $\forall X \in E : \exists Y \in F : X \to Y$, while $E \not\to F$ stands for $\forall X \in E : \neg\exists Y \in F : X \to Y$.
The set $\mathcal{M}_k$ consists of all terms whose left spine has length $\geq k$.

5.2.5. Definition. $\mathcal{M}_0 = \mathcal{M}; \mathcal{M}_{k+1} = \mathcal{M}_k \bullet \mathcal{M}$.

For example, $\mathcal{M}_0$ are all terms, and $\mathcal{M}_1$ are all terms except **S**.
The term S_k has a right spine of length $k-1$ and all its left children are **S**:

5.2.6. Definition. $\mathsf{S}_1 = \mathbf{S}; \mathsf{S}_{k+1} = \mathsf{S} \bullet \mathsf{S}_k$; $\mathcal{P}_k = \mathsf{S}_1 \ \cup\ \ldots\ \cup\ \mathsf{S}_k$; $\mathcal{Q}_k = \mathcal{M} \setminus \mathcal{P}_k$.

The set $\mathcal{P}_k$ is the collection of all S_i up to k. The complement of $\mathcal{P}_k$ is denoted by $\mathcal{Q}_k$. These are exactly the terms whose right spine has a length $\geq k$ or who have some subterm with left spine longer than 1. For example, $\mathcal{Q}_0$ are all terms, $\mathcal{Q}_2$ are all terms except **S** and T, and $\mathsf{A} = (\mathsf{S} \bullet \mathsf{S}) \bullet \mathsf{S} \in \mathcal{Q}_k$ for any k.
From the above description we immediately derive

5.2.7. Corollary. For all k we have $\mathcal{P}_k \subseteq \mathcal{P}_{k+1}$ and $\mathcal{Q}_k \supseteq \mathcal{Q}_{k+1}$.

5.2.8. Corollary. If $X \in \mathcal{Q}_k$, and $X \lhd Y$, then $Y \in \mathcal{Q}_{k+1}$

We collect all of the $\mathcal{P}_*$ into the set $\mathcal{P}$. These are all terms whose left children of the right spine are leaves. The complement of $\mathcal{P}$ is denoted by $\mathcal{Q}$.

5.2.9. Definition. $\mathcal{P} = \bigcup_{k\geq 1} \mathsf{S}_k; \quad \mathcal{Q} = \mathcal{M} \setminus \mathcal{P}$

5.2.10. Corollary. For all k we have $\mathcal{P}_k \subseteq \mathcal{P}$ and $\mathcal{Q} \subseteq \mathcal{Q}_k$.

5.2.11. Corollary. $\mathcal{P} = \mathbf{S} \cup \mathbf{S}\mathcal{P}; \mathcal{Q} = \mathcal{M}_2 \cup \mathbf{S}\mathcal{Q}$.

5.2.12. Corollary. If $X \in \mathcal{Q}$ and $X \lhd Y$, then $Y \in \mathcal{Q}$.

Patterns

Certain subterm positions are related to certain reduction properties. If Y is some expression, we'll define the set $\langle Y\rangle$ of *triads* over Y. It has the property $\langle Y\rangle Z \lhook\joinrel\twoheadrightarrow YZ$.

5.2.13. Definition. $\langle Y\rangle = Y \cup \mathbf{S}\langle Y\rangle\mathcal{M} \cup \mathbf{S}\mathcal{M}\langle Y\rangle$

5.2.14. Proposition. $\langle Y\rangle Z \lhook\joinrel\twoheadrightarrow YZ$

Proof. The claim is proved by structural induction. Let $Y' \in \langle Y\rangle$.

1. $Y' = Y$. Then the claim is vacuously true.
2. $Y' \in \mathbf{S}\langle Y\rangle *$. Then $Y'Z = \mathbf{S}\langle Y\rangle * Z \to \boxed{\langle Y\rangle Z}(*Z)$.
3. $Y' \in \mathbf{S} * \langle Y\rangle$. Then $Y'Z = \mathbf{S} * \langle Y\rangle Z \to (*Z)\boxed{\langle Y\rangle Z}$.

□

Another construction is the stairway pattern $\langle {}^{X}/_{Y}\rangle$.

5.2.15. Definition. $\langle {}^{X}/_{Y}\rangle = Y \cup X\langle {}^{X}/_{Y}\rangle$

In the literature this is also denoted by X^*Y but we chose this notation because of the similarity to $\langle Y\rangle$, of which the stairways are special case in the following situation:

5.2.16. Corollary. If $X \in \mathbf{S}*$, then $\langle {}^{X}/_{Y}\rangle \subseteq \langle Y\rangle$.

5.3 Predecessors and Successors

5.3.1. Proposition. For any **S**-term Y, the set of predecessors $\Delta^{\twoheadleftarrow}(Y)$ is finite and computable.

Proof. If $X \to Y$, then $|X| \leq |Y|$. There are only finitely many terms X that fulfill this condition. They can all be enumerated, and $X \to Y$ can be checked. This generates all *immediate* predecessors of Y.

Their immediate predecessors can be computed as well, and this process can be continued, but it must eventually stop (when all terms no larger than $|Y|$ have been checked). □

5.3.2. Remark. So far, this says nothing about the presence or absence of *cycles* $Y \to^+ Y$. This will be dealt with in proposition 6.2.2.

5.3.3. Proposition. Given **S**-terms X and Y, the question "does $X \twoheadrightarrow Y$ hold?" is decidable.

Proof. Just enumerate all predecessors of Y and check whether X occurs. □

5.3.4. Remark. The question "does $X \leftrightarrow^* Y$ hold?" is the *word problem.* We conjecture that it is decidable as well. However, with current knowledge, a proof seems out of reach.

5.3.5. Remark. What happens with $\Delta^{\twoheadleftarrow}(L)$ for rational L? We do not know. It it were rational (and effectively computable from an automaton for L), then $\Delta^{\twoheadleftarrow}(\mathcal{N})$ would immediately solve the question of normalization.

Because **CL**(**S**) is duplicating, $\Delta^{\twoheadrightarrow}(L)$ and even $\Delta^{\rightarrow}(L)$ need not be rational for rational L.

5.3.6. Example. $\mathsf{A}\mathcal{P}$ is rational, but $\Delta^{\rightarrow}\left(\mathsf{A}\mathcal{P}\right) = \Delta^{\twoheadrightarrow}\left(\mathsf{A}\mathcal{P}\right) = \left\{\mathsf{S}_k\mathsf{S}_k \mid k > 1\right\}$ is not.

While for singleton L, the set $\Delta^{\twoheadleftarrow}(L)$ is finite (therefore rational), even in this case $\Delta^{\twoheadrightarrow}(L)$ need not be rational, as shown by terms with infinite reductions.

However there *are* rational sets of **S**-terms that behave well under reduction.

5.3.7. Definition. $\mathcal{N}_{00} = \mathbf{S} \cup \mathsf{T}\mathcal{N}_{00} \cup \mathcal{N}_{00}\mathsf{S}$

5.3.8. Proposition. $\mathcal{N}_{00}$ is forward and backward closed under reduction: $\mathcal{N}_{00} = \Delta^{=}(\mathcal{N}_{00})$.

Proof. As set forth in section 4.6, saturation of $\mathcal{N}_{00}$ w.r.t. **CL**(**S**) (with the arrow pointing forwards and backwards) can be checked. We indicate the check for the forward closure: From the grammar for $\mathcal{N}_{00}$ we derive that a redex in $\mathcal{N}_{00}$ must have the form $((\mathsf{S}\mathsf{S})\mathcal{N}_{00})\mathsf{S}$. Its contractum is $(\mathsf{S}\mathsf{S})(\mathcal{N}_{00}\mathsf{S}) = \mathsf{T}(\mathcal{N}_{00}\mathsf{S}) \subseteq \mathsf{T}\mathcal{N}_{00} \subseteq \mathcal{N}_{00}$. □

5.4 Monsters

5.4.1. Definition. A *monster* is a term with normal form that is very large and many reduction steps away.

To get an impression of what *very* and *many* could mean, we look at an example.

5.4.2. Definition. $\mathsf{A}_0 = \mathsf{S}; \mathsf{A}_{n+1} = (\mathsf{T}\ \mathsf{A}_n) \quad \mathsf{T}_1 = \mathsf{T}; \mathsf{T}_{m+1} = (\mathsf{T}\ \mathsf{T}_m)$

5.4.3. Example. Names are chosen such that $\mathsf{A}_1 = \mathsf{A}; \mathsf{T}_1 = \mathsf{T}$

5.4.4. Proposition. $((\mathsf{A}_n\ \mathsf{A}_m)\ \mathsf{S})$ are monsters. Their normal form has size $n^{\Theta(m)}$ and is reachable in $\Theta(n \cdot m)$ steps.

For the proof, we additionally define

$$\mathsf{B}_0(X) = (\mathsf{S}\ X); \mathsf{B}_{m+1}(X) = ((\mathsf{S}\ X)\ \mathsf{B}_m(X)); \mathsf{C}_0(X) = (\mathsf{T}\ X); \mathsf{C}_{m+1}(X) = ((\mathsf{S}\ X)\ \mathsf{C}_m(X))$$
$$\mathsf{D}_{0,n} = ((\mathsf{S}\ \mathsf{A}_n)\ \mathsf{S}); \mathsf{D}_{m+1,n} = \mathsf{C}_{n+1}(\mathsf{D}_{m,n})$$

Note that all these are normal forms (if X is a normal form).

5.4.5. Lemma. $\mathrm{nf}\left((\mathsf{A}_m\ X)\right) = \mathsf{B}_m(X)$ as well as $\mathrm{nf}\left((\mathsf{T}_{m+1}\ X)\right) = \mathsf{C}_m(X)$, and both reductions take m steps.

Proof. $(\mathsf{A}_0\ X) = (\mathsf{S}\ X) = \mathsf{B}_0(X)$.

$(\mathsf{A}_{m+1}\ X) = ((\mathsf{T}\ \mathsf{A}_m)\ X) \to ((\mathsf{S}\ X)\ \boxed{(\mathsf{A}_m\ X)}) \twoheadrightarrow ((\mathsf{S}\ X)\ \mathsf{B}_m(X)) = \mathsf{B}_{m+1}(X)$.

$(\mathsf{T}_1\ X) = (\mathsf{T}\ X) = \mathsf{C}_0(X)$.

$(\mathsf{T}_{m+1}\ X) = ((\mathsf{T}\ \mathsf{T}_m)\ X) \to ((\mathsf{S}\ X)\ \boxed{(\mathsf{T}_m\ X)}) \twoheadrightarrow ((\mathsf{S}\ X)\ \mathsf{C}_m(X)) = \mathsf{C}_{m+1}(X)$. □

By replacing X with S in this proof, and using the identities $(\mathsf{A}_0\ \mathsf{S}) = (\mathsf{S}\ \mathsf{S}) = \mathsf{T}_1$ and $(\mathsf{T}_1\ \mathsf{S}) = ((\mathsf{S}\ \mathsf{S})\ \mathsf{S}) = \mathsf{A}_1$, we obtain

5.4.6. Lemma. $\mathrm{nf}\left((\mathsf{A}_m\ \mathsf{S})\right) = \mathsf{T}_{m+1}$ as well as $\mathrm{nf}\left((\mathsf{T}_m\ \mathsf{S})\right) = \mathsf{A}_m$, and reductions take m and $m-1$ steps, respectively.

5.4.7. Lemma. $\mathrm{nf}\left((\mathsf{B}_m(\mathsf{A}_n)\ \mathsf{S})\right) = \mathsf{D}_{m,n}$. It is reached in $2 \cdot m \cdot n$ reduction steps.

Proof. $(\mathsf{B}_0(\mathsf{A}_n)\ \mathsf{S}) = ((\mathsf{S}\ \mathsf{A}_n)\ \mathsf{S})$.

$(\mathsf{B}_{m+1}(\mathsf{A}_n)\ \mathsf{S}) = (((\mathsf{S}\ \mathsf{A}_n)\ \mathsf{B}_m(\mathsf{A}_n))\ \mathsf{S}) \to (\boxed{(\mathsf{A}_n\ \mathsf{S})}\ \boxed{(\mathsf{B}_m(\mathsf{A}_n)\ \mathsf{S})})$, using one reduction step.

By lemma 5.4.6, $(\mathsf{A}_n\ \mathsf{S}) \twoheadrightarrow \mathsf{T}_{n+1}$ in n steps.

By induction, $(\mathsf{B}_m(\mathsf{A}_n)\ \mathsf{S}) \twoheadrightarrow \mathsf{D}_{m,n}$ in $2 \cdot m \cdot n$ steps.

Therefore the whole expression reduces to $(\mathsf{T}_{n+1}\ \mathsf{D}_{m,n})$ in $(2 \cdot m + 1) \cdot n$ steps.

This again has normal form $\mathsf{C}_{n+1}(\mathsf{D}_{m,n}) = \mathsf{D}_{m+1,n}$ which is reached in n more steps by lemma 5.4.5.

In all, this were $(2 \cdot m + 2) \cdot n$ steps. □

So we proved that the normal forms are $\mathsf{D}_{m,n}$ indeed. Now we compute their sizes.

5.4.8. Proposition. $|\mathsf{C}_m\,(X)| = \Theta\,(m \cdot |X|)$

Proof. $|\mathsf{C}_0\,(X)| = \left| \mathsf{S}\ \mathsf{S}\ X \right| = 4 + |X|$.

$|\mathsf{C}_{m+1}\,(X)| = \left| \mathsf{S}\ X\ \mathsf{C}_m\,(X) \right| = 3 + |X| + |\mathsf{C}_m\,(X)|$. □

5.4.9. Proposition. $|\mathsf{D}_{m,n}| = n^{\Theta(m)}$

Proof. $|\mathsf{D}_{0,n}| = \left| \mathsf{S}\ \mathsf{A}_n\ \mathsf{S} \right| = \Theta\,(n)$.

$|\mathsf{D}_{m+1,n}| = |\mathsf{C}_{n+1}\,(\mathsf{D}_{m,n})| = \Theta\,((n+1) \cdot |\mathsf{D}_{m,n}|)$. □

Proof of proposition 5.4.4. $\mathsf{A}_m\ \mathsf{A}_n\ \mathsf{S} \twoheadrightarrow \mathsf{B}_m\,(\mathsf{A}_n)\ \mathsf{S}$, using m steps, by lemma 5.4.5. $\mathsf{B}_m\,(\mathsf{A}_n)\ \mathsf{S} \twoheadrightarrow \mathsf{D}_{m,\mathsf{A}_n}$, using $2 \cdot m \cdot n$ steps, by lemma 5.4.7. □

5.4.10. Proposition. $\mathbf{S}\ \mathsf{A}\ \mathsf{T}_n\ \mathbf{S}\ \mathbf{S}\ \mathbf{S}$ are monsters. Their normal form has size $n^{\Theta(n)}$, and it is reached in at least $\Theta\,(n^2)$ steps.

Proof. $\mathsf{S}\ \mathsf{A}\ \mathsf{T}_n\ \mathsf{S}\ \mathsf{S}\ \mathsf{S} \rightarrow \mathsf{A}\ \mathsf{S}\ \mathsf{T}_n\ \mathsf{S}\ \mathsf{S} \rightarrow \mathsf{T}\ \mathsf{T}\ \mathsf{T}_n\ \mathsf{S}\ \mathsf{S} \rightarrow \mathsf{S}\ \mathsf{T}_n\ (\mathsf{T}\ \mathsf{T}_n)\ \mathsf{S}\ \mathsf{S} = \mathsf{S}\ \mathsf{T}_n\ \mathsf{T}_{n+1}\ \mathsf{S}\ \mathsf{S} \rightarrow \underline{\mathsf{T}_n\ \mathsf{S}}\ \underline{(\mathsf{T}_{n+1}\ \mathsf{S})}\ \mathsf{S} \twoheadrightarrow \mathsf{A}_n\ \mathsf{A}_{n+1}\ \mathsf{S}$. □

We note that T_n and A_n are normal forms of *flat* **S**-terms, for instance $\mathbf{S}\,\mathbf{S}\,\mathbf{S}\,\mathbf{S}\,\mathbf{S}\,\mathbf{S} \twoheadrightarrow \mathsf{T}\ (\mathsf{T}\ \mathsf{T}) = \mathsf{T}_3$.

5.4.11. Example. $\mathsf{S}\ \mathsf{A}\ (\mathbf{S}\,\mathbf{S}\,\mathbf{S}\,\mathbf{S}\,\mathbf{S}\,\mathbf{S})\ \mathbf{S}\ \mathbf{S}\ \mathbf{S}$ has a normal form of size 6131, which is reached after 55 steps.

We conjecture that there are no monsters that grow substantially larger. However, a proof for a statement like "all terms X reach their normal form in $O\left(|X|^2\right)$ innermost reduction steps" doesn't seem immediately at hand. It would entail an algorithm that decides termination.

A History lesson on monsters

The following information was kindly provided by Henk Barendregt.

van der Poel and him were looking for monsters in **CL**(**S**) w.r.t. *strong* reduction (see section 5.6) and found $\mathbf{S}\ \mathsf{T}\ \mathbf{S}\ \mathsf{T}\ \mathbf{S}$ and $\mathbf{S}\ \mathsf{T}\ \mathsf{T}\ \mathsf{T}\ \mathbf{S}\ \mathbf{S}$. See also [vdPSvdM80], Appendix E.

The first expression has a strong normal form roughly of size 10^6. The second one reduces in one step to $\mathsf{T}\ \mathsf{T}\ (\mathsf{T}\ \mathsf{T})\ \mathbf{S} = \mathsf{T}_2\ \mathsf{T}_2\ \mathbf{S}$, where $\mathsf{T}_m\ \mathsf{T}_n\ \mathbf{S}$ can be shown to be monsters by similar arguments as presented above.

5.5 Terms without normal form

We give two examples of infinite reductions in **CL**(**S**).

An infinite, unique reduction

5.5.1. Example. The expression $X_0 X_0$ where $\mathsf{T} = \mathsf{S}\mathsf{S}$; $X_0 = \mathsf{S}\mathsf{T}\mathsf{T}$; $X_{n+1} = \mathsf{T}X_n$ has an infinite reduction. It has already been described by E. Zachos in [Zac78]. All redexes that occur during the reduction have the form $X_n X_m$ for some n and m. There are two types of reduction steps: one goes *downward*: $X_{n+1} X_m = \mathsf{S}\mathsf{S}X_n X_m \to \mathsf{S}X_m \boxed{X_n X_m}$

and the other goes *upward*: $X_0 X_m = \mathsf{S}\mathsf{T}\mathsf{T}X_m \to \mathsf{T}X_m(\mathsf{T}X_m) = X_{m+1} X_{m+1}$

It is interesting to note that the reduction graph of $X_0 X_0$ is a *pure line* (each expression contains exactly one redex).

An infinite head reduction

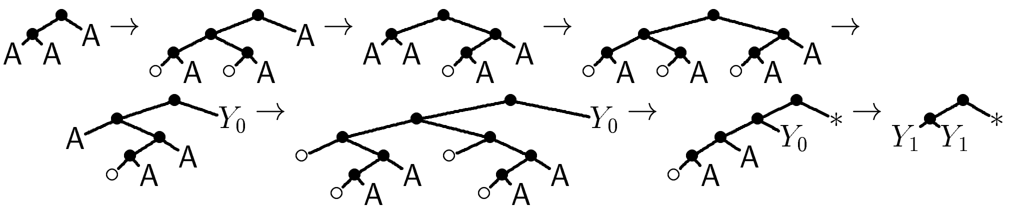
5.5.2. Example. The expression

$$\mathsf{A}\mathsf{A}\mathsf{A} \quad \text{where} \quad \mathsf{A} = \mathsf{S}\mathsf{S}\mathsf{S}$$

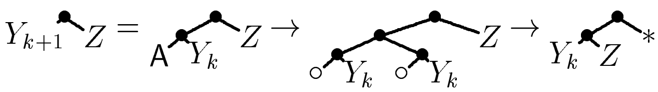
has an infinite reduction. It had been found by H. Barendregt, see [Bar84]. Again the reduction can be described by a pattern, but it is not so obvious:

$$Y_0 = \circ\mathsf{A}(\circ\mathsf{A}\mathsf{A});\ Y_{n+1} = \mathsf{A}Y_n$$

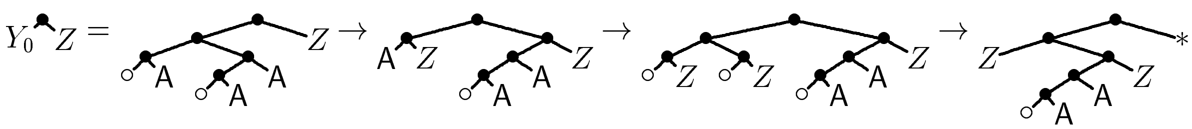
First we show that the pattern is reached.

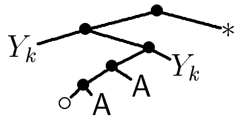


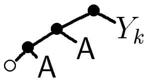
(Here, $*$ denotes a subexpression whose value is irrelevant because no later reduction touches it.) Then we show how the pattern reduces. For any Z, there is a *downwards* move that in general goes like



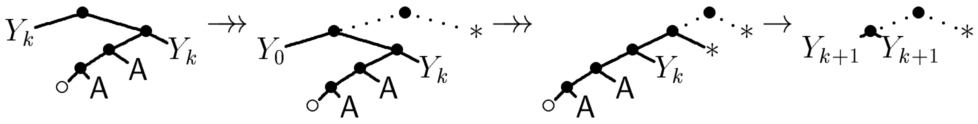
and an *upwards* move



The placeholder Z is instantiated with Y_k to give $Y_0 Y_k \twoheadrightarrow$ 

and now $Z =$ 

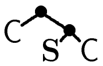
gives

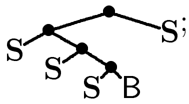


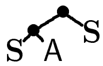
This is an infinite head reduction.

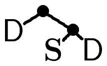
More infinite unique reductions

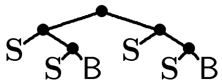
J. W. Klop raised the question whether there are any pure lines in **CL**(**S**) (question 3.3 in [Klo80]). We just saw one example $X_0 X_0$. Indeed there are more, and their typical start terms are:

1. 

where C = 

; B = 

2. 

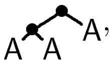
where D = 

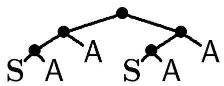
They have infinite, unique head reductions.

For a complete formulation of this statement, and a proof, see the *very* technical report [Wal96].

A history lesson on terms without normal form

(The following information was kindly provided by Henk Barendregt and Jan Willem Klop)

In 1975, Klop, Bergstra, Volken and Barendregt placed a 'bet' on finding an expression in **CL**(**S**) that has no normal form. Barendregt won with 

, and a similar term 

with a simpler proof.

Later Dobouè found the shortest term without normal form: **S T S S S S**, and its reduct **A T S S S**, but without proof, which was later supplied by Monique Baron, and simplified by Klop, Bergstra, Volken, and Barendregt.

The paper [vdPSvdM80], Appendix E, mentions the expression **S T A T S** of which no normal form could be found.

5.6 Weak and Strong reduction

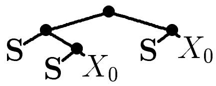
This thesis considers reductions in the $\mathbf{CL}(\mathbf{S})$ term model only. One could also consider the translation of $\mathbf{CL}(\mathbf{S})$ into Λ^0 via the function $\langle\cdot\rangle_\Lambda$, and then do rewriting there. In this respect, $\rightarrow_{\mathbf{CL}}$ is called *weak* reduction and $\rightarrow_\beta$ is called *strong* reduction. These two concepts differ substantially.

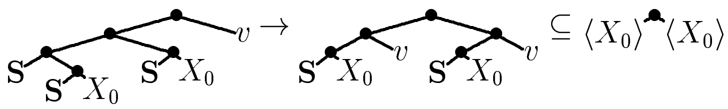
5.6.1. Proposition. A weak normal form need not be in strong normal form.

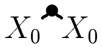
5.6.2. Example.  is in weak normal form. Its strong normal form is

$$\lambda\, x\, y\, z\, .\ (x\ z\ (y\ z))\ \mathbf{S} \rightarrow \lambda\, y\, z\, .\ (\mathbf{S}\ z\ (y\ z)) = \lambda\, y\, z\, .\ (\lambda\, x'\, y'\, z'\, .\ (x'\ z'\ (y'\ z'))\ z\ (y\ z)) \twoheadrightarrow \lambda\, y\, z\, .\ (\lambda\, z'\ (z\ z'\ (y\ z\ z')))$$

5.6.3. Proposition. There is a term that is in weak normal form, but has no strong normal form at all.

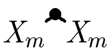
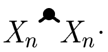
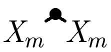
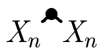
5.6.4. Example. We use notation of example 5.5.1. The term is $Y =$ . It is enough to show that  for a variable v, has no weak normal form.



by slightly abusing the definition of $\langle\cdot\rangle$. Roughly speaking, this behaves like  which has no normal form. (A full proof would apply proposition 8.2.3.)

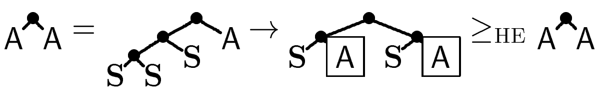
5.7 Homeomorphic Embeddings

By Kruskal's theorem 2.8.3, each infinite reduction $X_1 \rightarrow X_2 \rightarrow \ldots$ contains X_i and X_j with $i < j$ and $X_i \leq_{\mathrm{HE}} X_j$.

5.7.1. Example. See example 5.5.1. For all $m < n$, we have  $\hookrightarrow\!\!\!\!\twoheadrightarrow$ . From $X_m \lhd X_n$ we derive  $\leq_{\mathrm{HE}}$  which uniformly gives the desired embeddings.

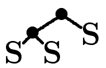
However the terminating reductions don't necessarily *omit* embeddings:

5.7.2. Proposition. There are $X, Y \in \mathbf{CL}(\mathbf{S})$ such that $X \rightarrow^+ Y$ and $X \leq_{\mathrm{HE}} Y$, but $\downarrow(X)$.

5.7.3. Example. 

So the homeomorphic embedding relation isn't really informative in $\mathbf{CL}(\mathbf{S})$.
This is also shown by

5.7.4. Proposition. There are $X, Y \in \mathbf{CL}(\mathbf{S})$ where $X \leq_{\mathrm{HE}} Y$, and $X \twoheadrightarrow \infty$, but $\downarrow(Y)$.

5.7.5. Example. $X = \mathbf{S}\ \mathsf{T}\ \mathbf{S}\ \mathbf{S}\ \mathbf{S}\ \mathbf{S}$ has no normal form, but $Y = \mathbf{S}\ \mathsf{A}\ \mathbf{S}\ \mathbf{S}\ \mathbf{S}\ \mathbf{S}$ has one. Still $X \leq_{\mathrm{HE}} Y$ because of $\mathsf{T} =$  $\leq_{\mathrm{HE}}$  $= \mathsf{A}$

Chapter 6

Loops

6.1 Introduction

Often infinite reductions in term rewriting systems can be described by patterns that re-appear, possibly in varied instantiations.

6.1.1. Definition. A *loop* is a reduction $X \to^+ C[X\sigma]$, where X is a term possibly containing variables, $C[\cdot]$ is a context, and σ is a substitution.

A loop can be composed with itself, giving an infinite reduction. For ground TRS, infinite reductions necessarily contain ground loops, and it is decidable whether such a TRS admits them.

In general, however, presence of loops is not decidable, and non-termination not necessarily means that loops exist.
For instance,

6.1.2. Conjecture. $\mathbf{CL}(\mathbf{S})$ admits no loops.

Empirical observations support that. We apply two different methods that prove weakened versions of that claim.

For completeness, we start by listing yet another method. It is already known in the literature. This is a potential method due to J. Bergstra that shows the absence of *cycles* from $\mathbf{CL}(\mathbf{S})$.
Then we define a notion of *monotone* term which helps to show the absence of certain (not all) *ground* loops.
This method can easily be applied to other rewrite systems. We do so for head reductions in $\mathbf{CL}(\mathbf{S}, \mathbf{K})$.
After that we'll turn to excluding *all* ground loops from $\mathbf{CL}(\mathbf{S})$, using quite a different approach.
This proof uses recursive path orderings. It should be applicable to other rewrite systems as well.
Yet all attempts fail to show the absence of *arbitrary* loops from $\mathbf{CL}(\mathbf{S})$.

Loops have been investigated by Zantema and Geser [ZG96]. They present a one-rule TRS that is non-terminating, but non-looping; and also a two-rule string rewriting system with that property.

6.2 Cycles

6.2.1. Definition. A *cycle* is a reduction $X \to^+ X$.

A cycle can be composed with itself infinitely often, so the presence of cycles shows non-termination of the rewriting system.

6.2.2. Proposition. $\mathbf{CL}(\mathbf{S})$ admits no cycles.

(The following proof is due to J. Bergstra and J. W. Klop, [BK79].)
In $L = \mathbf{S}\ X\ Y\ Z \to X\ Z\ (Y\ Z) = R$, we have $|L|$ with $|R|$ iff $Z = \mathbf{S}$, else $|L| < |R|$.
Now define a potential function $\|\mathbf{S}\| = 1; \|X\ Y\| = 2 \cdot \|X\| + \|Y\|$. Then

$$\begin{aligned}
\|L\| &= \|\mathbf{S}\ X\ Y\ Z\| = 8 + 4 \cdot \|X\| + 2 \cdot \|Y\| + \|Z\| \\
\|R\| &= \|X\ Z\ (Y\ Z)\| = 4 \cdot \|X\| + 2 \cdot \|Y\| + 3 \cdot \|Z\| \\
\|R\| &- \|L\| = 2 \cdot \|Z\| - 8
\end{aligned}$$

Observe that if $Z = \mathbf{S}$, then $\|L\| > \|R\|$.
Both $|\cdot|$ and $\|\cdot\|$ are strictly monotone w. r. t. term construction:

6.2.3. Lemma. If $\|Y\| < \|Z\|$, then for all X and for all $p \in \mathrm{Pos}(X)$: $\|X[p \leftarrow Y]\| < \|X[p \leftarrow Z]\|$.

Then we conclude:

6.2.4. Lemma. If $Y \to Z$, then

1. $|Y| < |Z|$
2. or $|Y| = |Z|$ and $\|Y\| > \|Z\|$.

This leads to

6.2.5. Lemma. If $Y \to^+ Z$, then $(|Y|, \|Y\|) \neq (|Z|, \|Z\|)$

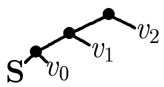
Therefore $X \neq Y$, and there is no $X \to^+ X$.
It would be nice to have more of such potential functions to solve other problems in $\mathbf{CL}(\mathbf{S})$.

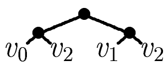
6.3 Monotone Terms

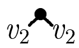
We consider $\mathbf{CL}(\mathbf{S}, V)$, terms with variables from a set V that is totally ordered: $v_0 < v_1 < v_2 < \ldots$.

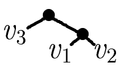
6.3.1. Definition. A term $X \in \mathbf{CL}(\mathbf{S}, V)$ is called *monotone* iff for all 

$\trianglelefteq X$ with $w \in V$, and for all $v \trianglelefteq Y$ with $v \in V$, we have $v < w$.

6.3.2. Example. 

is monotone. 

is monotone. 

is *not* monotone. 

is monotone.

6.3.3. Proposition. If $X \in \mathbf{CL}(\mathbf{S}, V)$ where X is monotone, and $X \to Y$, then Y is monotone.

Proof. Let $p +\!\!\!+ (1)$ be the position of a variable $w \trianglelefteq Y$, and $p +\!\!\!+ (0) +\!\!\!+ q$ be the position of a variable $v \trianglelefteq Y$.

Now let r be the position of the redex $\mathbf{S}\,A\,B\,C \trianglelefteq X$ (and the contractum $A\,C\,(B\,C)$ in Y).

If $r \| p$, then reduction did not change anything in $X\,[p]$, so $v < w$ remained intact because X is monotone.

If $p +\!\!\!+ (1)$ denotes a strict subterm of A, B, or one of the C in the contractum, then the subterm at position p was already present in X, so monotonicity holds.

The position $p +\!\!\!+ (1)$ cannot denote A or B in the contractum, because these are left subterms of their parent node.

If $p +\!\!\!+ (1)$ denotes one of the C in the contractum, then we have to verify that $v < w = C$. In X, the redex is $\mathbf{S}\,A\,B\,w$. By monotonicity, for all $v \trianglelefteq A \trianglelefteq \mathbf{S}\,A\,B$ and all $v \trianglelefteq B \trianglelefteq \mathbf{S}\,A\,B$ we know $v < w$. Therefore $A\,w$ and $B\,w$ in the contractum are monotone as well.

Finally, $p \sqsubseteq r$ is impossible because $Y\,[r]$ is a contractum, but $Y\,[p]$ is a variable. $\square$

6.3.4. Proposition. There is no term X in $\mathbf{CL}(\mathbf{S})$ such that $X\,v \lhook\joinrel\twoheadrightarrow v\,v$.

Proof. The term $X\,v$ is monotone, because X does not contain variables at all. So all its successors must be monotone as well, by proposition 6.3.3. None of them can contain $v\,v$, which is not monotone. $\square$

6.3.5. Remark. If there were such an X, then $X\,X \lhook\joinrel\twoheadrightarrow X\,X \lhook\joinrel\twoheadrightarrow \ldots$ would be a ground loop. Now we know that ground loops don't have this simple form. We will soon prove that $\mathbf{CL}(\mathbf{S})$ admits no ground loops at all, using quite a different technique however, because it is not clear how monotonicity could be defined for *ground* terms.

The notion of monotonicity can be applied to $\mathbf{CL}\,(\mathbf{S}, \mathbf{K})$, because it is respected at least by head reductions.

6.3.6. Proposition. If $X \in \mathbf{CL}\,(\mathbf{S}, \mathbf{K}, V)$ where X is monotone, and $X \rightarrow_h Y$, then Y is monotone.

Proof. If the redex is $\mathbf{S}\ A\ B\ C$, or the redex position is independent from the variable position, then the reasoning is exactly as in proposition 6.3.3.

If the redex is $\mathbf{K}\ A\ B$, then the contractum is A, which is harmless because it is at head position, and therefore cannot be a variable whose left neighbours had to be checked.

Of course this would be violated for non-head redexes. $\square$

This also proves a conjecture stated by Jan Willem Klop in [Klo80]:

6.3.7. Proposition. There is no $X \in \mathbf{CL}\,(\mathbf{S}, \mathbf{K})$ with $X\ a \lhook\joinrel\twoheadrightarrow_h a\ a$.

6.4 Ground Loops

6.4.1. Definition. A *ground loop* is a reduction $X \to^+ C[X]$, where $C[]$ is a context and X is a ground term.

A ground loop, composed with itself, would lead to an infinite reduction containing only finitely many different redexes, so the redex size would be bounded by some constant.

Yet we are going to show that for any fixed n, the restriction of $\mathbf{CL}(\mathbf{S})$ that only allows rewrites where the redex size does not exceed n, is terminating.

Restriction of CL(S) by labeling

In fact we are not bounding the *size* of the redexes but rather their *right depth* (length of their right spine). Moreover, we don't measure this depth exactly but use a safe approximation instead. This approximation is attached as a label to the application nodes and conservatively updated where reductions take place.

We fix some notation for this.

The labeled system $\mathbf{CL}_n(\mathbf{S})$ has the signature $\{\mathbf{S}, \circ_1, \circ_2, \ldots, \circ_n, \circ_\infty\}$ where $\mathbf{S}$ is as before, while the $\circ_i$ are labeled application nodes. There are $n+1$ of them. The rewrite rules of $\mathbf{CL}_n(\mathbf{S})$ are

$$\left\{ ((\mathbf{S} \circ_* x) \circ_* y) \circ_k z \to (x \circ_k z) \circ_{k+1} (y \circ_k z) \mid k \leq n \right\}$$

where $n+1 = \infty$ and the $*$ match arbitrary numbers.

Next, we prove that $\mathbf{CL}_n(\mathbf{S})$ is terminating.

After that we derive how this carries over to the original $\mathbf{CL}(\mathbf{S})$.

$\mathbf{CL}_n(\mathbf{S})$ is terminating

6.4.2. Theorem. For all n, $\mathbf{CL}_n(\mathbf{S})$ is terminating.

Proof. Order the signature by $\mathbf{S} > \circ_1 > \circ_2 > \ldots > \circ_n > \circ_\infty$.

This gives a recursive path order $>_{\text{rpo}}$ which is a reduction order for $\mathbf{CL}_n(\mathbf{S})$, as can be easily verified:

We need to show that

$$L = ((\mathbf{S} \circ_* x) \circ_* y) \circ_k z >_{\text{rpo}} (x \circ_k z) \circ_{k+1} (y \circ_k z) = R$$

The top symbols of L and R are $\circ_k > \circ_{k+1}$.

We verify $L >_{\text{rpo}} x \circ_k z$ and $L >_{\text{rpo}} y \circ_k z$. Here the top symbols coincide. We have to prove the multiset relation $\left\{ (\mathbf{S} \circ_* x) \circ_* y, z \right\} \gg \{x, z\}$ (and similar for $\{y, z\}$). This is true because $\mathbf{S} \circ_* x \rhd x$. $\square$

Consistency of labeling in $\mathbf{CL}_n(\mathbf{S})$

6.4.3. Definition. The *right depth* of a terms in $\mathbf{CL}(\mathbf{S})$ or $\mathbf{CL}_n(\mathbf{S})$ is

$$d_{\mathrm{r}}(\mathbf{S}) = 0;\, d_{\mathrm{r}}\left(x \overset{*}{\wedge} y\right) = 1 + d_{\mathrm{r}}(y)$$

We are going to show that the labeling of the application nodes is preserved as a correct approximation (from below) of the right depth of the terms.

6.4.4. Definition. A term X is called *consistent* iff for each subterm $X' \trianglelefteq X$ that has k as its top label, $d_{\mathrm{r}}(X') \geq k$.

Obviously, subterms of consistent terms are consistent.
Consistency is maintained by the reduction rules in $\mathbf{CL}_n(\mathbf{S})$.

6.4.5. Proposition. If X is consistent and $X \to_{\mathbf{CL}_n(\mathbf{S})} Y$, then Y is consistent.

Proof. Consider on reduction step $X \to_{\mathbf{CL}_n(\mathbf{S})} Y$. Assume the redex $((\mathbf{S} \overset{*}{\wedge} B) \overset{*}{\wedge} C) \overset{k}{\wedge} D \trianglelefteq X$ had been contracted.
Subterms of B, C, D just get copied and remain consistent.
Subterms whose position is independent from the redex position are not affected at all.
Subterms lying on the path from the redex position to the root might increase their right depth due to the reduction, but they don't change their top label. So the approximation remains correct.
We consider terms $B \overset{k}{\wedge} D$ and $C \overset{k}{\wedge} D$ in the contractum. We know that the redex $R = ((\mathbf{S} \overset{*}{\wedge} B) \overset{*}{\wedge} C) \overset{k}{\wedge} D$ was consistent, so $k \leq d_{\mathrm{r}}(R) = 1 + d_{\mathrm{r}}(D)$. It is allowed to attach the label k in $B \overset{k}{\wedge} D$ because $d_{\mathrm{r}}\left(B \overset{k}{\wedge} D\right) = 1 + d_{\mathrm{r}}(D) \geq k$. (The same reasoning is valid for $C \overset{k}{\wedge} D$.) So $B \overset{k}{\wedge} D$ and $C \overset{k}{\wedge} D$ are consistent.
Finally $Z = (B \overset{k}{\wedge} D) \overset{k+1}{\wedge} (C \overset{k}{\wedge} D)$ is consistent because $d_{\mathrm{r}}(Z) = 1 + d_{\mathrm{r}}\left(C \overset{k}{\wedge} D\right) \geq 1 + k$ and $C \overset{k}{\wedge} D$ has just been proven consistent. □

Translation between $\mathbf{CL}(\mathbf{S})$ and $\mathbf{CL}_n(\mathbf{S})$

From an unlabeled term in $\mathbf{CL}(\mathbf{S})$, we can create a consistently labeled term in $\mathbf{CL}_n(\mathbf{S})$ (for large enough n) by just tagging each node:

6.4.6. Definition. The mapping $\mathrm{tag}(\cdot) : \mathbf{CL}(\mathbf{S}) \to \mathbf{CL}_n(\mathbf{S})$ is performed by labeling each non-leaf subterm X with $d_{\mathrm{r}}(X)$.

6.4.7. Definition. The mapping $\mathrm{forget}(\cdot) : \mathbf{CL}_n(\mathbf{S}) \to \mathbf{CL}(\mathbf{S})$ just deletes all labels.

Obviously, reductions in $\mathbf{CL}_n(\mathbf{S})$ can be unlabeled.

6.4.8. Proposition. If $X \to_{\mathbf{CL}_n(\mathbf{S})} Y$ for some n, then $\mathrm{forget}(X) \to \mathrm{forget}(Y)$.

Can reductions in $\mathbf{CL}(\mathbf{S})$ be labeled? Yes, at least as long as the right depth of their redexes is bounded.

6.4.9. Proposition. Assume there is (finite or infinite) reduction

$$X_1 \stackrel{R_1}{\to} X_2 \stackrel{R_2}{\to} X_3 \stackrel{R_3}{\to} \dots$$

in $\mathbf{CL}(\mathbf{S})$ where $R_1, R_2, R_3, \dots$ are the redexes being reduced.
Assume that for all k, $d_{\mathrm{r}}(R_k) \leq n$. Then there is a reduction

$$X_1' \stackrel{R_1'}{\to}_{\mathbf{CL}_n(\mathbf{S})} X_2' \stackrel{R_2'}{\to}_{\mathbf{CL}_n(\mathbf{S})} X_3' \stackrel{R_3'}{\to}_{\mathbf{CL}_n(\mathbf{S})} \cdots$$

in $\mathbf{CL}_n(\mathbf{S})$ with $X_1' = \mathrm{tag}\,(X_1)$ and for all k, $\mathrm{forget}\,(X_k') = X_k$.

Proof. Each redex R_k' is a consistent subterm, by proposition 6.4.5. That's why the top label of R_k' is less or equal to $d_{\mathrm{r}}(R_k')$ but that is equal to $d_{\mathrm{r}}(R_k) \leq n$. Therefore R_k' really was an $\mathbf{CL}_n(\mathbf{S})$-redex. □

6.4.10. Theorem. Assume there is an infinite reduction

$$X_1 \stackrel{R_1}{\to} X_2 \stackrel{R_2}{\to} X_3 \stackrel{R_3}{\to} \dots$$

in $\mathbf{CL}(\mathbf{S})$ where $R_1, R_2, R_3, \dots$ are the redexes being reduced. Then the sequence $d_{\mathrm{r}}(R_k)$ is not bounded.

Proof. Otherwise, proposition 6.4.9 would apply. Then all reductions could be placed in some $\mathbf{CL}_n(\mathbf{S})$. But this is a terminating system by theorem 6.4.2, so the reduction chain could not be infinite. □

6.4.11. Corollary. $\mathbf{CL}(\mathbf{S})$ admits no ground loops.

Proof. A ground loop would give an infinite reduction chain with only finitely many different redexes, so their right depth would be bounded. This is impossible by theorem 6.4.10 □

Sadly, this method doesn't directly allow to show the absence of non-ground loops because they could produce a sequence of redexes with unbounded right depth.

Chapter 7

Top-termination

7.1 Introduction

If a TRS is non-terminating, then there are infinite reductions. These reductions can be classified.

7.1.1. Definition. An reduction sequence is *perpetual* iff it rewrites the root infinitely often.

A desirable property of non-terminating TRS is the absence of perpetual reductions.

7.1.2. Definition. A TRS R is top-terminating iff there is no perpetual reduction.

Trivially, every terminating TRS is top-terminating.
While termination guarantees that redexes eventually vanish, top termination does allow infinite reductions but redex positions move deeper and deeper. This means that for each level of the term, eventually reductions will only change nodes below that level. That's why we can speak of the limit of an infinite reduction.
In this chapter, we will show that **CL**(**S**) indeed is top-terminating. We conjecture that limits of infinite reductions can be finitely described, and these descriptions should allow solving the word problem in **CL**(**S**). But at the moment no proof seems within reach.

Since **CL**(**S**) is a combinatory TRS, it enjoys the standardization property, which loosely says "you can do all head reductions first". We will show that in **CL**(**S**), there are no perpetual *head* reductions. From that it follows that there are no perpetual reductions at all.

For a start, we'll give a very short proof for $(\mathsf{A}\mathsf{A})\mathsf{A}$ not having a head normal form. Indeed no $X \in (\mathcal{Q}\mathcal{Q})\mathcal{Q}$ has. Yet this proof does *not* say anything on top termination of the head reduction sequence, so we need a more involved argument.
Our plan is to assume that there were a reduction chain that touched the root infinitely often. We prove theorem 7.3.6 that says that this reduction eventually contains a redex of the form $((\mathbf{S}\mathcal{Q})*)\mathcal{Q}$. This goes to $(\mathcal{Q}\mathcal{Q})\mathcal{Q}$ in the next step. Then we prove theorem 7.5.10 telling us that in a head reduction from $(\mathcal{Q}\mathcal{Q})\mathcal{Q}$ on, the root can only be touched finitely often.

7.2 A First Attempt At $\mathcal{Q}\mathcal{Q}\mathcal{Q}$

We begin by giving a different proof for $\mathsf{A}\mathsf{A}\mathsf{A}$ having an infinite head reduction. Note that $\mathsf{A} \in \mathcal{Q}$. We have in fact

7.2.1. Proposition. $\mathcal{Q}\mathcal{Q}\mathcal{Q}$ and $\mathbf{S}\mathcal{Q}*\mathcal{Q}$ have infinite head reduction.

Proof. Consider $X\mathcal{Q}\mathcal{Q}$ with $X \in \mathcal{Q}$.

1. $X \in \mathbf{S}\mathcal{Q}$. Then $X\mathcal{Q}\mathcal{Q} = \mathbf{S}\mathcal{Q}\mathcal{Q}\mathcal{Q} \to \mathcal{Q}\mathcal{Q}(\mathcal{Q}\mathcal{Q}) \subseteq \mathcal{Q}\mathcal{Q}\mathcal{Q}$.
2. X has a head normal form $\mathbf{S}**$. Then $X\mathcal{Q}\mathcal{Q} \twoheadrightarrow_h \mathbf{S}**\mathcal{Q}\mathcal{Q} \to_h *\mathcal{Q}(*\mathcal{Q})\mathcal{Q} \subseteq \mathcal{Q}\mathcal{Q}\mathcal{Q}$.
3. X has no head normal form. Then $X\mathcal{Q}\mathcal{Q}$ has an infinite head reduction as well, and the arguments of X are never touched.

In each case except the last one, we do one head reduction step that leads to an expression that is in $\mathcal{Q}\mathcal{Q}\mathcal{Q}$ again. This proves that $\mathcal{Q}\mathcal{Q}\mathcal{Q}$ has an infinite head reduction.
On the other hand, $\mathbf{S}\mathcal{Q}*\mathcal{Q} \to_h \mathcal{Q}\mathcal{Q}(*\mathcal{Q}) \subseteq \mathcal{Q}\mathcal{Q}\mathcal{Q}$. $\square$

Note that this proof does not say much about where the head redexes really are—compared to the explicit derivation of $\mathsf{A}\mathsf{A}\mathsf{A}$ given in example 5.5.2 which explicitly shows that the left spine grows longer and longer. Thus we really need some different ideas to show top-termination of $\mathbf{CL}(\mathbf{S})$.

7.3 Reaching $\mathcal{Q}\mathcal{Q}\mathcal{Q}$

7.3.1. Proposition. If $\mathbf{S}BCD \to_h^+ *FG$, then $D \trianglelefteq F$ and $CD \trianglelefteq G$.

Proof. The reduction starts $\mathbf{S}BCD \to_h BD(CD)$. We are going to show by induction on the length of the reduction chain that the required property remains invariant from now on. In each reduction step, the head redex could be at these positions:

1. at the top: $\mathbf{S}*FG \to_h *G(FG) = *F'G'$

   We have $D \triangleleft CD \trianglelefteq G = F'$ and $CD \trianglelefteq G \triangleleft G'$.

2. one step down the left spine: $(((\mathbf{S}\,*)\,*)\,F)\,G \to_h ((*\,F)\,(*\,F))\,G = (*\,F')\,G'$

   We have $D \trianglelefteq F \triangleleft F'$ and $C\,D \trianglelefteq G = G'$.

3. deeper on the left spine: $(*\,F)\,G \to_h (*\,F)\,G$. This does not change F and G.

$\square$

7.3.2. Proposition. If $*\,D \to_h^+ (\mathbf{S}\,F)\,G$, then $D \trianglelefteq F$ and $D \triangleleft G$.

Proof. Consider the first of these reductions that touches D. This reduction must exist. Its redex has the form $((\mathbf{S}\,*)\,*)\,D$. To this, apply proposition 7.3.1 $\square$

Assume now that there were a head reduction including infinitely many top redexes. This could be written as

$$V_{(1,1)} \twoheadrightarrow_h V_{(1,n_1)} \to_h U_{(1,1)} \twoheadrightarrow_h U_{(1,m_1)} \to_h$$
$$V_{(2,1)} \twoheadrightarrow_h V_{(2,n_1)} \to_h U_{(2,1)} \twoheadrightarrow_h \dots$$

where each of the $V_{(i,k)}$ is a top redex, and none of the $U_{(i,k)}$ is a redex.
In such a chain, each block of consecutive top redexes $V_{(i,*)}$ is finite. By assumption, there should still be infinitely many top redexes, so each of the blocks of consecutive non-top redexes $U_{(i,*)}$ must be finite as well.
We now want to compare $V_{(i,1)}$ to $V_{(i+1,1)}$, using $U_{(i,1)}$ as intermediate step.

7.3.3. Proposition. For all k, we have $V_{(k,1)} = (*\,B)\,C \to_h^+ (*\,D)\,E = U_{(k+1,1)}$ with $C \trianglelefteq D$ and $B\,C \trianglelefteq E$.

Proof. By proposition 7.3.1. $\square$

7.3.4. Proposition. For all k, we have $U_{(k,1)} = (*\,B)\,C \to_h^+ ((\mathbf{S}\,D)\,E)\,C = V_{(k+1,1)}$ with $B \trianglelefteq D$ and $B \trianglelefteq E$.

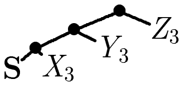
Proof. None of the reductions take place at the top, so C is not touched and we can apply proposition 7.3.2 to the left subterm of the root. $\square$

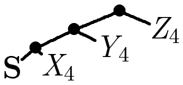
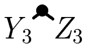
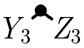
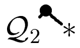
The last two propositions can be merged into

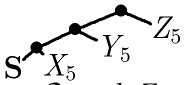
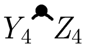
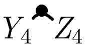
7.3.5. Proposition. For all k, we have $V_{(k,1)} = ((\mathbf{S}\,*)\,Y)\,Z \twoheadrightarrow_h ((\mathbf{S}\,X')\,Y')\,Z' = V_{(k+1,1)}$ with $Z \trianglelefteq X'$, $Z \triangleleft Y'$, and $Y\,Z \trianglelefteq Z'$.

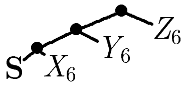
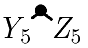
Now we apply this proposition a few times, starting with some arbitrary $V_{(1,1)} = ((\mathbf{S}\,X)\,Y)\,Z$.

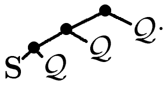
We obtain $V_{(2,1)} = ((\mathbf{S}\,X_2)\,Y_2)\,Z_2$ with $Y\,Z \trianglelefteq Z_2$. Therefore Z_2 cannot be $\mathbf{S}$, so it must be in $\mathcal{Q}_1$.

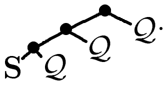
For $V_{(3,1)} =$  we have $Z_2 \lhd Y_3$. Since $Z_2 \in \mathcal{Q}_1$ this means $Y_3 \in \mathcal{Q}_2$.

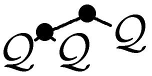
For $V_{(4,1)} =$  we know  $\unlhd Z_4$. But  $\in$  $\subseteq \mathcal{Q}$, leading to $Z_4 \in \mathcal{Q}$.

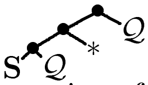
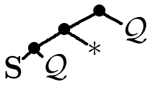
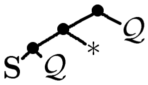
For $V_{(5,1)} =$  again $Z_4 \lhd Y_5$ and  $\unlhd Z_5$. Since $Z_4 \in \mathcal{Q}$ and  $\in$  $\subseteq \mathcal{Q}$ we know $Y_5 \in \mathcal{Q}$ and $Z_5 \in \mathcal{Q}$.

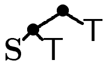
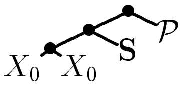
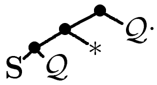
Finally, $V_{(6,1)} =$  has $Y_5 \unlhd X_6$ as well as $Z_5 \lhd Y_6$ and  $\unlhd Z_6$.

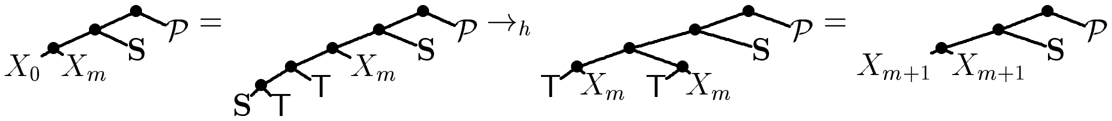
This says that $V_{(6,1)} \in$ . So we proved

7.3.6. Theorem. In a head reduction that has infinitely many steps at the top, eventually there is a top redex .

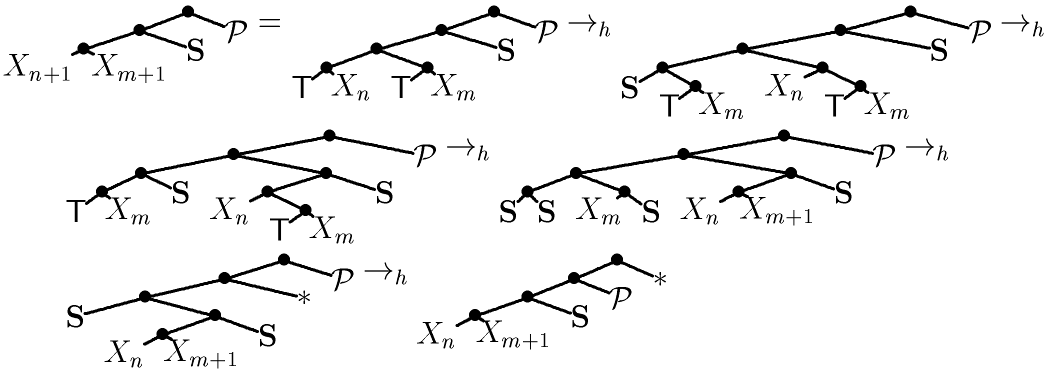
7.4 Avoiding 

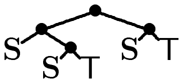
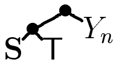
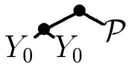
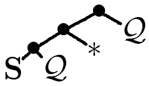
The previous arguments might have created the impression that all infinite head reductions eventually contain some  redex. But the existence of the  redex had only been shown under the assumption of a head reduction that has enough top steps. There are reductions that avoid the  pattern completely by just descending to the left soon enough.

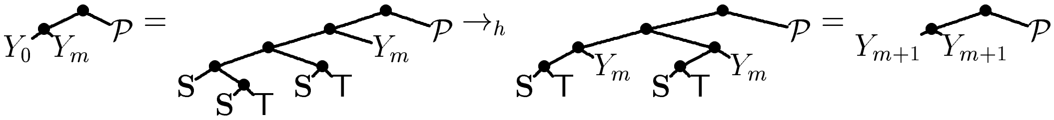
7.4.1. Example. $X_0 =$ ; $X_{n+1} =$ . Then  has an infinite head reduction where no redex has form . The reduction consists of the upward step



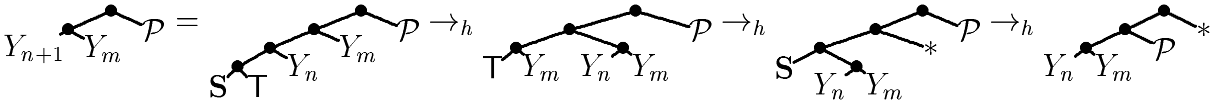
and the downward step



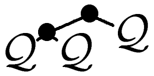
7.4.2. Example. $Y_0 =$ ; $Y_{n+1} =$ . Then  has an infinite head reduction where no redex has form . Here the upward step is



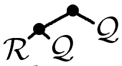
while the downward step goes



Yet these exceptions don't hurt because nevertheless the reductions are top-terminating.

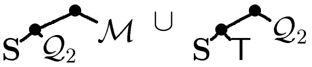
7.5 Inside 

We are going to show that each head reduction starting from  touches the root only finitely often.

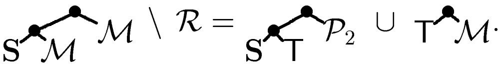
In order to do so, we present a set $\mathcal{R} \subseteq \mathcal{Q}$ such that  has the desired property immediately, i. e. no further knowledge (than being in $\mathcal{Q}$) about the two arguments of $\mathcal{R}$ is required.

Finally we deal with the complementary cases . Here we need to iterate the reasoning as some of the $\mathcal{Q}$ arguments might now get in front position. But while doing so, they also grow and therefore the process cannot be repeated infinitely often.

We give the definition now

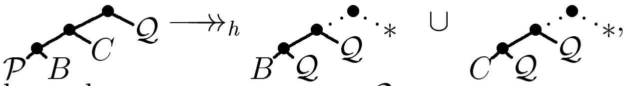
7.5.1. Definition. $\mathcal{R} =$ 

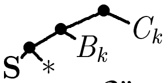
and justify it after preparing some tools. As an immediate consequence we obtain

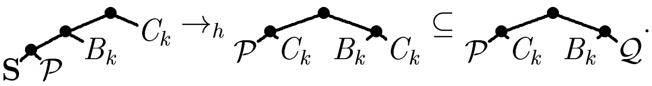
7.5.2. Definition. $\overline{\mathcal{R}} =$ .

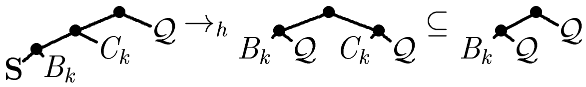
This can be checked with the program `RX` (see Appendix).

7.5.3. Proposition. If $B \in \mathcal{Q}$ and $C \in \mathcal{Q}$, then there is a reduction

, i. e. the result starts with B or C and is followed by at least two arguments $\mathcal{Q}$.

Proof. We show that in all redexes  that occur during this reduction, both B_k and C_k have the form "B or C, followed by some $\mathcal{Q}$".

This is invariant because .

We stop the reduction at . □

7.5.4. Proposition. $((\mathcal{Q}_2\,\mathcal{Q})\,\mathcal{Q})\,\mathcal{Q}$ has infinite head reduction where the root is rewritten only finitely often.

Proof. We have $\mathcal{Q}_2 = \mathcal{Q} \cup \mathsf{S}\,(\mathsf{S}\,\mathcal{P})$.
In the first case, in the reduction of $((\mathcal{Q}\,\mathcal{Q})\,\mathcal{Q})\,\mathcal{Q}$ the root is never rewritten because $(\mathcal{Q}\,\mathcal{Q})\,\mathcal{Q}$ alone has no head normal form.
Otherwise, $(((\mathsf{S}\,(\mathsf{S}\,\mathcal{P}))\,\mathcal{Q})\,\mathcal{Q})\,\mathcal{Q} \to_h (((\mathsf{S}\,\mathcal{P})\,\mathcal{Q})\,(\mathcal{Q}\,\mathcal{Q}))\,\mathcal{Q} \to_h ((\mathcal{P}\,(\mathcal{Q}\,\mathcal{Q}))\,(\mathcal{Q}\,(\mathcal{Q}\,\mathcal{Q})))\,\mathcal{Q} \subseteq ((\mathcal{P}\,(\mathcal{Q}\,\mathcal{Q}))\,(\mathcal{Q}\,\mathcal{Q}))\,\mathcal{Q}$.
By proposition 7.5.3, this goes to $((\mathcal{Q}\,\mathcal{Q})\,\mathcal{Q})\,\mathcal{Q}$ in each case, and so we are done. ☐

Now we look at $(X\,\mathcal{Q})\,\mathcal{Q}$ with $X \in \mathcal{Q}$.
If $X = \mathsf{S}\,X'$, then $(X\,\mathcal{Q})\,\mathcal{Q} = ((\mathsf{S}\,X')\,\mathcal{Q})\,\mathcal{Q} \to_h (X'\,\mathcal{Q})\,(\mathcal{Q}\,\mathcal{Q}) \subseteq (X'\,\mathcal{Q})\,\mathcal{Q}$ and we consider X' instead (this transition can happen only finitely often).
So assume $X \in (\mathcal{M}\,\mathcal{M})\,\mathcal{M}$. If it has no head normal form, then head reductions in $(X\,\mathcal{Q})\,\mathcal{Q}$ never touch the root of the term.
Therefore let X have head normal form $H \in (\mathsf{S}\,*)\,*$.
If $H \in \mathcal{R}$, then we will apply proposition 7.5.5 which comes next. Otherwise $H \in \overline{\mathcal{R}}$ and we will use proposition 7.5.9 which is stated and proved later.

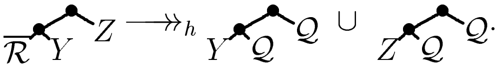
7.5.5. Proposition. $(\mathcal{R}\,\mathcal{Q})\,\mathcal{Q}$ has infinite head reduction where the root is rewritten only finitely often.

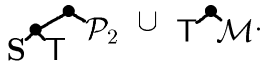
Proof. Recall $\mathcal{R} = (\mathsf{S}\,\mathsf{T})\,\mathcal{Q}_2 \cup (\mathsf{S}\,\mathcal{Q}_2)\,\mathcal{M}$.

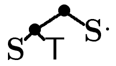
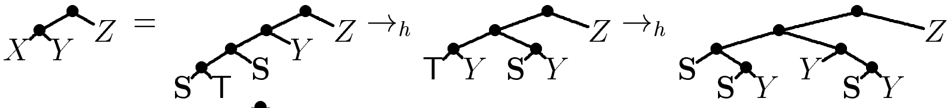
1. $(((\mathsf{S}\,\mathsf{T})\,\mathcal{Q}_2)\,\mathcal{Q})\,\mathcal{Q} \to_h ((\mathsf{T}\,\mathcal{Q})\,(\mathcal{Q}_2\,\mathcal{Q}))\,\mathcal{Q} \to_h ((\mathsf{S}\,(\mathcal{Q}_2\,\mathcal{Q}))\,*)\,\mathcal{Q} \to_h ((\mathcal{Q}_2\,\mathcal{Q})\,\mathcal{Q})\,(*\,\mathcal{Q}) \subseteq ((\mathcal{Q}_2\,\mathcal{Q})\,\mathcal{Q})\,\mathcal{Q}$
2. $(((\mathsf{S}\,\mathcal{Q}_2)\,*)\,\mathcal{Q})\,\mathcal{Q} \to_h ((\mathcal{Q}_2\,\mathcal{Q})\,(*\,\mathcal{Q}))\,\mathcal{Q} \subseteq ((\mathcal{Q}_2\,\mathcal{Q})\,\mathcal{Q})\,\mathcal{Q}$

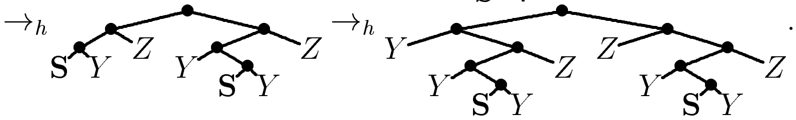
and we apply proposition 7.5.4 in both cases. ☐

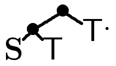
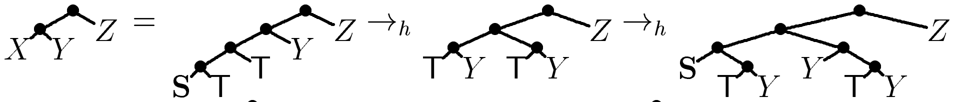
To deal with the case $(\overline{\mathcal{R}}\,\mathcal{Q})\,\mathcal{Q}$, we need more tools.

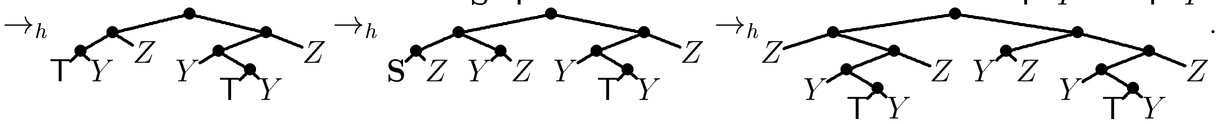
7.5.6. Proposition. If $Y \in \mathcal{Q}$ and $Z \in \mathcal{Q}$, then 

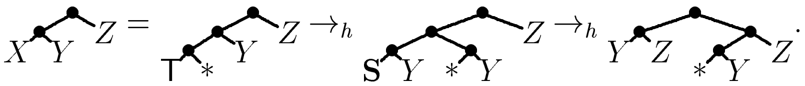
Proof. Recall $\overline{\mathcal{R}}$ = 

1. X =  Then 

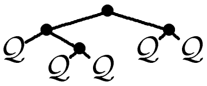


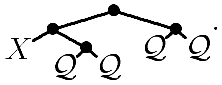
2. X =  Then 

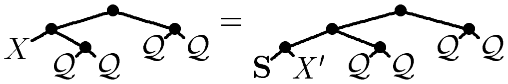


3. $X \in$  Then 

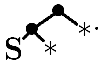
□

7.5.7. Proposition.  has infinite head reduction where the root is rewritten only finitely often.

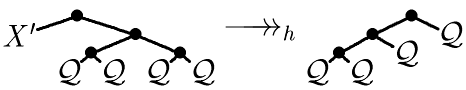
Proof. Let the expression be 

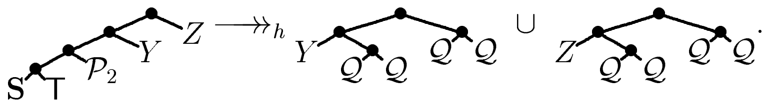
If $X =$ , then 

 and repeat the argument for X'. This transition can only happen finitely often.

Now $X \in$  If this has no head normal form, then we're done. Assume it has head normal form $X' \in$ 

If $X' \in \mathcal{R}$ then the claim is established by proposition 7.5.5.

If $X' \in \overline{\mathcal{R}}$ then by proposition 7.5.6  and again the root cannot be rewritten any more. □

7.5.8. Proposition. If $Y \in \mathcal{Q}$ and $Z \in \mathcal{Q}$,
then 

Proof. Follows from the proof of proposition 7.5.6. □

7.5.9. Proposition. $\overline{\mathcal{R}}\mathcal{Q}\mathcal{Q}$ has infinite head reduction where the root is rewritten only finitely often.

Proof. Let $XYZ \in \overline{\mathcal{R}}\mathcal{Q}\mathcal{Q}$. If $X \in \mathbf{S}\mathsf{T}\mathcal{P}_2$, then $\mathbf{S}\mathsf{T}\mathcal{P}_2\mathcal{Q}\mathcal{Q} \twoheadrightarrow_h \mathcal{Q}(\mathcal{Q}\mathcal{Q})(\mathcal{Q}\mathcal{Q})$ by proposition 7.5.8 and this has the required property by proposition 7.5.7.
Therefore assume $X \in \mathsf{T}*$.
Then $XY\mathcal{Q} = \mathsf{T}*Y\mathcal{Q} \to_h \mathbf{S}Y(*Y)\mathcal{Q} \to_h Y\mathcal{Q}(*Y\mathcal{Q}) \subseteq Y\mathcal{Q}(\mathcal{Q}\mathcal{Q})$.
If $Y = \mathbf{S}Y'$, then $Y\mathcal{Q}(\mathcal{Q}\mathcal{Q}) = \mathbf{S}Y'\mathcal{Q}(\mathcal{Q}\mathcal{Q}) \to_h Y'(\mathcal{Q}\mathcal{Q})(\mathcal{Q}(\mathcal{Q}\mathcal{Q})) \subseteq \mathcal{Q}(\mathcal{Q}\mathcal{Q})(\mathcal{Q}\mathcal{Q})$ and we apply proposition 7.5.7.
So let $Y \in \mathcal{M}\mathcal{M}\mathcal{M}$. If this has no head normal form, then we are done. Assume it has a head normal form $Y' \in \mathbf{S}*$.
If $Y' \in \mathcal{R}$ then we use proposition 7.5.5. Therefore let $Y' \in \overline{\mathcal{R}}$.
If $Y' \in \mathbf{S}\mathsf{T}\mathcal{P}_2$, then by proposition 7.5.8, $Y'\mathcal{Q}(\mathcal{Q}\mathcal{Q}) \subseteq Y'\mathcal{Q}\mathcal{Q} \twoheadrightarrow_h \mathcal{Q}(\mathcal{Q}\mathcal{Q})(\mathcal{Q}\mathcal{Q})$ and we now use proposition 7.5.7.
Thus, assume $Y' \in \mathsf{T}*$. We have $Y'\mathcal{Q}(\mathcal{Q}\mathcal{Q}) = \mathsf{T}*\mathcal{Q}(\mathcal{Q}\mathcal{Q}) \to_h \mathbf{S}\mathcal{Q}(*\mathcal{Q})(\mathcal{Q}\mathcal{Q})$
$\to_h \mathcal{Q}(\mathcal{Q}\mathcal{Q})(*\mathcal{Q}(\mathcal{Q}\mathcal{Q})) \subseteq \mathcal{Q}(\mathcal{Q}\mathcal{Q})(\mathcal{Q}\mathcal{Q})$. Then we are done, using proposition 7.5.7. □

So we proved

7.5.10. Theorem. $\mathcal{Q}\mathcal{Q}\mathcal{Q}$ has infinite head reduction where the root is rewritten only finitely often.

In light of theorem 7.3.6, this completes the proof of top-termination for $\mathbf{CL}(\mathbf{S})$.

7.6 Infinite normal forms

(This section is just an aside; it will not be used in the sequel. We use concepts and results from infinitary rewriting, as described in [KKSdV95], without repeating their definitions here.)

Top-termination guarantees that infinite reductions are *strongly convergent.*

(That's good. We don't have to reason about Cauchy convergent reductions as defined by Dershowitz in [DKP91].)

Since $\mathbf{CL}(\mathbf{S})$ is orthogonal and non-collapsing, the infinitary Church-Rosser property holds, and we can take the limit of a fair reduction (that does not disregard some redex completely) as the unique (possibly infinite) normal form of an $\mathbf{S}$-term.

We conjecture that such limits of $\mathbf{S}$-terms can always be described by patterns similar to that given in the examples 5.5.1 and 5.5.2 above.

Moreover, they seem to have enough leaves at finite depths so that the word problem in $\mathbf{CL}(\mathbf{S})$ might be decidable by the comparison of these patterns.

7.6.1. Example. We return to $\mathsf{S}\ \mathsf{T}\ \mathsf{T}\ (\mathsf{S}\ \mathsf{T}\ \mathsf{T})$ (example 5.5.1) with its X_k patterns.

$$f\,(m,n) = \lim X_m X_n$$
$$f\,(m+1,n) = \lim (\mathsf{T} X_m) X_n = \lim (\mathsf{S} X_n)(X_m X_n) = (\mathsf{S} X_n)\,\boxed{\lim X_m X_n} = (\mathsf{S} X_n)\, f\,(m,n)$$
$$f\,(0,n) = \lim ((\mathsf{S}\,\mathsf{T})\,\mathsf{T})\, X_n = \lim (\mathsf{T} X_n)(\mathsf{T} X_n) = f\,(n+1,n+1)$$

If we drop all subscripts and arguments, then we obtain a grammar for a set of infinite trees.

$$X \to (\mathsf{S}\,\mathsf{T})\,\mathsf{T} \ \cup\ \mathsf{T} X;\ f \to (\mathsf{S} X)\, f$$

This grammar, possibly with some decoration, should be sufficient to characterize $\mathsf{S}\ \mathsf{T}\ \mathsf{T}\ (\mathsf{S}\ \mathsf{T}\ \mathsf{T})$.

The limits so obtained don't contain redexes. Still they might contain subterms with infinite left spines. Such subterms *are* in head normal form, where *head normal form* is defined as a term that cannot be reduced to a redex.

On the other hand, terms M where $\lim M$ has an infinite left spine behave like *unsolvable* terms because $\lim M N = \lim M (\lim N)$. This property will be used in the next example.

7.6.2. Example. Now we want to describe $\lim (\mathsf{A}\ \mathsf{A}\ \mathsf{A})$ (example 5.5.2).

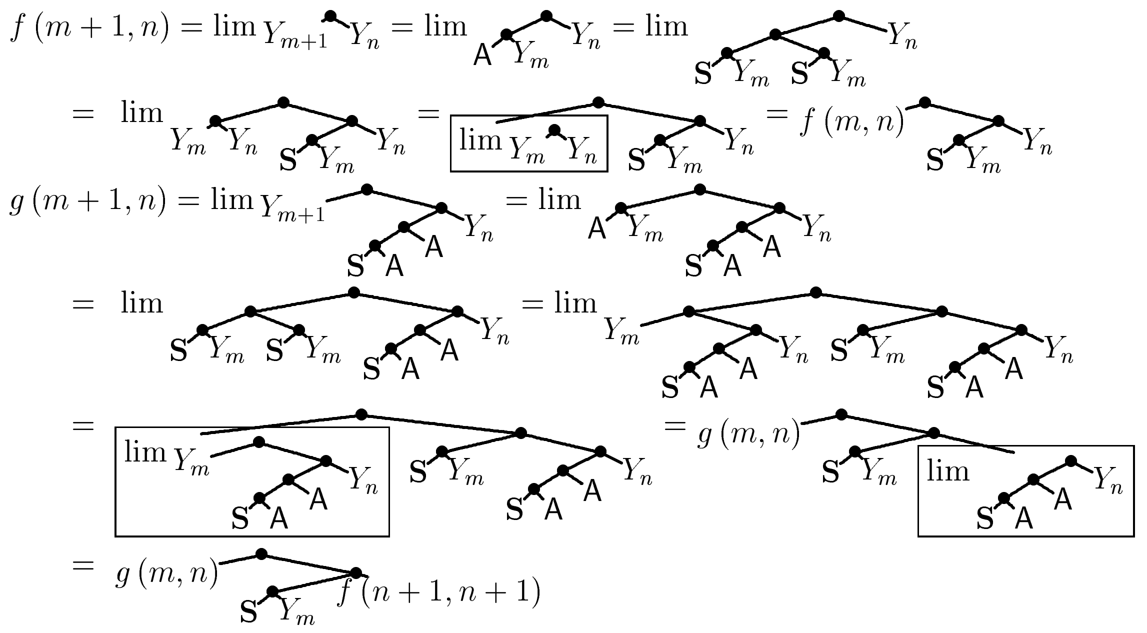
Because of $(\mathsf{A}\,\mathsf{A})\,\mathsf{A} \twoheadrightarrow_h (Y_1 Y_1)\,((\mathsf{S}\,((\mathsf{S}\,\mathsf{A})\,\mathsf{A}))\, Y_0)$, this amounts to determining $f\,(1,1)$ where

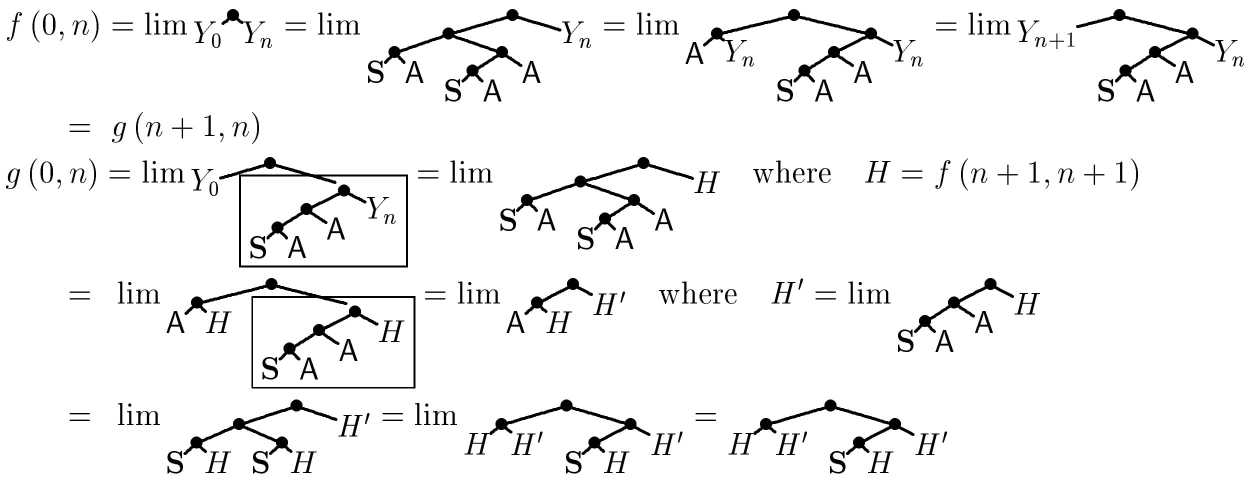
7.6.3. Definition. $f\,(m,n) = \lim Y_m Y_n$ and the auxiliary $g\,(m,n) = \lim Y_m\,(((\mathsf{S}\,\mathsf{A})\,\mathsf{A})\, Y_n)$

As $Y_m Y_n \in (Q\,Q)\,Q$ and $Y_m\,(((\mathsf{S}\,\mathsf{A})\,\mathsf{A})\, Y_n) \in (Q\,Q)\,Q$, both have an infinite head reduction, as has just been shown. Therefore both $f\,(m,n)$ and $g\,(m,n)$ have an infinite left spine. The following is immediate

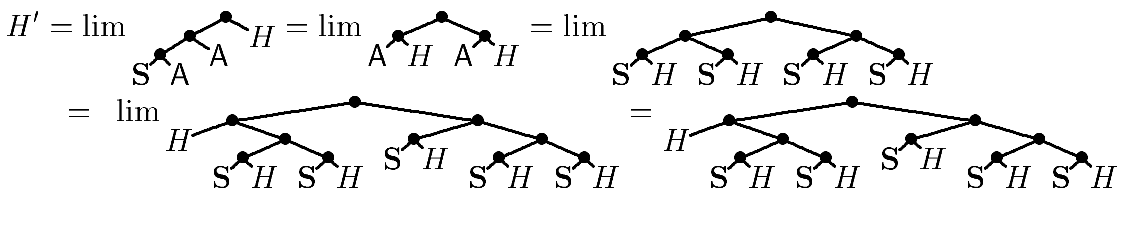
7.6.4. Lemma. $\lim ((\mathsf{S}\,\mathsf{A})\,\mathsf{A})\, Y_n = \lim (\mathsf{A} Y_n)(\mathsf{A} Y_n) = \lim Y_{n+1} Y_{n+1} = f\,(n+1,n+1)$

The reductions already given can be translated to





The names H and H' are used only locally, for convenience of notation. We still need to determine H', using the fact that H has an infinite left spine.



Again, dropping all subscripts and arguments in Y_*, $f(*,*)$, $g(*,*)$, H, H' results in a grammar for infinite trees.

Chapter 8

Decidability of Normalization

8.1 Introduction

To decide whether an arbitrary **S**-term from $\mathcal{M}$ normalizes, it is enough to have a procedure that decides this for the case $\mathcal{N}{\bullet}\mathcal{N}$, i.e. it may assume that both children of the root already are in normal form. If they were not, we could first apply the procedure to these subterms recursively and compute their normal forms if the procedure says they exist.

Since $\mathbf{CL}(\mathbf{S})$ is orthogonal and non-erasing, it enjoys the conservation theorem which says that if there is some infinite reduction from X, then all reductions from X are infinite. So it is sufficient to exhibit one infinite reduction in order to show that X has no normal form.

We start by giving three classes of expressions with infinite reductions.

Continuing, we will construct subsets of $\mathcal{N}{\bullet}\mathcal{N}$ that contain terms with similar reduction properties. By some sort of magic, these subsets cover all of $\mathcal{N}{\bullet}\mathcal{N}$.

Actually a small range of terms in $\mathcal{N}{\bullet}\mathcal{N}$ remains uncovered but we will show that after some reductions these terms move into the colonized areas.

The construction begins by partitioning the set $\mathcal{N}$ itself into classes $\mathcal{N}_0$, $\mathcal{N}_1$ and $\mathcal{N}_2$, and then combines pairs of these. It turns out that $\mathcal{N}_0$ is responsible for finite reductions, and $\mathcal{N}_2$ for infinite ones.

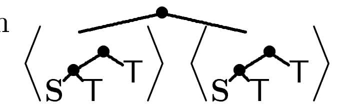
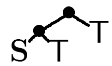
During the process, the case $\mathcal{N}_1{\bullet}\mathcal{N}_1$ needs special attention. Here, a further partition of $\mathcal{N}_1$ into the sets $\mathcal{L}_0$, $\mathcal{L}_1$ and $\mathcal{L}_2$ is called for. By considering all possible combinations of these three sets, the investigation is completed.

Most of these constructions have been found empirically, by staring at thousands of examples. So the presentation here necessarily generates some ad-hoc feeling. A straight top-down approach might be more desirable, but that seems to transcend today's knowledge of $\mathbf{CL}(\mathbf{S})$.

8.2 Infinite Reductions

We now give three classes of expressions not having a normal form.

They are all of the shape $\langle * \rangle{\bullet}*$.

We start by generalizing the reduction shown in example 5.5.1. We note that each redex has the form  and that  is the smallest term in the set $\mathcal{Q}_2{\bullet}\mathcal{Q}_1$.

8.2.1. Definition. $F = \left\langle \mathcal{Q}_2 \mathcal{Q}_1 \right\rangle$

We are going to show that $FF \twoheadrightarrow \infty$. We need

8.2.2. Proposition. $\mathcal{Q}_1 F \subseteq F$

Proof. We use the fact that $\mathcal{Q}_1 = \mathsf{T} \cup \mathcal{Q}_2$. Let $X \in \mathcal{Q}_1; Y \in F$.

1. $X = \mathsf{T}$. Then $XY \in \mathsf{S}{*}F \subseteq \langle F \rangle \subseteq F$.
2. $X \in \mathcal{Q}_2$. As $Y \in F \subseteq \mathcal{Q}_1$, we have $XY \in \mathcal{Q}_2 \mathcal{Q}_1 \subseteq F$.

□

This will now be used in the proof of

8.2.3. Proposition. $\left\langle \mathcal{Q}_2 \mathcal{Q}_1 \right\rangle \left\langle \mathcal{Q}_2 \mathcal{Q}_1 \right\rangle \twoheadrightarrow \infty$

Proof. We have $\left\langle \mathcal{Q}_2 \mathcal{Q}_1 \right\rangle \left\langle \mathcal{Q}_2 \mathcal{Q}_1 \right\rangle \hookrightarrow\!\!\!\!\twoheadrightarrow \mathcal{Q}_2 \mathcal{Q}_1 F \ni XYZ$.
If $X \twoheadrightarrow \infty$, then the $XYZ \twoheadrightarrow \infty$. So assume X has a normal form.
It could have shape $\mathsf{S}{*}$ or $\mathsf{S}{*}{*}$.

1. $X \twoheadrightarrow \mathsf{S}\mathcal{Q}_1$. $\mathsf{S}\mathcal{Q}_1 Y F \to \mathcal{Q}_1 F (Y F) \subseteq \mathcal{Q}_1 F (\mathcal{Q}_1 F) \subseteq FF \twoheadrightarrow \infty$
2. $X \twoheadrightarrow \mathsf{S}{*}{*}$. $\mathsf{S}{*}{*}\mathcal{Q}_1 F \to ({*}\mathcal{Q}_1)({*}\mathcal{Q}_1)F$.

   By corollaries 5.2.8 and 5.2.7 we have ${*}\mathcal{Q}_1 \subseteq \mathcal{Q}_2 \subseteq \mathcal{Q}_1$,
   so $({*}\mathcal{Q}_1)({*}\mathcal{Q}_1)F \subseteq \mathcal{Q}_2 \mathcal{Q}_1 F \subseteq FF \twoheadrightarrow \infty$.

In either case we find an infinite reduction because we've done one reduction step whose result again has the form FF.

□

The following pattern turns out to be a generalization of $\mathsf{A}\mathsf{A}\mathsf{A} \twoheadrightarrow \infty$. We recall that $\mathsf{A} \in \mathcal{Q}_k$ for any k.

8.2.4. Proposition. $\left\langle \mathcal{Q}_3 \mathcal{Q}_2 \right\rangle \mathcal{Q}_1 \twoheadrightarrow \infty$

Proof. Let $\left\langle \mathcal{Q}_3 \mathcal{Q}_2 \right\rangle \mathcal{Q}_1 \hookrightarrow\!\!\!\!\twoheadrightarrow \mathcal{Q}_3 \mathcal{Q}_2 \mathcal{Q}_1 \ni XYZ$.
If $X \twoheadrightarrow \infty$, then the $XYZ \twoheadrightarrow \infty$. So assume X has a normal form. It could be $\mathsf{S}\mathcal{Q}_2$ or $\mathsf{S}{*}{*}$.

1. $X \twoheadrightarrow \mathbf{S}\,\mathcal{Q}_2$. $((\mathbf{S}\,\mathcal{Q}_2)\,\mathcal{Q}_2)\,\mathcal{Q}_1 \to (\mathcal{Q}_2\,\mathcal{Q}_1)\,(\mathcal{Q}_2\,\mathcal{Q}_1) \subseteq F\,F \twoheadrightarrow \infty$ by proposition 8.2.3

2. $X \twoheadrightarrow (\mathbf{S}\,*)\,*$. $(((\mathbf{S}\,*)\,*)\,\mathcal{Q}_2)\,\mathcal{Q}_1 \to ((*\,\mathcal{Q}_2)\,(*\,\mathcal{Q}_2))\,\mathcal{Q}_1$ By corollaries 5.2.8 and 5.2.7 we have $*\,\mathcal{Q}_2 \subseteq \mathcal{Q}_3 \subseteq \mathcal{Q}_2$, so $((*\,\mathcal{Q}_2)\,(*\,\mathcal{Q}_2))\,\mathcal{Q}_1 \subseteq (\mathcal{Q}_3\,\mathcal{Q}_2)\,\mathcal{Q}_1 \twoheadrightarrow \infty$ by recursive application of proposition 8.2.4.

$\square$

Finally we mention a variant of the $(\mathcal{Q}_3\,\mathcal{Q}_2)\,\mathcal{Q}_1$ pattern:

8.2.5. Proposition. $\left\langle (\mathbf{S}\,\mathcal{Q}_3)\,\mathcal{Q}_0 \right\rangle \mathcal{Q}_2 \twoheadrightarrow \infty$

Proof. $\left\langle (\mathbf{S}\,\mathcal{Q}_3)\,\mathcal{Q}_0 \right\rangle \mathcal{Q}_2 \hookrightarrow\!\!\!\!\twoheadrightarrow ((\mathbf{S}\,\mathcal{Q}_3)\,\mathcal{Q}_0)\,\mathcal{Q}_2 \to (\mathcal{Q}_3\,\mathcal{Q}_2)\,(\mathcal{Q}_0\,\mathcal{Q}_2) \in (\mathcal{Q}_3\,\mathcal{Q}_2)\,\mathcal{Q}_1$. Then apply proposition 8.2.4.

$\square$

8.2.6. Remark. We have shown

$$\left\langle \mathcal{Q}_2\,\mathcal{Q}_1 \right\rangle \left\langle \mathcal{Q}_2\,\mathcal{Q}_1 \right\rangle \twoheadrightarrow \infty; \quad \left\langle \mathcal{Q}_3\,\mathcal{Q}_2 \right\rangle \mathcal{Q}_1 \twoheadrightarrow \infty; \quad \left\langle (\mathbf{S}\,\mathcal{Q}_3)\,\mathcal{Q}_0 \right\rangle \mathcal{Q}_2 \twoheadrightarrow \infty$$

Is it a coincidence that the sum of the indices of these expression is "nearly" always 6? There should be a better way to describe these patterns.

8.3 The case $\mathcal{N}\,\mathcal{N}$

We will classify the set $\mathcal{N}$ of normal forms according the existence of certain subterms that might or might not occur during the reduction of $X\,a$ where $X \in \mathcal{N}$ and a is a variable.
The most restricted class is

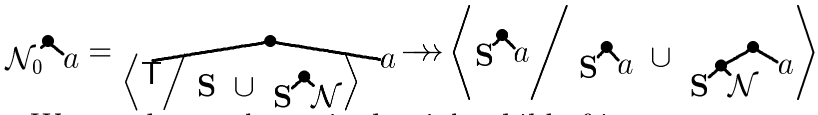
8.3.1. Definition. $\mathcal{N}_0 = \left\langle \mathsf{T} \middle/ \mathbf{S} \;\cup\; \mathbf{S}\,\mathcal{N} \right\rangle$

8.3.2. Example. $\mathsf{A} = (\mathbf{S}\,\mathbf{S})\,\mathbf{S} = \mathsf{T}\,\mathbf{S} \in \langle \mathsf{T} / \mathbf{S} \rangle \subseteq \mathcal{N}_0$.

A term from $\mathcal{N}_0$ never moves its argument into functional position:

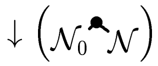
8.3.3. Example. $\mathsf{A}\,a = ((\mathbf{S}\,\mathbf{S})\,\mathbf{S})\,a \to (\mathbf{S}\,a)\,(\mathbf{S}\,a)$

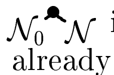
8.3.4. Proposition. $\mathcal{N}_0\,a$ has a normal form, and $\mathcal{N}_0\,a \not\twoheadrightarrow a\,*$

Proof.  and no further reductions can happen. We see that a always is the right child of its parent.

□

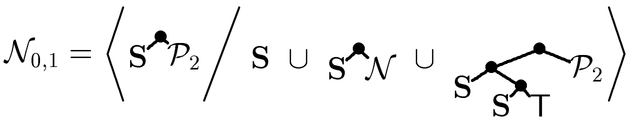
Since $\mathcal{N}_0$ effectively ignores its argument, we have

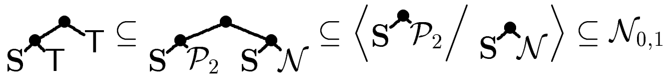
8.3.5. Proposition. 

Proof. The normal form of  is that of  with a replaced by $\mathcal{N}$. No reductions can happen inside $\mathcal{N}$, because it already is a normal form. No new redexes can be created.

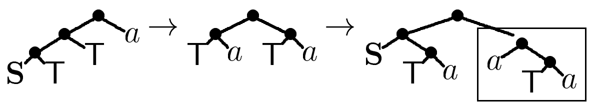
□

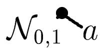
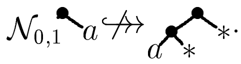
A slightly more extended class is

8.3.6. Definition. 

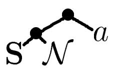
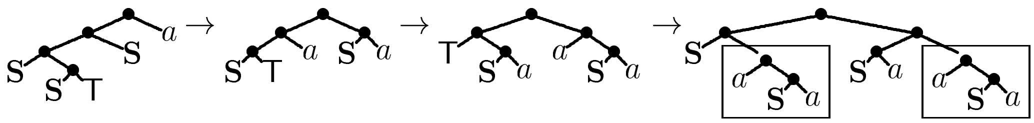
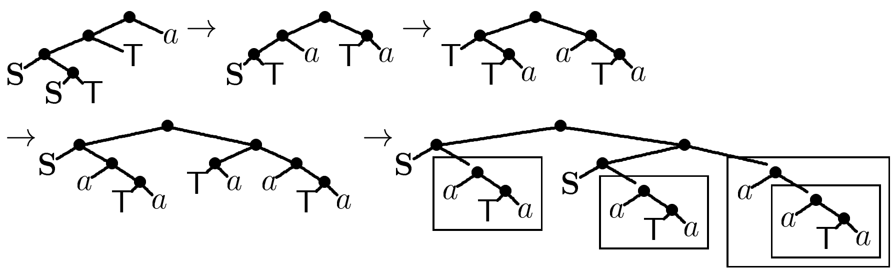
8.3.7. Example. 

A term from $\mathcal{N}_{0,1}$ might move its argument into functional position, but there at most one argument is supplied.

8.3.8. Example. 

8.3.9. Proposition.  has a normal form, and 

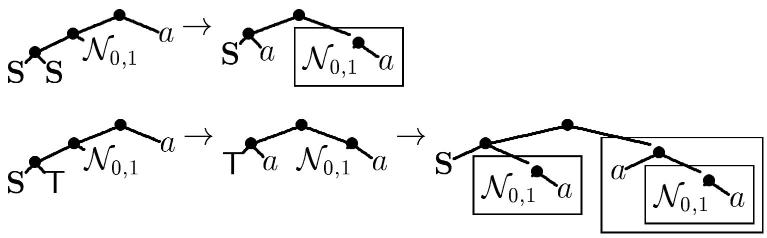
Proof. The base cases are

1.  is in normal form.
2.  is in normal form.
3. 
4. 

where the claim can be verified at the boxed subterms.

The inductive steps are

1.
2.



$\square$

If we look at the proof of proposition 8.3.9, we notice

8.3.10. Corollary. If $\mathcal{N}_{0,1}\,a \hookrightarrow\!\!\!\rightarrow a\,X$, then $a \lhd X$.

which will be used later on.

Proposition 8.3.9 will now be used to exhibit another set of normalizing terms:

8.3.11. Proposition. $\downarrow \left(\mathcal{N}_{0,1}\,\mathcal{N}_0\right)$

Proof. First reduce $\mathcal{N}_{0,1}\,\mathcal{N}_0$ as if $\mathcal{N}_0$ were a variable. By proposition 8.3.9, the only redexes that might arise are of the form $\mathcal{N}_0\,*$. They might be nested, but all of them can be reduced to normal form by proposition 8.3.5.

$\square$

By definition we have $\mathcal{N}_0 \subseteq \mathcal{N}_{0,1}$, and for later convenience we define

8.3.12. Definition. $\mathcal{N}_1 = \mathcal{N}_{0,1} \setminus \mathcal{N}_0; \mathcal{N}_2 = \mathcal{N} \setminus \mathcal{N}_{0,1}$.

So $\mathcal{N}$ is the disjoint sum of $\mathcal{N}_0$, $\mathcal{N}_1$, and $\mathcal{N}_2$.

Essentially, $\mathcal{N}_0$ is responsible for normalization, and $\mathcal{N}_2$ for non-normalization, as will be highlighted in the next two statements.

8.3.13. Proposition. $\mathcal{N}_2\,\mathcal{Q}_2 \twoheadrightarrow \infty$.

Proof. By a manipulation of rational tree languages it can be shown that

$$\mathcal{N}_2 \subseteq \left\langle \mathcal{Q}_3\,\mathcal{Q}_2 \right\rangle \cup \left\langle (\mathsf{S}\,\mathcal{Q}_3)\,\mathcal{Q}_0 \right\rangle$$

This is exactly the right form to apply propositions 8.2.4 and 8.2.5.

(Moreover, this is exactly the kind of statement that can be verified by the program `RX`.) $\square$

8.3.14. Proposition. $\mathcal{N}_1\,\mathcal{N}_2 \twoheadrightarrow \infty$.

Proof. By corollary 8.3.10, somewhere in a reduction from $\mathcal{N}_1\,\mathcal{N}_2$ there is a subexpression $\mathcal{N}_2\,X$ with some $\mathcal{N}_2$ contained in X. Clearly then $X \in \mathcal{Q}_2$, so we have $\mathcal{N}_2\,\mathcal{Q}_2$ and can apply proposition 8.3.13.

$\square$

Combining propositions 8.3.5, 8.3.11 on one hand and 8.3.13, 8.3.14 on the other, the cases still missing in the analysis of $\mathcal{N}\,\mathcal{N}$ are $\mathcal{N}_2\,\mathcal{P}_2$ and $\mathcal{N}_1\,\mathcal{N}_1$.

8.4 The sub-case $\mathcal{N}_1\,\mathcal{N}_1$

We look deeper into the set $\mathcal{N}_1$. We single out the subset of terms that move their argument a into functional position but which only supply them with the argument $\mathsf{S}\,a$.

8.4.1. Definition. $\mathcal{L}_0 = \left\langle \mathsf{T} \middle/ \mathsf{S}\,(\mathsf{S}\,\mathcal{P}_2)\,\mathsf{S} \right\rangle$

8.4.2. Proposition. $\mathcal{L}_0\,a$ has a normal form, and $\left\{Z \mid \mathcal{L}_0\,a \hookrightarrow\!\!\!\!\twoheadrightarrow a\,Z\right\} = \{\mathsf{S}\,a\}$.

Proof. The base cases are

1. $\mathsf{S}\,\mathsf{T}\,\mathsf{S}\,a \to \mathsf{T}\,a\,(\mathsf{S}\,a) \to \mathsf{S}\,(\mathsf{S}\,a)\,\boxed{a\,(\mathsf{S}\,a)}$

2. $\mathsf{S}\,(\mathsf{S}\,\mathsf{T})\,\mathsf{S}\,a \to \mathsf{S}\,\mathsf{T}\,a\,(\mathsf{S}\,a) \to \mathsf{T}\,(\mathsf{S}\,a)\,(a\,(\mathsf{S}\,a)) \to \mathsf{S}\,\boxed{a\,(\mathsf{S}\,a)}\,(\mathsf{S}\,a\,\boxed{a\,(\mathsf{S}\,a)})$

Then inductive step reads $\mathsf{T}\,\mathcal{L}_0\,a \to \mathsf{S}\,a\,\boxed{\mathcal{L}_0\,a}$.

□

Using $\mathcal{L}_0$, we find another class of normalizing terms.

8.4.3. Proposition. $\downarrow\left(\mathcal{L}_0\,\mathcal{N}_1\right)$.

Proof. Reduce $\mathcal{L}_0\,\mathcal{N}_1$ as if $\mathcal{N}_1$ were a variable. By proposition 8.4.2, the only redexes that may arise have the form $\mathcal{N}_1\,(\mathsf{S}\,\mathcal{N}_1)$. But $\mathsf{S}\,\mathcal{N}_1 \subseteq \mathsf{S}\,\mathcal{N} \subseteq \mathcal{N}_0$, so $\mathcal{N}_1\,(\mathsf{S}\,\mathcal{N}_1) \subseteq \mathcal{N}_1\,\mathcal{N}_0$ and therefore they normalize by proposition 8.3.11.

□

Another useful subset of $\mathcal{N}_1$ is

8.4.4. Definition. $\mathcal{L}_1 = \left\langle \mathsf{T} \middle/ \mathsf{S}\,\mathsf{T}\,(\mathsf{S}\,\mathcal{P}_2) \right\rangle$.

It has the following property:

8.4.5. Proposition. $\mathcal{L}_1\,a$ has a normal form, and $\left\{Z \mid \mathcal{L}_1\,a \hookrightarrow\!\!\!\!\twoheadrightarrow a\,Z\right\} = \left\{\mathsf{S}\,\mathcal{P}_2\,a\right\}$.

Proof. The base case is $\mathsf{S}\,\mathsf{T}\,(\mathsf{S}\,\mathcal{P}_2)\,a \to \mathsf{T}\,a\,(\mathsf{S}\,\mathcal{P}_2\,a) \to \mathsf{S}\,(\mathsf{S}\,\mathcal{P}_2\,a)\,\boxed{a\,(\mathsf{S}\,\mathcal{P}_2\,a)}$.

The inductive step is $\mathsf{T}\,\mathcal{L}_1\,a \to \mathsf{S}\,a\,\boxed{\mathcal{L}_1\,a}$.

□

This leads to yet another set of normalizing **S**-terms.

8.4.6. Proposition. $\downarrow\left(\mathcal{L}_1\mathcal{L}_0\right)$

Proof. Reduce $\mathcal{L}_1\mathcal{L}_0$, treating $\mathcal{L}_0$ as a variable. By proposition 8.4.5, the only new redexes have the form $\mathcal{L}_0(\mathsf{S}\mathcal{P}_2\mathcal{L}_0)$. We have $\mathcal{L}_0 \subseteq \mathcal{N}_1$, therefore $\mathsf{S}\mathcal{P}_2\mathcal{L}_0 \subseteq \mathsf{S}\mathcal{P}_2\mathcal{N}_1 \subseteq \mathcal{N}_1$ by definition. But them $\mathcal{L}_0(\mathsf{S}\mathcal{P}_2\mathcal{L}_0) \subseteq \mathcal{L}_0\mathcal{N}_1$ which normalizes by proposition 8.4.3.

□

Up to here we found some normalizing subsets of $\mathcal{N}_1\mathcal{N}_1$. Now we will show that all other elements of $\mathcal{N}_1\mathcal{N}_1$ have an infinite reduction.
For convenience, we take

8.4.7. Definition. $\mathcal{L}_{1,2} = \mathcal{N}_1 \setminus \mathcal{L}_0; \mathcal{L}_2 = \mathcal{L}_{1,2} \setminus \mathcal{L}_1$.

So $\mathcal{N}_1$ is the disjoint sum of $\mathcal{L}_0$, $\mathcal{L}_1$, and $\mathcal{L}_2$.
From $\mathcal{L}_{1,2}$ we build more non-normalizing terms.

8.4.8. Proposition. $\mathcal{L}_{1,2}\mathcal{L}_{1,2} \twoheadrightarrow \infty$

Proof. By manipulating rational languages we derive $\mathcal{L}_{1,2} \subseteq \left\langle \mathcal{Q}_2\mathcal{Q}_1 \right\rangle$. So proposition 8.2.3 is applicable.

□

Combining propositions 8.4.3, 8.4.6, and 8.4.8, there's exactly one case remaining in the analysis of $\mathcal{N}_1\mathcal{N}_1$, and that's $\mathcal{L}_2\mathcal{L}_0$. It is covered by the following

8.4.9. Proposition. $\mathcal{L}_2\mathcal{N}_1 \twoheadrightarrow \infty$

Proof. First we observe that

$$\mathcal{L}_2 \subseteq \left\langle \mathsf{T} \middle/ \mathsf{S}(\mathsf{S}\mathsf{T})\mathsf{T} \cup \mathsf{S}\mathsf{T}\mathcal{Q}_3 \right\rangle$$

(This can be verified, using `RX`.) Then we have two base cases:

1. $\mathsf{S}\mathsf{T}\mathcal{Q}_3\mathcal{N}_1 \to \mathsf{T}\mathcal{N}_1(\mathcal{Q}_3\mathcal{N}_1) \to \mathsf{S}(\mathcal{Q}_3\mathcal{N}_1)\boxed{\mathcal{N}_1(\mathcal{Q}_3\mathcal{N}_1)}$. By corollary 8.3.10, applied to the boxed expression, there is a reduction $\mathcal{N}_1(\mathcal{Q}_3\mathcal{N}_1) \hookrightarrow\!\!\!\twoheadrightarrow \mathcal{Q}_3\mathcal{N}_1X$ with $\mathcal{Q}_3\mathcal{N}_1 \lhd X$, so that $X \subseteq \mathcal{Q}_2$. Moreover we sure have $\mathcal{N}_1 \subseteq \mathcal{Q}_2$, therefore the $\mathcal{Q}_3\mathcal{N}_1X \subseteq \mathcal{Q}_3\mathcal{Q}_2\mathcal{Q}_1$ and proposition 8.2.4 can be applied.

2. $\mathsf{S}(\mathsf{S}\mathsf{T})\mathsf{T}\mathcal{N}_1 \to \mathsf{S}\mathsf{T}\mathcal{N}_1(\mathsf{T}\mathcal{N}_1) \subseteq \mathsf{S}\mathsf{T}\mathcal{Q}_3\mathcal{N}_1$ which gets us back to the previous case.

The inductive step is just $\mathsf{T}\mathcal{L}_2\mathcal{N}_1 \to \mathsf{S}\mathcal{N}_1\boxed{\mathcal{L}_2\mathcal{N}_1}$.

□

8.5 The sub-case $\mathcal{N}_2\,\mathcal{P}_2$

Combining all of the previous, the analysis for $\mathcal{N}_2{}^\wedge\mathcal{Q}_2$ is complete, but $\mathcal{N}_2{}^\wedge\mathcal{P}_2 = \mathcal{N}_2{}^\wedge\mathsf{S} \cup \mathsf{T}$ is missing. We will show that after doing some reductions, the decision procedure for $\mathcal{N}_2{}^\wedge\mathcal{Q}_2$ can be used.

We define a notion of *complete reduction*. After one complete reduction step of an $X{}^\wedge\mathsf{S}$ redex, with X in normal form, no further redexes $X{}^\wedge\mathsf{S}$ can occur. And after a complete reduction step of an $X{}^\wedge\mathsf{T}$ redex, no further redexes $X{}^\wedge\mathcal{P}_2$ are possible. That's why we then may use the procedure that decides normalization for $\mathcal{N}{}^\wedge\mathcal{Q}_2$.

8.5.1. Definition. The complete reduction relation $X \stackrel{!}{\to} W$ is defined for all $X \in \mathcal{N}{}^\wedge\mathcal{M}$ and $W \in \mathcal{M}$ by

$$\mathsf{S}{}^\wedge Z \stackrel{!}{\to} \mathsf{S}{}^\wedge Z;\quad (\mathsf{S}{}^\wedge X){}^\wedge Z \stackrel{!}{\to} (\mathsf{S}{}^\wedge X){}^\wedge Z$$

$$((\mathsf{S}{}^\wedge X){}^\wedge Y){}^\wedge Z \stackrel{!}{\to} X'{}^\wedge Y' \quad \text{where} \quad X{}^\wedge Z \stackrel{!}{\to} X' \text{ and } Y{}^\wedge Z \stackrel{!}{\to} Y'$$

8.5.2. Corollary. Obviously $X \stackrel{!}{\to} W$ implies $X \twoheadrightarrow W$. For each $X \in \mathcal{N}{}^\wedge\mathcal{M}$, there is exactly one W with $X \stackrel{!}{\to} W$, so we can view this relation as a function, defined by structural recursion on the normal forms.

In a redex $((\mathsf{S}{}^\wedge X){}^\wedge Y){}^\wedge Z$, call Z the *argument* of the redex. We are also interested in subterms that might become arguments later. In a term, call a subterm (position) an *argument candidate* if it occurs as the right child of a subterm that has a left spine longer than two (i. e. whose left child is in $\mathcal{M}_2$). The purpose of $\stackrel{!}{\to}$ is that in $X{}^\wedge Y \stackrel{!}{\to} Z$, all argument candidates in terms reachable from Z are larger than Y. This is made precise by

8.5.3. Proposition. If $X \in \mathcal{N}$, and $X{}^\wedge\mathsf{S}_k \stackrel{!}{\to} Z_0 \to Z_1 \to \ldots$, then none of the Z_i contains a subterm $\mathcal{M}_2{}^\wedge\mathcal{P}_k$.

Proof. We prove that the property holds for Z_0. Then we proceed by induction.

We are going to show that $X{}^\wedge\mathsf{S}_k \stackrel{!}{\to} Z_0$ implies that no $\mathcal{M}_2{}^\wedge\mathcal{P}_k \trianglelefteq Z_0$. We use structural induction on X.

1. $X = \mathsf{S} \cup \mathsf{S}{}^\wedge X'$. Then $X{}^\wedge\mathsf{S}_k \stackrel{!}{\to} X{}^\wedge\mathsf{S}_k$ and does not contain $\mathcal{M}_2{}^\wedge\mathcal{P}_k$ because it's a normal form and $\mathcal{M}_2{}^\wedge\mathcal{P}_k$ is not.

2. $X = (\mathsf{S}{}^\wedge X_1){}^\wedge X_2$. Then $X{}^\wedge\mathsf{S}_k \stackrel{!}{\to} X_1'{}^\wedge X_2'$ where $X_1{}^\wedge\mathsf{S}_k \stackrel{!}{\to} X_1'$ and $X_2{}^\wedge\mathsf{S}_k \stackrel{!}{\to} X_2'$. By induction, the claim is true for X_1' and X_2'. We have $\mathsf{S}_k \lhd X_2'$, therefore $X_2' \in \mathcal{Q}_k$ and $X_1'{}^\wedge X_2' \notin \mathcal{M}_2{}^\wedge\mathcal{P}_k$.

Now chose the smallest i such that Z_i contains $\mathcal{M}_2{}^\wedge\mathcal{P}_k$. (We have $i > 0$.) This subexpression must have been created in the reduction $Z_{i-1} \to Z_i$. We list the possible relative positions of the redex $((\mathsf{S}{}^\wedge C){}^\wedge D){}^\wedge E$ (resp. its contractum $(C{}^\wedge E){}^\wedge(D{}^\wedge E)$) and the pattern $\mathcal{M}_2{}^\wedge\mathcal{P}_k$.

1. The pattern occurs inside C, D, or E. Then it would already have occured in Z_{i-1}, so i is not minimal.

2. $C^\wedge E \in \mathcal{M}_2{}^\wedge\mathcal{P}_k$ or $D^\wedge E \in \mathcal{M}_2{}^\wedge\mathcal{P}_k$ Then $E \in \mathcal{P}_k$ and in Z_{i-1} the redex $((\mathbf{S}^\wedge C)^\wedge D)^\wedge E \in ((\mathbf{S}^\wedge *)^\wedge *)^\wedge\mathcal{P}_k \subseteq \mathcal{M}_2{}^\wedge\mathcal{P}_k$ and i is not minimal.

3. $(C^\wedge E)^\wedge(D^\wedge E) \in \mathcal{M}_2{}^\wedge\mathcal{P}_k$. In Z_{i-1}, we had $((\mathbf{S}^\wedge C)^\wedge D)^\wedge E \in ((\mathbf{S}^\wedge *)^\wedge *)^\wedge E \subseteq \mathcal{M}_2{}^\wedge E$. Because i is minimal, we have $E \in \mathcal{Q}_k$ leading to $D^\wedge E \in \mathcal{Q}_k$ and the contradiction $(C^\wedge E)^\wedge(D^\wedge E) \in \mathcal{M}_2{}^\wedge\mathcal{Q}_k$.

4. $(\cdots((C^\wedge E)^\wedge(D^\wedge E))\cdots{}^\wedge *)^\wedge G \in \mathcal{M}_2{}^\wedge\mathcal{P}_k$. That is, $G \in \mathcal{P}_k$. But then $(\cdots(((\mathbf{S}^\wedge C)^\wedge D)^\wedge E)\cdots{}^\wedge *)^\wedge G \in \mathcal{M}_2{}^\wedge\mathcal{P}_k$ in Z_{i-1} as well, contradicting the minimality of i.

□

8.6 Deciding Normalization

8.6.1. Theorem. There is a procedure that decides whether an expression in $\mathbf{CL}(\mathbf{S})$ has a normal form.

Proof. $\mathbf{S}$ has a normal form. Therefore assume the expression is $X^\wedge Y$.
Recursively apply the procedure to X and Y and compute their normal forms if they exist. This recursion is terminating because it's structural. If one (or both) of the normal forms don't exist, halt and answer *not normalizing*. Otherwise call them X' and Y', respectively.
If $Y' \in \mathcal{Q}_2$, answer *normalizing* if $X'^\wedge Y' \in \mathcal{N}_0{}^\wedge\mathcal{N} \cup \mathcal{L}_0{}^\wedge\mathcal{N}_1 \cup \mathcal{L}_1{}^\wedge\mathcal{L}_0 \cup \mathcal{N}_1{}^\wedge\mathcal{N}_0$, and *not normalizing* else.
If $Y' \in \mathcal{P}_2$, then case T or $\mathbf{S}$ applies:
Case T: If $Y' = \mathsf{T}$, determine Z from $X'^\wedge Y' \xrightarrow{!} Z$. Then apply the procedure to Z. By proposition 8.5.3, no redex $\mathcal{M}_2{}^\wedge\mathsf{T}$ will occur in reductions from Z. Therefore case T will not happen again, so recursion terminates.
Case $\mathbf{S}$: If $Y' = \mathbf{S}$, determine Z from $X'^\wedge Y' \xrightarrow{!} Z$. Then apply the procedure to Z. By proposition 8.5.3, no redex $\mathcal{M}_2{}^\wedge\mathbf{S}$ will occur in reductions from Z. Therefore case $\mathbf{S}$ will not happen again, so recursion terminates. □

8.6.2. Remark. It is possible to omit cases $\mathbf{S}$ and T from the decision procedure, by explicitly giving rational sets X_S and X_T such that $X^\wedge\mathbf{S}$ (for $X \in \mathcal{N}$) normalizes iff $X \in X_S$, and similar for T. The exact construction has been omitted here because it is quite technical.

In the procedure given above, of course one would like to remove the restriction that X and Y be already in normal form, to save all those recursive calls. Indeed this will be done in chapter 9 where we give an explicit construction for the set of all normalizing $\mathbf{S}$-terms.

8.7 Related work

The Starling

The combinator **S** is called the *starling* in Smullyan's book [Smu85].

The decidability of normalization of **S**-terms had already been conjectured by E. Zachos in 1978.

Zachos completely analyzes the normalization of **S**-terms of size up to 9, by either finding the normal form or exhibiting patterns (similar to those given in the examples 5.5.1 or 5.5.2) that lead to an infinite reduction.

Still a proof that each non-normalizing term really admits such a pattern would be desirable.

The Lark

The combinator **L** with reduction rule $\mathbf{L}xy \to x(yy)$ is called the *lark* by Smullyan. It has been investigated in depth by R. Statman [Sta89]. He showed that the word problem is decidable for $\mathbf{CL}(\mathbf{L})$, using a very technical argument.

Later, Sprenger and Wymann-Böni [SWB93] gave a different proof. Essentially, they considered the system $x(yy) \to \mathbf{L}xy$ which is terminating, but not confluent. Then they try to make the system confluent. But completion does not stop. So they make an ingenious guess on the resulting infinite TRS, which can later be verified.

We conjecture that the word problem should also be decidable for the starling but it might be substantially more difficult than for the lark: If deciding normalization is a step towards deciding convertibility, than this first step is trivial for the lark, but, as we've seen, quite hard for the starling.

Chapter 9

Normalizing terms in $\mathbf{CL}(\mathbf{S})$

9.1 Introduction

This chapter presents the rational language of normalizing terms in $\mathbf{CL}(\mathbf{S})$.

This amounts to finding pred $(\mathcal{N})$, the set of (many-step) predecessors of $\mathcal{N}$. In general, this need not be rational set, (see section 4.6,) and even if it were, there is no method known for finding it.

So what we're really doing here is to present a rational set of normalizing terms which is saturated (forward and backward closed under reduction) and which contains $\mathcal{N}$.

The construction is based on the subsets of $\mathcal{N}$ that have already been investigated in chapter 8. It has been found by doing some "educated guesses" and later verifying them by the program `RX`.

While this approach solves the problem for this case, a general method for constructing (backward) closures of rational languages (here: normal forms) would be desirable, so that we could avoid the guesswork.

It has to be admitted that the verification of the given grammar was only partial, as will be explained. Therefore we adopt the view that the main result of the thesis is *decidability* (chapter 8, which does not depend on machine proofs) rather than *rationality* (this chapter) of pred $(\mathcal{N})$.

9.2 A first look at pred$(\mathcal{N})$

For this section, we agree on the following notation.

9.2.1. Definition. $\mathrm{pred}_Y(Z) = \{X \mid X\,Y \twoheadrightarrow Z\}$; $\mathrm{pred}(Z) = \{X \mid X \twoheadrightarrow Z\}$

That is, $\mathrm{pred}_Y(Z)$ are all terms X that reduce to Z when supplied with an argument from Y. Earlier, $\mathrm{pred}(Z)$ had been called $\Delta^{\twoheadleftarrow}(Z)$.
By theorem 8.6.1, we have

9.2.2. Proposition.

$$\begin{aligned}\mathrm{pred}(\mathcal{N}) = \mathbf{S} &\cup \mathrm{pred}(\mathcal{N}_0)\,\mathrm{pred}(\mathcal{N}) \cup \mathrm{pred}(\mathcal{L}_0)\,\mathrm{pred}(\mathcal{N}_{0,1})\\ &\cup \mathrm{pred}(\mathcal{L}_1)\,\mathrm{pred}(\mathcal{L}_0) \cup \mathrm{pred}(\mathcal{N}_{0,1})\,\mathrm{pred}(\mathcal{N}_0)\\ &\cup \mathrm{pred}_{\mathbf{S}}(\mathcal{N})\,\mathbf{S} \cup \mathrm{pred}_{\mathsf{T}}(\mathcal{N})\,\mathsf{T}\end{aligned}$$

Indeed this can be completed to a regular grammar for $\mathrm{pred}(\mathcal{N})$.

9.2.3. Proposition.

$$\begin{aligned}\mathrm{pred}(\mathcal{N}_0) = \mathbf{S} &\cup \mathbf{S}\,\mathrm{pred}(\mathcal{N}) \cup \mathsf{T}\,\mathrm{pred}(\mathcal{N}_0)\\ &\cup \mathrm{pred}_{\mathbf{S}}(\mathcal{N}_0)\,\mathbf{S} \cup \mathrm{pred}_{\mathsf{T}}(\mathcal{N}_0)\,\mathsf{T}\end{aligned}$$

Proof. The expression matches the structure of $\mathcal{N}_0 = \mathbf{S} \cup \mathbf{S}\,\mathcal{N} \cup \mathsf{T}\,\mathcal{N}_0$. $\square$

Similarly,

9.2.4. Proposition.

$$\begin{aligned}\mathrm{pred}(\mathcal{N}_{0,1}) = \mathbf{S} &\cup \mathbf{S}\,\mathrm{pred}(\mathcal{N}) \cup (\mathbf{S}\,(\mathbf{S}\,\mathsf{T}))\,\mathcal{P}_2 \cup (\mathbf{S}\,\mathcal{P}_2)\,\mathrm{pred}(\mathcal{N}_{0,1})\\ &\cup \mathrm{pred}_{\mathbf{S}}(\mathcal{N}_{0,1})\,\mathbf{S} \cup \mathrm{pred}_{\mathsf{T}}(\mathcal{N}_{0,1})\,\mathsf{T}\end{aligned}$$

Proof. The expression matches the structure of $\mathcal{N}_{0,1} = \mathbf{S} \cup \mathbf{S}\,\mathcal{N} \cup (\mathbf{S}\,(\mathbf{S}\,\mathsf{T}))\,\mathcal{P}_2 \cup (\mathbf{S}\,\mathcal{P}_2)\,\mathcal{N}_{0,1}$. $\square$

9.2.5. Proposition.

$$\begin{aligned}\mathrm{pred}_{\mathbf{S}}(\mathcal{N}) = \mathbf{S} &\cup \mathbf{S}\,\mathrm{pred}(\mathcal{N})\\ &\cup \mathrm{pred}\left((\mathbf{S}\,\mathrm{pred}_{\mathbf{S}}(\mathcal{N}_0))\,\mathrm{pred}_{\mathbf{S}}(\mathcal{N})\right) \cup \mathrm{pred}\left((\mathbf{S}\,\mathrm{pred}_{\mathbf{S}}(\mathcal{L}_0))\,\mathrm{pred}_{\mathbf{S}}(\mathcal{N}_{0,1})\right)\\ &\cup \mathrm{pred}\left((\mathbf{S}\,\mathrm{pred}_{\mathbf{S}}(\mathcal{L}_1))\,\mathrm{pred}_{\mathbf{S}}(\mathcal{L}_0)\right) \cup \mathrm{pred}\left((\mathbf{S}\,\mathrm{pred}_{\mathbf{S}}(\mathcal{N}_{0,1}))\,\mathrm{pred}_{\mathbf{S}}(\mathcal{N}_0)\right)\\ &\cup \mathrm{pred}_{\mathbf{S}}(\mathrm{pred}_{\mathbf{S}}(\mathcal{N}))\,\mathbf{S} \cup \mathrm{pred}_{\mathsf{T}}(\mathrm{pred}_{\mathbf{S}}(\mathcal{N}))\,\mathsf{T}\end{aligned}$$

Proof. Take one $X \in \mathrm{pred}_{\mathbf{S}}(\mathcal{N})$. In the reduction $X\,\mathbf{S} \twoheadrightarrow \mathcal{N}$, consider the first step where the redex is at the root of the term. It has the form $((\mathbf{S}\,Y)\,Z)\,\mathbf{S}$, and its contractum is $(Y\,\mathbf{S})\,(Z\,\mathbf{S})$.
This must be an element of $\mathrm{pred}(\mathcal{N})$, so we can apply the classification of proposition 9.2.2. $\square$

Similarly,

9.2.6. Proposition.

$$\begin{aligned}\mathrm{pred}_{\mathsf{T}}(\mathcal{N}) = {} & \mathsf{S} \cup \mathsf{S}\ \mathrm{pred}(\mathcal{N}) \\ & \cup \mathrm{pred}\left(\mathsf{S}\ \mathrm{pred}_{\mathsf{T}}(\mathcal{N}_0)\ \mathrm{pred}_{\mathsf{T}}(\mathcal{N})\right) \cup \mathrm{pred}\left(\mathsf{S}\ \mathrm{pred}_{\mathsf{T}}(\mathcal{L}_0)\ \mathrm{pred}_{\mathsf{T}}(\mathcal{N}_{0,1})\right) \\ & \cup \mathrm{pred}\left(\mathsf{S}\ \mathrm{pred}_{\mathsf{T}}(\mathcal{L}_1)\ \mathrm{pred}_{\mathsf{T}}(\mathcal{L}_0)\right) \cup \mathrm{pred}\left(\mathsf{S}\ \mathrm{pred}_{\mathsf{T}}(\mathcal{N}_{0,1})\ \mathrm{pred}_{\mathsf{T}}(\mathcal{N}_0)\right) \\ & \cup \mathrm{pred}_{\mathsf{S}}(\mathrm{pred}_{\mathsf{T}}(\mathcal{N}))\ \mathsf{S} \cup \mathrm{pred}_{\mathsf{T}}(\mathrm{pred}_{\mathsf{T}}(\mathcal{N}))\ \mathsf{T}\end{aligned}$$

9.3 The case $\mathrm{pred}_{\mathsf{T}}(\mathcal{N})$

We immediately see

9.3.1. Proposition. $\mathrm{pred}_{\mathsf{T}}(\mathcal{L}_0) = \emptyset$

Therefore the expression reduces to

$$\begin{aligned}\mathrm{pred}_{\mathsf{T}}(\mathcal{N}) = {} & \mathsf{S} \cup \mathsf{S}\ \mathrm{pred}(\mathcal{N}) \\ & \cup \mathrm{pred}\left(\mathsf{S}\ \mathrm{pred}_{\mathsf{T}}(\mathcal{N}_0)\ \mathrm{pred}_{\mathsf{T}}(\mathcal{N})\right) \cup \mathrm{pred}\left(\mathsf{S}\ \mathrm{pred}_{\mathsf{T}}(\mathcal{N}_{0,1})\ \mathrm{pred}_{\mathsf{T}}(\mathcal{N}_0)\right) \\ & \cup \mathrm{pred}_{\mathsf{S}}(\mathrm{pred}_{\mathsf{T}}(\mathcal{N}))\ \mathsf{S} \cup \mathrm{pred}_{\mathsf{T}}(\mathrm{pred}_{\mathsf{T}}(\mathcal{N}))\ \mathsf{T}\end{aligned}$$

We now give the following explicit constructions. Their correctness can be shown by using the program `RX`, which follows the algorithm for checking saturation indicated in section 4.6.

9.3.2. Proposition.

$$\begin{aligned}\mathrm{pred}_{\mathsf{S}}(\mathcal{L}_0) &= \left\langle \mathsf{T} \middle/ \mathsf{S}\ \mathsf{S}\ \mathcal{P}_2 \right\rangle \\ \mathrm{pred}(\mathcal{L}_0) &= \left\langle \mathsf{T} \middle/ \mathrm{pred}_{\mathsf{S}}(\mathcal{L}_0)\ \mathsf{S} \right\rangle\end{aligned}$$

9.3.3. Proposition.

$$\begin{aligned}\mathrm{pred}_{\mathsf{T}}(\mathcal{L}_1) &= \mathsf{S}\ \mathsf{T} \cup \mathsf{A} \\ \mathrm{pred}_{\mathsf{S}}(\mathcal{L}_1) &= \left\langle \mathsf{T} \middle/ \left\langle \mathsf{T} \middle/ \mathsf{S}\ \mathsf{T} \right\rangle\ \mathsf{S} \right\rangle \\ \mathrm{pred}(\mathcal{L}_1) &= \left\langle \mathsf{T} \middle/ \mathsf{S}\ \mathsf{T}\ \mathsf{S}\ \mathsf{T} \cup \mathrm{pred}_{\mathsf{T}}(\mathcal{L}_1)\ \mathsf{T} \cup \mathrm{pred}_{\mathsf{S}}(\mathcal{L}_1)\ \mathsf{S} \right\rangle\end{aligned}$$

9.3.4. Proposition.

$$\mathrm{pred}_{\mathsf{T}}(\mathcal{N}_0) = \mathcal{P}_2; \mathrm{pred}_{\mathsf{S}}(\mathcal{N}_0) = \mathcal{N}_{00}$$

9.3.5. Proposition.

$$\mathrm{pred}_{\mathsf{T}}(\mathcal{N}_{0,1}) = \mathcal{N}_{00} \;\cup\; \mathrm{pred}_{\mathsf{S}}(\mathcal{L}_0)$$
$$\mathrm{pred}_{\mathsf{S}}(\mathcal{N}_{0,1}) = \mathrm{pred}_{\mathsf{T}}(\mathcal{N}_{0,1}) \;\cup\; \left\langle \mathsf{T} \middle/ \left(\left\langle \mathsf{T} \middle/ (\mathsf{S}\,\mathcal{N}_{00})\right\rangle\,\mathsf{S}\right)\right\rangle$$

From proposition 9.2.6 we obtain

9.3.6. Proposition. $\mathrm{pred}_{\mathsf{T}}(\mathcal{N}) \supseteq \mathrm{pred}\left(\left(\mathsf{S}\,\mathrm{pred}_{\mathsf{T}}(\mathcal{N}_0)\right)\,\mathrm{pred}_{\mathsf{T}}(\mathcal{N})\right) \supseteq (\mathsf{S}\,\mathcal{P}_2)\,\mathrm{pred}_{\mathsf{T}}(\mathcal{N})$

In fact we have

9.3.7. Proposition.

$$\mathrm{pred}_{\mathsf{T}}(\mathcal{N}) = \mathsf{S} \;\cup\; (\mathsf{S}\,\mathrm{pred}(\mathcal{N})) \;\cup\; (\mathrm{pred}_{\mathsf{T}}(\mathcal{N}_{0,1})\,\mathcal{P}_2) \;\cup\; ((\mathsf{S}\,\mathcal{P}_2)\,\mathrm{pred}_{\mathsf{T}}(\mathcal{N}))$$
$$\cup\; (\mathrm{pred}_{\mathsf{S}}(\mathrm{pred}_{\mathsf{T}}(\mathcal{N}))\,\mathsf{S}) \;\cup\; (\mathrm{pred}_{\mathsf{T}}(\mathrm{pred}_{\mathsf{T}}(\mathcal{N}))\,\mathsf{T})$$

As had been observed in proving theorem 8.6.1, the difficult cases are $(X\,\mathsf{S})$ and $(X\,\mathsf{T})$, or rather $\mathrm{pred}_{\mathsf{S}}(\mathcal{N})$ and $\mathrm{pred}_{\mathsf{T}}(\mathcal{N})$. That's why we need to evaluate stacks of $\mathrm{pred}_{\mathcal{P}_2}(\cdot)$.

9.3.8. Proposition.

$$\mathrm{pred}_{\mathsf{S}}(\mathrm{pred}_{\mathsf{T}}(\mathcal{N})) = \mathrm{pred}_{\mathsf{S}}(\mathcal{N}_{0,1}) \;\cup\; \left\langle \mathsf{T} \middle/ (\mathsf{S}\,\mathrm{pred}_{\mathsf{T}}(\mathcal{N}_{0,1}))\right\rangle$$
$$\mathrm{pred}_{\mathsf{T}}(\mathrm{pred}_{\mathsf{T}}(\mathcal{N})) = \left\langle \mathsf{T} \middle/ (\mathsf{S}\,\mathrm{pred}_{\mathsf{T}}(\mathcal{N}_{0,1}))\right\rangle$$
$$\mathrm{pred}_{\mathsf{T}}(\mathrm{pred}_{\mathsf{S}}(\mathcal{N})) = \left\langle \mathsf{T} \middle/ ((\mathsf{S}\,(\mathsf{S}\,\mathsf{T}))\,\mathsf{T})\right\rangle \;\cup\; \mathrm{pred}_{\mathsf{S}}(\mathcal{L}_1)$$

That completes the expression for $\mathrm{pred}_{\mathsf{T}}(\mathcal{N})$, but we're still missing $\mathrm{pred}_{\mathsf{S}}(\mathcal{N})$.

9.4 The case $\mathrm{pred}_{\mathsf{S}}(\mathcal{N})$

Intricate expressions are necessary for stacks of $\mathrm{pred}_{\mathsf{S}}(\cdot)$, so they'll deserve names in their own right.

9.4.1. Definition. *one* $= \mathrm{pred}_{\mathsf{S}}(\mathcal{N})$, *two* $= \mathrm{pred}_{\mathsf{S}}(\mathrm{pred}_{\mathsf{S}}(\mathcal{N}))$, *three* $= \mathrm{pred}_{\mathsf{S}}(\mathrm{pred}_{\mathsf{S}}(\mathrm{pred}_{\mathsf{S}}(\mathcal{N})))$, *four* $= \mathrm{pred}_{\mathsf{S}}(\mathrm{pred}_{\mathsf{S}}(\mathrm{pred}_{\mathsf{S}}(\mathrm{pred}_{\mathsf{S}}(\mathcal{N}))))$

Fortunately, we don't need larger numbers, as *five* = *six* = ... $= \mathcal{N}_{00}$.
From proposition 9.2.5 we derive

9.4.2. Proposition. *one* $= \mathrm{pred}_{\mathsf{S}}(\mathcal{N}) \supseteq \mathrm{pred}\left(\left(\mathsf{S}\,\mathrm{pred}_{\mathsf{S}}(\mathcal{N}_0)\right)\,\mathrm{pred}_{\mathsf{S}}(\mathcal{N})\right) \supseteq (\mathsf{S}\,\mathcal{N}_{00})\,\mathit{one}$

and we easily verify

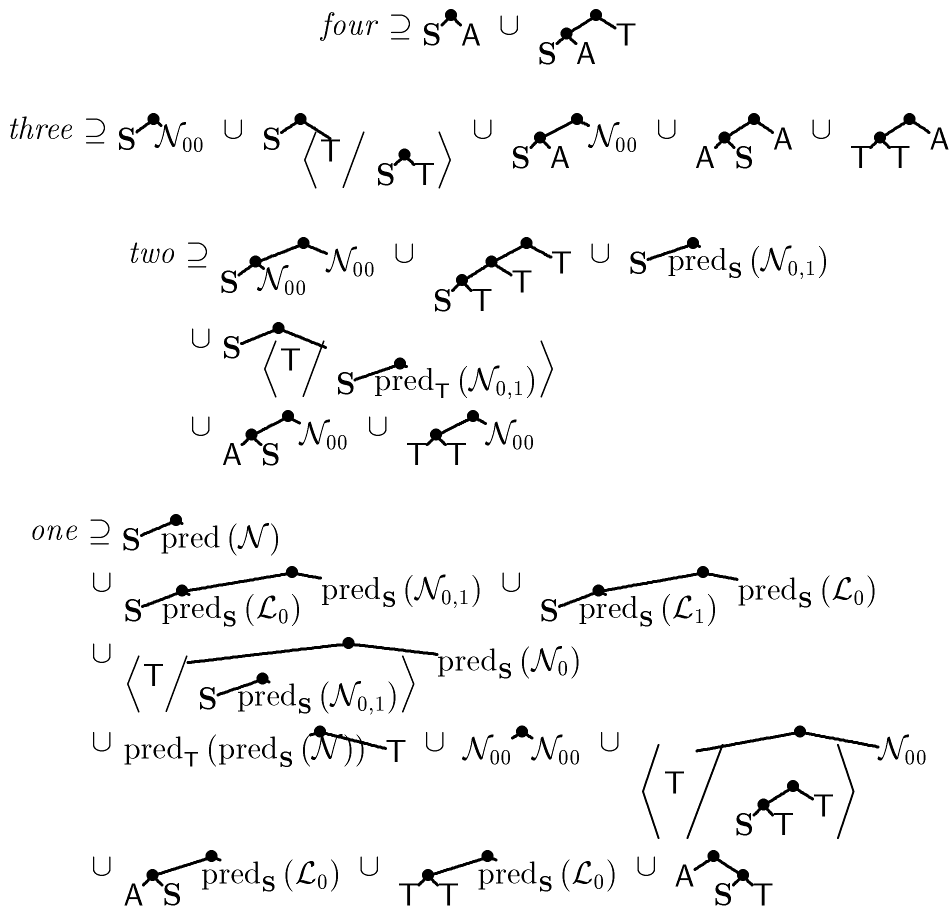
9.4.3. Proposition. $\mathit{two} \supseteq \mathsf{T}\,\mathit{two}, \mathit{three} \supseteq \mathsf{T}\,\mathit{three}, \mathit{four} \supseteq \mathsf{T}\,\mathit{four}$

9.4.4. Proposition. $\mathit{one} \supseteq \mathit{two}\,\mathsf{S}$, $\mathit{two} \supseteq \mathit{three}\,\mathsf{S}$, $\mathit{three} \supseteq \mathit{four}\,\mathsf{S}$

9.4.5. Proposition. $\mathit{four} \supseteq \mathcal{N}_{00}$, $\mathit{three} \supseteq \mathcal{N}_{00}$, $\mathit{two} \supseteq \mathcal{N}_{00}$, $\mathit{one} \supseteq \mathcal{N}_{00}$

The propositions 9.4.2, 9.4.3, 9.4.4, 9.4.5 are interpreted as rules in the grammars for *one*, *two*, *three*, *four*. They have to be completed by the following

9.4.6. Proposition.



That completes the regular grammar that generates pred $(\mathcal{N})$.

All of the constructions can be checked with `RX`.

9.5 Conclusion

To be fair, it has to be admitted that the a complete check has not been achieved so far. The grammar is rather large, and pushes `RX`'s limits.

However, extensive saturation checking has been done for reasonable sub-grammars. Further, it can be verified by pen and paper that all **S**-terms generated by the full grammar really are normalizing. The question is of course whether we forgot to produce some normalizing terms. Again extensive testing of example terms shows that this doesn't seem to be the case. As a conclusion, what really had been shown here is that pred $(\mathcal{N})$ *strongly* seems to be rational, and its grammar looks more or less like the one presented above, and it could be obtained and verified by using some more accuracy and/or time in machine or human computation.