Normalizing S-Terms can be Generated by a Context-free Grammar

Stathis Zachos[1,2,3], John Ramirez[3], and Panos Hilaris[1]

[1] Computer Science, ECE, National Technical University of Athens
[2] CIS, Brooklyn College, CUNY
[3] Computer Science, GradCenter, CUNY
Email: {zachos, philaris}@cs.ntua.gr

Abstract. Curry's Combinatory Logic is a functional calculus which may serve as a foundation to the theory of computations, even to computational complexity. Combinatory Logic, which is based on the two combinators S and K, is an undecidable theory. The theory based only on S has been proven decidable by J. Waldmann. Zachos simplified the proof and gave a decision algorithm. Here we make a further step in lowering the complexity of the decision algorithm. We present a context-free grammar which fully characterizes all normalizing S-terms. Thus the complexity of deciding whether an S-term X has a normal form is $O(|X|^3)$ given by the CYK algorithm.

1 Related Work and Motivation

"Given an S-term, is it normalizable?" The question was answered positively [15]. In [16, 17], a simpler proof without the use of rational tree languages and an actual decision algorithm were presented. Many people have been involved in similar investigations, for example: Barendregt, Bergstra, Klop, Statman, Dershowitz, Jouannaud, Smullyan etc. [18, 1–3, 7–9, 11, 13, 14, 12].

The original motivation of this problem was the need to create a *Functional Calculus* instead of *Set Theory* as a foundation for Theory of Computation, i.e. for Computability (Thue, Schönfinkel [10], Curry [5, 6], Church [4], Turing, Markov) but even for Computational Complexity. In such a functional calculus only *one operation* is needed: application $f(g)$. We write (fg) instead of $f(g)$.

Schönfinkel made the following observation: functions of one argument are enough, e.g., $f(g, h) = ((fg)h)$. We use left association for dropping some parentheses, i.e., instead of $((((fg)(h(gh)))((gh)((fh)f))))$ we write $fg(h(gh))(gh(fhf))$.

An example is the SK-Calculus or Combinatory Logic of H. Curry [5]. It postulates combinators and rewriting rules, e.g. $St_1t_2t_3 = t_1t_3(t_2t_3)$ and $Kt_1t_2 = t_1$. Actually there are many other combinators and rewriting rules but the system $\{S, K\}$ is complete.

2 Introduction

We call S-terms the elements of a system generated by one symbol S and one non-associative and non-commutative (implicit) operation that we call application. We construct S-terms as strings of S 's and parentheses with the following rules:

- S is an S-term.
- If m_1 and m_2 are S-terms, $(m_1 m_2)$ is an S-term.

In this construction we say that m_1 and m_2 are *proper sub-terms* of the constructed S-term. S is a sub-term of any S-term. Also, we say that an S-term is a sub-term of itself. We may abbreviate by omitting parentheses by using left association. For example, we write $SS(S(SSS))S$ instead of $(((SS)(S((SS)S)))S)$ and $xyz = (xy)z \neq x(yz)$.

Here, we use lower case italic letters to represent S-terms. We use upper case calligraphic letters to represent sets of S-terms. For any sets of S-terms $\mathcal{A}$ and $\mathcal{C}$, we will write $\mathcal{AC} = \{\, ac \mid a \in \mathcal{A} \text{ and } c \in \mathcal{C}\,\}$.

We define the length of an S-term to be the number of occurrences of the symbol S in the term. For any S-term x, we write $|x|$ to denote the length of x.

The reduction relation $\rightarrow$ is defined here by the S-rule:

$$Sadc \stackrel{\text{def}}{\longrightarrow} ac(dc)\,.$$

The left hand side, $Sadc$, is sometimes called *redex* and the right hand side, $ac(dc)$, *reductum*. In particular,

$$SSdc \longrightarrow Sc(dc)\,.$$

In general we write $x \rightarrow y$ if y can be written by replacing some redex, sub-term, in x by the corresponding reductum of the S-rule.

When reducing by the S-rule we eliminate one symbol S from an S-term and introduce a replica of a sub-term in the S-term, hence, if $x \rightarrow y$, $|x| \leq |y|$. Thus a reduction step certainly does not reduce the length of the S-term.

Here, we will use an abbreviation $B \stackrel{\text{def}}{=} S(SS)$. Using the S-rule twice we get:

$$Bad \;=\; S(SS)ad \;\longrightarrow\; SSd(ad) \;\longrightarrow\; S(ad)(d(ad))\,.$$

that we will write:

$$Bad \stackrel{2}{\longrightarrow} S(ad)(d(ad))\,.$$

In general, for $k \geq 0$, we write $\stackrel{k}{\longrightarrow}$ to represent k reduction steps.

Here we describe other extensions of the relation $\rightarrow$. The transitive closure of $\rightarrow$ is denoted by $\stackrel{+}{\longrightarrow}$ and its reflexive transitive closure is denoted by $\stackrel{*}{\longrightarrow}$. For two sets of S-terms $\mathcal{X}$ and $\mathcal{Y}$, we will write $\mathcal{X} \rightarrow \mathcal{Y}$ if for any $x \in \mathcal{X}$ we can apply the S-rule on some redex sub-term of x so that $x \rightarrow y$ for some $y \in \mathcal{Y}$. Similarly, we will extend the other relations described above to sets of S-terms.

We say that an S-term x *is in normal form* if the S-rule cannot be applied to any sub-term of x, i.e., there is no redex in x. We say that x *has a normal form* and write $x\!\downarrow$ if $x \stackrel{*}{\longrightarrow} n$ for some n in normal form; we write $x\!\uparrow$ otherwise, i.e., if x *does not have a normal form*, which is equivalent to: there is a non-terminating reduction chain starting with x. The following was proven by Waldmann [15] and improved by Zachos [16].

Theorem. *There is an algorithm that decides if a given a S-term has a normal form and in that case produces the corresponding unique normal form.*

3 Notations

We first introduce some further notation and state some necessary technical facts.

Suppose $x \to y$. Then, for any sub-term z of y we will write $x \xrightarrow{\circ} z$. For example:

$$Sadc \xrightarrow{\circ} dc\,.$$

As extensions of $\xrightarrow{\circ}$, we will denote its transitive closure by $\xrightarrow{\oplus}$ and its reflexive transitive closure by $\xrightarrow{\circledast}$. Using this notation we have the following fact:

> Suppose $\mathcal{X} \xrightarrow{\oplus} \mathcal{X}$. Then, there is an infinite reduction chain starting with any $x \in \mathcal{X}$, i.e., $\mathcal{X}\uparrow$.

We are using a notation similar to regular expressions, e.g., we write S instead of $\{S\}$, we write $\mathcal{X}+\mathcal{Y}$ instead of $\mathcal{X}\cup\mathcal{Y}$, etc. $\mathcal{M}$ is the set of all S-terms; $\mathcal{N}$ is the set of all S-terms that are in normal form.

$$\mathcal{M} \stackrel{\text{def}}{=} S + \mathcal{M}\mathcal{M}, \qquad \mathcal{N} \stackrel{\text{def}}{=} S + S\mathcal{N} + S\mathcal{N}\mathcal{N}$$

For any set $\mathcal{A}$ we define $\overline{\mathcal{A}} = \mathcal{M} - \mathcal{A}$.

With this notation, we will also define the sets:

$$\mathcal{Q}_1 \stackrel{\text{def}}{=} \overline{S}, \qquad \mathcal{Q}_2 \stackrel{\text{def}}{=} \overline{S+SS}, \qquad \mathcal{Q}_3 \stackrel{\text{def}}{=} \overline{S+SS+S(SS)} = \overline{S+SS+B}$$

So $\mathcal{Q}_1$ is the set of all S-terms of length greater than one; $\mathcal{Q}_2$ is the set of all S-terms of length greater than two. Some immediate facts are:

$$\mathcal{Q}_1 = SS + \mathcal{Q}_2, \qquad \mathcal{M}\mathcal{Q}_i \subseteq \mathcal{Q}_{i+1} \subseteq \mathcal{Q}_i \quad \text{for } i \in \{1,2\}, \qquad \mathcal{M}\mathcal{Q}_3 \subseteq \mathcal{Q}_3$$

Since every reductum is in $\mathcal{M}\mathcal{M}\mathcal{M} \subseteq \mathcal{Q}_3 \subseteq \mathcal{Q}_2 \subseteq \mathcal{Q}_1$, we can always write $x \to \mathcal{Q}_i$ for any redex x (or any term x that has a redex!) and $i = 1, 2$, and 3.

For sets of S-terms $\mathcal{X}$ (the *prefix* set) and $\mathcal{Y}$ (the *base* set) we recursively define $\mathcal{X}^n[\mathcal{Y}]$ for all $n \geq 0$ by: $\mathcal{X}^0[\mathcal{Y}] = \mathcal{Y}$ and $\mathcal{X}^{k+1}[\mathcal{Y}] = \mathcal{X}(\mathcal{X}^k[\mathcal{Y}])$, for $k \geq 0$. The set of all terms defined above is: $\mathcal{X}^*[\mathcal{Y}] = \sum_{n\geq 0} \mathcal{X}^n[\mathcal{Y}] = \mathcal{X}^0[\mathcal{Y}] + \mathcal{X}^1[\mathcal{Y}] + \mathcal{X}^2[\mathcal{Y}] + \cdots$, which is the (least) solution of the fixpoint equation: $\mathcal{X}^*[\mathcal{Y}] = \mathcal{Y} + \mathcal{X}(\mathcal{X}^*[\mathcal{Y}])$. For example:

$$(SS+B)^*[\mathcal{X}] = \mathcal{X} + SS\mathcal{X} + B\mathcal{X} + SS(SS\mathcal{X}) + SS(B\mathcal{X}) + B(SS\mathcal{X}) + B(B\mathcal{X}) + \cdots$$

Definition. $\mathcal{E} \stackrel{\text{def}}{=} (SS)^*[\mathcal{Q}_2\mathcal{Q}_1]$ which is equal to $(\mathcal{Q}_1)^*[\mathcal{Q}_2\mathcal{Q}_1]$.

4 Easy Facts

Proposition R. *For any S-terms sets $\mathcal{X}$ and $\mathcal{Y}$: $(S\mathcal{M})^*[\mathcal{X}]\,\mathcal{Y} \xrightarrow{\circledast} \mathcal{X}\mathcal{Y}$, in particular, if $\mathcal{X}\mathcal{Y}\uparrow$ then $(S\mathcal{M})^*[\mathcal{X}]\,\mathcal{Y}\uparrow$.*

Results ([16]). $\mathcal{E}\mathcal{E}\uparrow, \quad \mathcal{Q}_3\mathcal{Q}_2\mathcal{Q}_1\uparrow, \quad (\mathcal{Q}_3\mathcal{Q}_2\mathcal{Q}_1 + \mathcal{E}\mathcal{E})\uparrow.$

5 Classification

We can limit ourselves to S-terms of the form $\mathcal{N}\mathcal{N}$. We proceed by classifying all S-terms in $\mathcal{N}$ into different classes $\mathcal{H}_0$, $\mathcal{H}_1$, $\mathcal{L}_0$, $\mathcal{L}_1$, ...

$$\mathcal{H}_0 \stackrel{\text{def}}{=} (SS+B)^*[S+S\mathcal{N}+SBS+SB(SS)], \quad \mathcal{H}_1 \stackrel{\text{def}}{=} (SS+B)^*[\mathcal{Q}_3\mathcal{Q}_2+S\mathcal{Q}_3\mathcal{M}]$$

Facts. $\mathcal{H}_0 \subseteq \mathcal{N}$, unlike $\mathcal{H}_1$, and $\mathcal{H}_0$, $\mathcal{H}_1$ are disjoint.

Result 1 ([16]). *$\mathcal{H}_0$ and $\mathcal{H}_1$ cover $\mathcal{N}$*

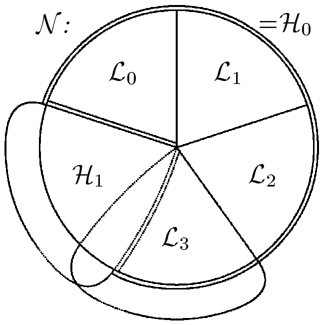
We further refine and dissect $\mathcal{H}_0$ into more mutually disjoint sets:

$$\mathcal{L}_0 \stackrel{\text{def}}{=} (SS)^*[S+S\mathcal{N}] \qquad \mathcal{L}_2 \stackrel{\text{def}}{=} (SS)^*[B(SS)+BB]$$
$$\mathcal{L}_1 \stackrel{\text{def}}{=} (SS)^*[BS+SBS] \qquad \mathcal{L}_3 \stackrel{\text{def}}{=} (SS)^*[SB(SS)+B\mathcal{Q}_3]$$

Facts. $\mathcal{L}_{012} \stackrel{\text{def}}{=} \mathcal{L}_0+\mathcal{L}_1+\mathcal{L}_2 \subseteq \mathcal{H}_0$ and $\mathcal{L}_0$, $\mathcal{L}_1$, $\mathcal{L}_2$, $\mathcal{L}_3$ are mutually disjoint.

Result 2 ([16]). *$\mathcal{L}_0$, $\mathcal{L}_1$, $\mathcal{L}_2$, and $\mathcal{L}_3$ cover $\mathcal{H}_0$.*

From Results 1 and 2, we partition $\mathcal{N}$ according to the sideways diagram (the circle represents $\mathcal{N}$, double lines surround $\mathcal{H}_0$; notice that $\mathcal{H}_1$ and $\mathcal{L}_3$ intersect both inside and outside of $\mathcal{N}$, but that will not be a problem).



6 Structure of the Proof

With the above definitions and letting $\mathcal{L}_{23} \stackrel{\text{def}}{=} \mathcal{L}_2+\mathcal{L}_3$ and $\mathcal{L}_{123} \stackrel{\text{def}}{=} \mathcal{L}_1+\mathcal{L}_2+\mathcal{L}_3$, the Theorem can be shown by proving (see [16]) the following parts:

Part 1. $\mathcal{L}_0\mathcal{N}\downarrow$

Part 2. $\mathcal{H}_0\mathcal{L}_0\downarrow$

Part 3. $\mathcal{L}_1\mathcal{H}_0\downarrow$

Part 4. $\mathcal{L}_2\mathcal{L}_1\downarrow$

Part 5. $\mathcal{L}_{23}\mathcal{L}_{23}\uparrow$

Part 6. $\mathcal{L}_3\mathcal{L}_1\uparrow$

Part 7. $\mathcal{H}_1\mathcal{Q}_2\uparrow$

Part 8. $\mathcal{L}_{123}\mathcal{H}_1\uparrow$

Part 9. Whether $(\mathcal{H}_1 \cap \mathcal{N})(SS+S)\uparrow$ can be decided by reduction to parts 1 to 8.

7 A Grammar for $\langle \mathcal{H}_0 \rangle$

¿From this point on, we use the angle brackets $\langle\cdot\rangle$ to denote the set of *"predecessors"* for a given set. That is, for any set $\mathcal{A}$,

$$\langle \mathcal{A} \rangle \stackrel{\text{def}}{=} \{\, x \in \mathcal{M} \mid x \stackrel{*}{\longrightarrow} \mathcal{A} \,\}\,.$$

It is our objective to develop a context free grammar to recognize $\langle\mathcal{N}\rangle$. Naturally, we will use these sets of predecessors as non-terminal symbols in our grammar. Many of the technical proofs have been omitted.

The terms in $\langle\mathcal{N}\rangle$ are either S or terms of from $\langle\mathcal{N}\rangle\langle\mathcal{N}\rangle$. To describe the terms of the second form we use the classification of $\mathcal{N}$ into $\mathcal{H}_0$ and $\mathcal{H}_1$, and the further classification of $\mathcal{H}_0$ into $\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2$, and $\mathcal{L}_3$. With these classes, the applications described in PARTS 1 through 6 cover all possible terms in $\mathcal{N}\mathcal{N}$. Because of the incomplete nature of the result in Part 6 we need to the define the following sets:

$$\mathcal{N}^{-S} \stackrel{\text{def}}{=} \{\, n \in \mathcal{N} \mid nS \stackrel{*}{\longrightarrow} \mathcal{N}\,\} \qquad \text{and}$$
$$\mathcal{N}^{-SS} \stackrel{\text{def}}{=} \{\, n \in \mathcal{N} \mid n(SS) \stackrel{*}{\longrightarrow} \mathcal{N}\,\}\,.$$

Then, the results from PARTS 1-6 prove the sufficiency of:

$$\langle\mathcal{N}\rangle \ ::= \ S \mid \langle\mathcal{L}_0\rangle\,\langle\mathcal{N}\rangle \mid \langle\mathcal{H}_0\rangle\,\langle\mathcal{L}_0\rangle \mid \langle\mathcal{L}_1\rangle\,\langle\mathcal{H}_0\rangle \mid \langle\mathcal{L}_2\rangle\,\langle\mathcal{L}_1\rangle \mid \langle\mathcal{N}^{-S}\rangle\,S \mid \langle\mathcal{N}^{-SS}\rangle\,(\,S\,S\,)$$

(Here, the production symbol ::= can be substituted by equality when the disjunction symbols | are substituted by union.) In this section we will expand $\langle\mathcal{H}_0\rangle$.

Before proceeding, we need:

Proposition 1. *The sets $\mathcal{Q}_3\mathcal{Q}_2$, $S\mathcal{Q}_3\mathcal{M}$, and, by extension, $\mathcal{H}_1$ are closed under reduction, that is, if a term in either set reduces, the resulting term lies in the same set.*

The result above reveals that $\langle\mathcal{H}_0\rangle$ and $\mathcal{Q}_3\mathcal{Q}_2$ are disjoint. Since $\mathcal{H}_0 \subseteq \mathcal{N}$, this fact is equivalent to $\langle\mathcal{H}_0\rangle \subseteq (S{+}SS{+}B)\langle\mathcal{N}\rangle + \langle\mathcal{N}\rangle(S{+}SS)$. This will facilitate the grammatical description of $\langle\mathcal{H}_0\rangle$. For this, we introduce the following sets:

$$\mathcal{H}_0^{-S} \stackrel{\text{def}}{=} \{\, n \in \mathcal{N} \mid nS \stackrel{*}{\longrightarrow} \mathcal{H}_0\,\} \qquad \text{and}$$
$$\mathcal{H}_0^{-SS} \stackrel{\text{def}}{=} \{\, n \in \mathcal{N} \mid n(SS) \stackrel{*}{\longrightarrow} \mathcal{H}_0\,\}\,.$$

Hence, we can describe $\langle\mathcal{H}_0\rangle$ completely with:

$$\langle\mathcal{H}_0\rangle \ ::= \ S \mid S\,\langle\mathcal{N}\rangle \mid S\,S\,\langle\mathcal{H}_0\rangle \mid B\,\langle\mathcal{H}_0\rangle \mid \langle\mathcal{H}_0^{-S}\rangle\,S \mid \langle\mathcal{H}_0^{-SS}\rangle\,(\,S\,S\,)$$

because the application of other pairs of strings in $\langle\mathcal{N}\rangle\langle\mathcal{N}\rangle$ results in $\mathcal{H}_1$. However, we are left with the task of producing rules for $\langle\mathcal{H}_0^{-S}\rangle$ and $\langle\mathcal{H}_0^{-SS}\rangle$. We start with:

Proposition 2. $\mathcal{H}_0^{-SS} = (SS)^*[S + SS + B + SB\,]$.

Readily, $\mathcal{H}_0^{-SS} \subseteq \mathcal{L}_0 \subseteq \mathcal{H}_0$. We classify the terms in $\mathcal{H}_0^{-SS}$ as follows: $S \in \mathcal{H}_0^{-SS}$; the terms in $\mathcal{H}_0^{-SS}$ of the form $S\mathcal{N}$ are $SS + B + SB$; and the terms in $\mathcal{H}_0^{-SS}$ of the form $S\mathcal{N}\mathcal{N}$ are the terms in $SS\mathcal{H}_0^{-SS}$.

Proposition 3. $\{\, n \in \mathcal{N} \mid n(SS) \stackrel{*}{\longrightarrow} \mathcal{H}_0^{-SS}\,\} = S + SS$.

Let

$$\mathcal{K}_0 \stackrel{\text{def}}{=} (SS)^*[S], \qquad \mathcal{K}_1 \stackrel{\text{def}}{=} (SS)^*[SS], \quad \text{and} \qquad \mathcal{K}_{01} \stackrel{\text{def}}{=} \mathcal{K}_0 + \mathcal{K}_1 = (SS)^*[S+SS].$$

Proposition 4. $\{ n \in \mathcal{N} \mid nS \stackrel{*}{\longrightarrow} \mathcal{H}_0^{-SS} \} = \mathcal{K}_{01}$.

Proposition 5. $\mathcal{H}_0^{-S} = (SS)^*[S + SS + B + SB + S\,\mathcal{K}_{01}\,S]$.

We can define $\mathcal{L}_0^{-S}$, $\mathcal{L}_1^{-S}$, $\mathcal{L}_2^{-S}$, and $\mathcal{L}_3^{-S}$ in a manner similar to that of $\mathcal{H}_0^{-S}$. I.e.,

$$\mathcal{L}_i^{-S} \stackrel{\text{def}}{=} \{ n \in \mathcal{N} \mid nS \stackrel{*}{\longrightarrow} \mathcal{L}_i \}$$

for $i = 0, 1, 2, 3$. With these definitions:

Corollary 1.

$$\begin{aligned}
\mathcal{L}_0^{-S} &= (SS)^*[S + SS] = \mathcal{K}_{01}, \\
\mathcal{L}_1^{-S} &= (SS)^*[B + SB], \\
\mathcal{L}_2^{-S} &= (SS)^*[BS] \subseteq (SS)^*[S\,\mathcal{K}_{01}\,S], \qquad \textit{and} \\
\mathcal{L}_3^{-S} &= (SS)^*[S\,\mathcal{K}_{01}\,S] - (SS)^*[SSS] - (SS)^*[BS].
\end{aligned}$$

For any S-term m if $m = m_l m_r$ for some terms m_l and m_r, we say m_r is a *right sub-term* of m. We extend the notion of right sub-term to include its reflexive transitive closure. We now show some results about right sub-term ahead.

Proposition 6. *Given any S-term m. For any m' such that $m \stackrel{*}{\longrightarrow} m'$, every $n \in \mathcal{N}$ that is a right sub-term of m is also a right sub-term of m'.*

Corollary 2. *For any $n \in \mathcal{N}$, suppose $mc \stackrel{*}{\longrightarrow} n$ for some $m \in \mathcal{M}$ and $c \in \mathcal{N}$. Then, c is a right sub-term of n.*

Corollary 3. *Suppose m is the reductum reduced from some redex $Sadc$ with $c \in \mathcal{N}$ and $m \stackrel{*}{\longrightarrow} m_1 m_2$. Then c is a proper right sub-term of both m_1 and m_2.*

Remark 1. Given $\mathcal{A} = (SS)^*[\mathcal{D}]$ with $\mathcal{D} \subseteq \mathcal{N}$, suppose $xy \stackrel{*}{\longrightarrow} \mathcal{A}$ for some x and y. Then, either $xy \stackrel{*}{\longrightarrow} \mathcal{D}$, or $x = SS$, or $y = S$. This is justified as follows: The first two options are trivial. If none of those two options are satisfied, we would face the reduction $xy \stackrel{*}{\longrightarrow} SSa \in \mathcal{A}$ (for some $a \in \mathcal{A}$!). In such case, Corollary 3 imposes the third choice by stating (the normal form of) y is a proper right sub-term of SS. (Note for any $d \in \mathcal{D}$, we can search for all pairs (x, y) so that $xy \stackrel{*}{\longrightarrow} d$ by exhausting all pairs that satisfy $|xy| \leq |d|$.)

Corollary 3 shows that no term is $S\mathcal{N}S$ is a reductum. Neither terms in $S + S\mathcal{N}$ are reducta. After this, Remark 1 above establishes how to determine all xy such that $xy \stackrel{*}{\longrightarrow} (SS)^*[S + S\mathcal{N} + B + SB + S\mathcal{L}_0^{-S}S] = \mathcal{H}_0^{-S}$ (recall Proposition 5 and Corollary 1). For instance, let:

$$\mathcal{K}_2 \stackrel{\text{def}}{=} (SS)^*[S\,\mathcal{L}_0^{-S}] \qquad \text{and} \qquad \mathcal{K}_3 \stackrel{\text{def}}{=} (SS)^*[S\,\mathcal{L}_0^{-S}\,S].$$

For these, Proposition R tells us:

$$\mathcal{K}_2\, S \;\xrightarrow{*}\; \mathcal{K}_3\,.$$

while Remark 1 ensures no other reductions result in $\mathcal{K}_3$. That ensures the sufficiency of:

$$\langle \mathcal{K}_3 \rangle \;::=\; \langle \mathcal{K}_2 \rangle\, S \mid S\, S\, \langle \mathcal{K}_3 \rangle$$

This way, we can produce a "complete" context free grammar for $\langle \mathcal{H}_0 \rangle$ using on the following set of rules:

$$\begin{array}{rlcl}
\text{(i)} & \langle \mathcal{H}_0 \rangle & ::= & S \mid S\, \langle \mathcal{N} \rangle \mid S\, S\, \langle \mathcal{H}_0 \rangle \mid B\, \langle \mathcal{H}_0 \rangle \mid \langle \mathcal{H}_0^{-S} \rangle\, S \mid \langle \mathcal{H}_0^{-SS} \rangle\, (\, S\, S\,) \\
\text{(ii)} & \langle \mathcal{H}_0^{-S} \rangle & ::= & \langle \mathcal{H}_0^{-SS} \rangle \mid \langle \mathcal{K}_3 \rangle \\
\text{(iii)} & \langle \mathcal{H}_0^{-SS} \rangle & ::= & \langle \mathcal{L}_0^{-S} \rangle \mid \langle \mathcal{L}_1^{-S} \rangle \\
\text{(iv)} & \langle \mathcal{L}_0^{-S} \rangle & ::= & \langle \mathcal{K}_0 \rangle \mid \langle \mathcal{K}_1 \rangle \\
\text{(v)} & \langle \mathcal{L}_1^{-S} \rangle & ::= & B \mid S\, B \mid S\, S\, \langle \mathcal{L}_1^{-S} \rangle \\
\text{(vi)} & \langle \mathcal{K}_3 \rangle & ::= & \langle \mathcal{K}_2 \rangle\, S \mid S\, S\, \langle \mathcal{K}_3 \rangle \\
\text{(vii)} & \langle \mathcal{K}_2 \rangle & ::= & \langle \mathcal{K}_1 \rangle \mid S\, \langle \mathcal{L}_0^{-S} \rangle \mid S\, S\, \langle \mathcal{K}_2 \rangle \\
\text{(viii)} & \langle \mathcal{K}_1 \rangle & ::= & \langle \mathcal{K}_0 \rangle\, S \mid S\, S\, \langle \mathcal{K}_1 \rangle \\
\text{(ix)} & \langle \mathcal{K}_0 \rangle & ::= & S \mid \langle \mathcal{K}_1 \rangle\, S \mid S\, S\, \langle \mathcal{K}_0 \rangle \\
\text{(x)} & B & ::= & S\, (\, S\, S\,)
\end{array}$$

Our grammar for $\langle \mathcal{N} \rangle$ also needs rules for $\langle \mathcal{L}_0 \rangle$, $\langle \mathcal{L}_1 \rangle$, and $\langle \mathcal{L}_2 \rangle$. These rules also follow easily after Remark 1. We have shown already $\mathcal{L}_0^{-S}$, $\mathcal{L}_1^{-S}$, and $\mathcal{L}_2^{-S}$ in Corollary 1. Indeed, we already have rules for $\langle \mathcal{L}_0^{-S} \rangle$ and $\langle \mathcal{L}_1^{-S} \rangle$. However, $\mathcal{L}_2^{-S}$ is simply presented as a subset of $\mathcal{K}_3$. For that reason, we introduce:

$$\mathcal{K}_4 \;\stackrel{\text{def}}{=}\; (SS)^*[B]\,.$$

Then (note $SSS(SS) \to BB \in \mathcal{L}_2$ but no $xy \to B(SS)$!),

$$\begin{array}{rlcl}
\text{(xi)} & \langle \mathcal{L}_0 \rangle & ::= & S \mid S\, \langle \mathcal{N} \rangle \mid \langle \mathcal{L}_0^{-S} \rangle\, S \mid S\, S\, \langle \mathcal{L}_0 \rangle \\
\text{(xii)} & \langle \mathcal{L}_1 \rangle & ::= & \langle \mathcal{L}_1^{-S} \rangle\, S \mid S\, S\, \langle \mathcal{L}_1 \rangle \\
\text{(xiii)} & \langle \mathcal{L}_2 \rangle & ::= & B\, (\, S\, S\,) \mid S\, S\, S\, (\, S\, S\,) \mid B\, B \mid \langle \mathcal{L}_2^{-S} \rangle\, S \mid S\, S\, \langle \mathcal{L}_2 \rangle \\
\text{(xiv)} & \langle \mathcal{L}_2^{-S} \rangle & ::= & \langle \mathcal{K}_4 \rangle\, S \mid S\, S\, \langle \mathcal{L}_2^{-S} \rangle \\
\text{(xv)} & \langle \mathcal{K}_4 \rangle & ::= & B \mid S\, S\, \langle \mathcal{K}_4 \rangle
\end{array}$$

8 Beyond $\mathcal{H}_0$: the Sets $\mathcal{N}^{-S}$ and $\mathcal{N}^{-SS}$

Recall:

$$\mathcal{N}^{-S} \stackrel{\text{def}}{=} \{\, n \in \mathcal{N} \mid nS \stackrel{*}{\longrightarrow} \mathcal{N} \,\} \qquad \text{and}$$
$$\mathcal{N}^{-SS} \stackrel{\text{def}}{=} \{\, n \in \mathcal{N} \mid n(SS) \stackrel{*}{\longrightarrow} \mathcal{N} \,\}\,.$$

Since $S + SS \subseteq \mathcal{L}_0$, Part 2, $\mathcal{H}_0\,\mathcal{L}_0\!\downarrow$, easily shows $\mathcal{H}_0 \subseteq \mathcal{N}^{-S}$ and $\mathcal{H}_0 \subseteq \mathcal{N}^{-SS}$. Now, we proceed to complete the representation of the terms in $\mathcal{N}^{-SS}$. Let:

$$\mathcal{L}_4 \stackrel{\text{def}}{=} (SS + B)^*[S\,\mathcal{H}_0^{-SS}(S + SS)]\,.$$

We will show $\mathcal{N}^{-SS} = \mathcal{H}_0 + \mathcal{L}_4$. After this, we can revise the grammatical rule for $\langle\mathcal{N}\rangle$ to:

$$\langle\mathcal{N}\rangle \;::=\; S \mid \langle\mathcal{L}_0\rangle\,\langle\mathcal{N}\rangle \mid \langle\mathcal{H}_0\rangle\,\langle\mathcal{L}_0\rangle \mid \langle\mathcal{L}_1\rangle\,\langle\mathcal{H}_0\rangle \mid \langle\mathcal{L}_2\rangle\,\langle\mathcal{L}_1\rangle \mid \langle\mathcal{N}^{-S}\rangle\,S \mid \langle\mathcal{L}_4\rangle\,(\,S\,S\,)$$

Indeed, we show:

Part 10. $\mathcal{L}_4(SS)\!\downarrow$ *and* $(\,\mathcal{H}_1 \cap \mathcal{N} - \mathcal{L}_4)\,(SS)\!\uparrow$.

Corollary 4. $\mathcal{H}_1 \cap \mathcal{N}^{-SS} \;=\; (SS + B)^*[S(SB + SS\,\mathcal{H}_0^{-SS})(S + SS)]\,.$

The right sub-terms of $\mathcal{L}_4$ are S, SS, and $\mathcal{L}_4$. Because of Corollary 2, for the grammar of $\langle\mathcal{L}_4\rangle$ we only need to describe the sets:

$$\mathcal{L}_4^{-S} \stackrel{\text{def}}{=} \{\, n \in \mathcal{N} \mid nS \stackrel{*}{\longrightarrow} \mathcal{L}_4 \,\} \qquad \text{and}$$
$$\mathcal{L}_4^{-SS} \stackrel{\text{def}}{=} \{\, n \in \mathcal{N} \mid n(SS) \stackrel{*}{\longrightarrow} \mathcal{L}_4 \,\}$$

so we could write:

$$\langle\mathcal{L}_4\rangle \;::=\; S\,S\,\langle\mathcal{L}_4\rangle \mid B\,\langle\mathcal{L}_4\rangle \mid \langle\mathcal{L}_4^{-S}\rangle\,S \mid \langle\mathcal{L}_4^{-SS}\rangle\,(\,S\,S\,)$$

We start by proving:

Proposition 7. $\mathcal{L}_4^{-SS} = (SS)^*[S\,\mathcal{H}_0^{-SS}]\,.$

Recall: $\mathcal{K}_0 \stackrel{\text{def}}{=} (SS)^*[S] \;\subseteq\; \mathcal{K}_{01} \stackrel{\text{def}}{=} (SS)^*[S + SS] \;=\; \mathcal{L}_0^{-S}\,.$

Proposition 8. $\{\, n \in \mathcal{N} \mid nS \in \mathcal{L}_4^{-SS} \,\} \;=\; \mathcal{K}_0\,.$

Proposition 9. $\mathcal{L}_4^{-S} \;=\; (SS)^*[S\,\mathcal{H}_0^{-SS} \;+\; S\mathcal{K}_0 S\,]\,.$

Note $\mathcal{L}_4^{-SS} \subseteq \mathcal{L}_4^{-S}$ and the difference $\mathcal{L}_4^{-S} - \mathcal{L}_4^{-SS} = (SS)^*[S\mathcal{K}_0 S] \subseteq \mathcal{H}_0^{-S}$ so $\{\, n \in \mathcal{N} \mid nS \stackrel{*}{\longrightarrow} \mathcal{H}_1 \cap \mathcal{N}^{-SS} \,\} \;\subseteq\; \mathcal{L}_4^{-SS}$. Let:

$$\mathcal{K}_5 \stackrel{\text{def}}{=} (SS)^*[S\,\mathcal{K}_0] \qquad \text{and} \qquad \mathcal{K}_6 \stackrel{\text{def}}{=} (SS)^*[S\,\mathcal{K}_0\,S]\,.$$

Then, we can specify a grammar for $\langle\mathcal{L}_4\rangle$:

$$\begin{array}{rlcl}
\text{(xvi)} & \langle\mathcal{L}_4\rangle & ::= & S\,S\,\langle\mathcal{L}_4\rangle \mid B\,\langle\mathcal{L}_4\rangle \mid \langle\mathcal{L}_4^{-s}\rangle\,S \mid \langle\mathcal{L}_4^{-ss}\rangle\,(\,S\,S\,) \\
\text{(xvii)} & \langle\mathcal{L}_4^{-s}\rangle & ::= & \langle\mathcal{L}_4^{-ss}\rangle \mid \langle\mathcal{K}_6\rangle \\
\text{(xviii)} & \langle\mathcal{L}_4^{-ss}\rangle & ::= & \langle\mathcal{K}_0\rangle\,S \mid S\,\langle\mathcal{H}_0^{-ss}\rangle \mid S\,S\,\langle\mathcal{L}_4^{-ss}\rangle \\
\text{(xix)} & \langle\mathcal{K}_6\rangle & ::= & \langle\mathcal{K}_5\rangle\,S \mid S\,S\,\langle\mathcal{K}_6\rangle \\
\text{(xx)} & \langle\mathcal{K}_5\rangle & ::= & \langle\mathcal{K}_0\rangle\,S \mid S\,\langle\mathcal{K}_0\rangle \mid S\,S\,\langle\mathcal{K}_5\rangle
\end{array}$$

Now, the only piece missing in our grammar for $\langle\mathcal{N}\rangle$ is $\langle\mathcal{N}^{-s}\rangle$. To fill this gap, we first prove:

Part 11. $\mathcal{N}^{-s} \;=\; (S\mathcal{L}_0^{-s})^*[S + S\mathcal{N} + S\mathcal{H}_0^{-s}\mathcal{L}_0^{-s} + S\mathcal{L}_1^{-s}\mathcal{H}_0^{-s} + S\mathcal{L}_2^{-s}\mathcal{L}_1^{-s} + S\mathcal{L}_4^{-s}S\,]$.

Remark 2. Recall $\mathcal{H}_0 \subseteq \mathcal{N}^{-s} \cap \mathcal{N}^{-ss}$. Now, we can easily check:

$$\mathcal{L}_4 \;=\; (SS+B)^*[S\mathcal{H}_0^{-ss}(S+SS)] \;\subseteq\; (S\mathcal{L}_0^{-s})^*[S\mathcal{H}_0^{-s}\mathcal{L}_0^{-s}] \;\subseteq\; \mathcal{N}^{-s}\,.$$

It is no surprise $\mathcal{N}^{-ss} \subseteq \mathcal{N}^{-s}$.

After this result, to investigate $\langle\mathcal{N}^{-s}\rangle$ we only need investigate the redexes that reduce into $\mathcal{N}^{-s}$. Suppose n_1n_2 is a redex with $n_1, n_2 \in \mathcal{N}$, and $n_1n_2 \xrightarrow{+} n_0 \in \mathcal{N}^{-s}$. Since n_0 is a reductum, $n_0 \notin S + S\mathcal{N} + S\mathcal{N}S$. Then, from the expression for $\mathcal{N}^{-s}$ in Part 11, $n_0 \in S\mathcal{L}_0^{-s}\mathcal{N}^{-s} + S\mathcal{H}_0^{-s}\mathcal{L}_0^{-s} + S\mathcal{L}_1^{-s}\mathcal{H}_0^{-s} + S\mathcal{L}_2^{-s}\mathcal{L}_1^{-s}$. In short, $n_1n_2 \xrightarrow{+} S\mathcal{H}_0^{-s}\mathcal{N}^{-s}$. We proceed to analyze exhaustively the choices for n_1: We have either $n_1 \in \mathcal{H}_0$, or $n_1 \in \mathcal{H}_1 \cap \mathcal{N}^{-ss}$ and $n_2 = SS$, or $n_1 \in \mathcal{H}_1 \cap \mathcal{N}^{-s}$ and $n_2 = S$.

#1. Suppose $n_1 \in \mathcal{H}_0$. We need to examine various cases from $n_1 \in (SS)^*[S + S\mathcal{N} + SBS + SB(SS) + B\mathcal{H}_0]$:

(i) Suppose $n_1 \in (SS)^*[S]$. If $n_1 = S$, the S-term $n_1n_2 \in S\mathcal{N}$ ($\subseteq \mathcal{N}^{-s}$!) is not a redex. Then, we are supposing $n_1 = (SS)^{k+1}[S] = SS\big((SS)^k[S]\big)$ for some $k \geq 0$. Then,

$$\begin{aligned} n_1n_2 \;=\; SS\big((SS)^k[S]\big)\,n_2 \;\longrightarrow\; Sn_2\big((SS)^k[S]\,n_2\big) \;\xrightarrow{*} \\ S\mathcal{L}_0^{-s}\mathcal{N}^{-s} + S\mathcal{H}_0^{-s}\mathcal{L}_0^{-s} + S\mathcal{L}_1^{-s}\mathcal{H}_0^{-s} + S\mathcal{L}_2^{-s}\mathcal{L}_1^{-s} \;\subseteq\; \mathcal{N}^{-s}\,. \end{aligned}$$

Therefore, either $n_2 \in \mathcal{L}_0^{-s}$ (and $(SS)^k[S]\,n_2 \xrightarrow{*} \mathcal{N}^{-s}$) or $(SS)^k[S]\,n_2 \xrightarrow{*} \mathcal{H}_0^{-s}$.

– Suppose $n_2 \in \mathcal{L}_0^{-s}$. Then, we may verify:

$$n_1\,n_2 \;\in\; (SS)^*[S]\,\mathcal{L}_0^{-s} \;\xrightarrow{*}\; (S\mathcal{L}_0^{-s})^*[S\mathcal{L}_0^{-s}] \;\subseteq\; (S\mathcal{L}_0^{-s})^*[S\mathcal{N}\,] \;\subseteq\; \mathcal{N}^{-s}\,.$$

– Suppose $(SS)^k[S]\,n_2 \xrightarrow{*} \mathcal{H}_0^{-s}$ and $n_2 \notin \mathcal{L}_0^{-s}$. Recalling $\mathcal{H}_0^{-s} = (SS)^*[S + SS + B + SB + S\mathcal{L}_0^{-s}S\,]$ and the rules for $\langle\mathcal{H}_0^{-s}\rangle$, we determine this needs $(SS)^k[S] = S$ and $n_2 = B$. Given this, we verify:

$$n_1\,n_2 \;=\; (SSS)B \;\longrightarrow\; SB(SB) \;\subseteq\; S\mathcal{L}_1^{-s}\mathcal{H}_0^{-s} \;\subseteq\; \mathcal{N}^{-s}\,.$$

(ii) Suppose $n_1 \in (SS)^*[Sn_1']$ for some $n_1' \in \mathcal{N}$. If $n_1 = Sn_1'$, the S-term $n_1 n_2 \in S\mathcal{N}\mathcal{N}$ is not a redex (the allowed values for n_1' and n_2 can be found in Part 11). Then, we are supposing $n_1 = (SS)^{k+1}[Sn_1'] = SS\big((SS)^k[Sn_1']\big)$ for some $k \geq 0$. Then,

$$n_1\, n_2 \;=\; SS\big((SS)^k[Sn_1']\big)\, n_2 \;\longrightarrow\; Sn_2\big((SS)^k[Sn_1']\, n_2\big) \;\stackrel{*}{\longrightarrow}\; S\mathcal{L}_0^{-S}\mathcal{N}^{-S} + S\mathcal{H}_0^{-S}\mathcal{L}_0^{-S} + S\mathcal{L}_1^{-S}\mathcal{H}_0^{-S} + S\mathcal{L}_2^{-S}\mathcal{L}_1^{-S} \;\subseteq\; \mathcal{N}^{-S}.$$

Therefore, either $n_2 \in \mathcal{L}_0^{-S}$ (and $(SS)^k[Sn_1]\, n_2 \stackrel{*}{\longrightarrow} \mathcal{N}^{-S}$) or $(SS)^k[Sn_1]\, n_2 \stackrel{*}{\longrightarrow} \mathcal{H}_0^{-S}$.

– Suppose $n_2 \in \mathcal{L}_0^{-S}$. Then,

$$n_1\, n_2 \in (SS)^*[Sn_1']\, \mathcal{L}_0^{-S} \;\stackrel{*}{\longrightarrow}\; (S\mathcal{L}_0^{-S})^*[Sn_1'\mathcal{L}_0^{-S}] \;\subseteq\; (S\mathcal{L}_0^{-S})^*[Sn_1'\mathcal{L}_0^{-S}] \;\subseteq\; \mathcal{N}^{-S}.$$

However, from Part 11, for the above statement, we need to match,

$$Sn_1'\mathcal{L}_0^{-S} \;\subseteq\; S\mathcal{L}_0^{-S}\mathcal{N}^{-S} + S\mathcal{H}_0^{-S}\mathcal{L}_0^{-S} + S\mathcal{L}_1^{-S}\mathcal{H}_0^{-S} + S\mathcal{L}_2^{-S}\mathcal{L}_1^{-S} + S\mathcal{L}_4^{-S}S\,.$$

This is satisfied only if $n_1' \in \mathcal{H}_0^{-S}$, for any $n_2 \in \mathcal{L}_0^{-S}$, or if $n_1' \in \mathcal{L}_4^{-S}$, for the particular case of $n_2 = S \in \mathcal{L}_0^{-S}$. These alternatives are verified with:

$$n_1\, n_2 \;\in\; (SS)^*[S\mathcal{H}_0^{-S}]\, \mathcal{L}_0^{-S} \;\stackrel{*}{\longrightarrow}\; (S\mathcal{L}_0^{-S})^*[S\mathcal{H}_0^{-S}\mathcal{L}_0^{-S}] \;\subseteq\; \mathcal{N}^{-S} \qquad \text{and}$$

$$n_1\, n_2 \;\in\; (SS)^*[S\mathcal{L}_4^{-S}]\, S \;\stackrel{*}{\longrightarrow}\; (S\mathcal{L}_0^{-S})^*[S\mathcal{L}_4^{-S}S] \;\subseteq\; \mathcal{N}^{-S}.$$

– Suppose $(SS)^k[Sn_1']\, n_2 \stackrel{*}{\longrightarrow} \mathcal{H}_0^{-S}$ and $n_2 \notin \mathcal{L}_0^{-S}$. Then, $(SS)^k[Sn_1'] = SS$ and $n_2 \in \mathcal{H}_0^{-S} - \mathcal{L}_0^{-S}$. With $n_1 = SS(SS)$ and $n_2 \in \mathcal{L}_i^{-S}$ for $i = 1$, 2, or 3,

$$n_1\, n_2 \;\in\; SS(SS)\, \mathcal{L}_i^{-S} \;\longrightarrow\; S\mathcal{L}_i^{-S}(SS\mathcal{L}_i^{-S}) \;\subseteq\; S\mathcal{L}_0^{-S}\mathcal{N}^{-S} + S\mathcal{H}_0^{-S}\mathcal{L}_0^{-S} + S\mathcal{L}_1^{-S}\mathcal{H}_0^{-S} + S\mathcal{L}_2^{-S}\mathcal{L}_1^{-S} \;\subseteq\; \mathcal{N}^{-S}.$$

This is only satisfied when $i = 1$ (note $i = 0$ is not an option now!). Then, we may verify,

$$n_1\, n_2 \;\in\; SS(SS)\, \mathcal{L}_1^{-S} \;\longrightarrow\; S\mathcal{L}_1^{-S}\mathcal{H}_0^{-S} \;\subseteq\; \mathcal{N}^{-S}.$$

(iii) Suppose $n_1 \in (SS)^*[SBS]$. Then,

$$n_1\, n_2 \;=\; (SS)^*[SBS]\, n_2 \;\stackrel{*}{\longrightarrow}\; (Sn_2)^*[SBS\, n_2] \;\stackrel{*}{\longrightarrow}\; \mathcal{N}^{-S}.$$

For this, we need to verify $SBS\, n_2 \stackrel{*}{\longrightarrow} \mathcal{N}^{-S}$ first:

$$SBS\, n_2 \;\longrightarrow\; Bn_2(Sn_2) \;\stackrel{2}{\longrightarrow}\; S\,\big(n_2(Sn_2)\big)\,\big(Sn_2(n_2(Sn_2))\big) \;\stackrel{*}{\longrightarrow}\; S\mathcal{H}_0^{-S}\mathcal{N}^{-S}.$$

Thus, $n_2(Sn_2) \stackrel{*}{\longrightarrow} \mathcal{H}_0^{-S}$. To satisfy this, we need $n_2 \in S + SS$. This being provided, we can verify $SBS\, n_2 \stackrel{*}{\longrightarrow} \mathcal{N}^{-S}$ with:

$$SBSS \stackrel{*}{\longrightarrow} SB(SSB) \in S\mathcal{L}_1^{-S}\mathcal{H}_0^{-S} \subseteq \mathcal{N}^{-S} \qquad \text{but}$$

$$SBS(SS) \notin \langle\mathcal{N}^{-S}\rangle, \text{ because } SBS(SS)S\uparrow \text{ (from part 8 after some reductions).}$$

Therefore, given $n_1 \in (SS)^*[SBS]$, only for $n_2 = S$ we may verify:

$$n_1\, n_2 \;\in\; (SS)^*[SBS]\, S \;\xrightarrow{*}\; (SS)^*[SBS\, S] \;\xrightarrow{*}\; (S\mathcal{L}_0^{-S})^*[\mathcal{N}^{-S}] \;\subseteq\; \mathcal{N}^{-S}.$$

(iv) Suppose $n_1 = (SS)^*[SB(SS)]$. Then:

$$n_1\, n_2 \;=\; (SS)^*[SB(SS)]\, n_2 \;\xrightarrow{*}\; (Sn_2)^*[SB(SS)\, n_2] \;\xrightarrow{*}\; \mathcal{N}^{-S}.$$

For this, we need to verify $SB(SS)\, n_2 \xrightarrow{*} \mathcal{N}^{-S}$ first:

$$SB(SS)\, n_2 \;\longrightarrow\; Bn_2(SSn_2) \;\xrightarrow{2}\; S\left(n_2(SSn_2)\right)\left(SSn_2(n_2(SSn_2))\right) \;\longrightarrow\; S\left(n_2(SSn_2)\right)\left(S\left(n_2(SSn_2)\right)\left(n_2(n_2(SSn_2))\right)\right) \;\xrightarrow{*}\; S\mathcal{H}_0^{-S}\mathcal{N}^{-S}.$$

Thus, $n_2(SSn_2) \xrightarrow{*} \mathcal{H}_0^{-S}$, but for this we need $n_2 = SS$. Then:

$$SB(SS)\, n_2 \xrightarrow{*} S\left(SS(SS(SS))\right)\left(S\left(SS(SS(SS))\right)\left(SS(SS(SS(SS)))\right)\right) \;\in\; (S\mathcal{L}_0^{-S})^*[S\mathcal{H}_0^{-S}\mathcal{L}_0^{-S}] \;\subseteq\; \mathcal{N}^{-S}.$$

Therefore, given $n_1 \in (SS)^*[SB(SS)]$, only for $n_2 = SS$ we may verify:

$$n_1\, n_2 \;\in\; (SS)^*[SB(SS)]\,(SS) \;\xrightarrow{*}\; (S(S{+}SS))^*[SB(SS)\,(SS)] \;\xrightarrow{*}\; (S\mathcal{L}_0^{-S})^*[\mathcal{N}^{-S}] \;\subseteq\; \mathcal{N}^{-S}.$$

(v) Suppose $n_1 = (SS)^* B n_1'$ for some $n_1' \in \mathcal{H}_0$. Then,

$$n_1\, n_2 \;=\; (SS)^*[Bn_1']\, n_2 \;\xrightarrow{*}\; (Sn_2)^*[Bn_1' n_2] \;\xrightarrow{*}\; \mathcal{N}^{-S}.$$

For this, we need to verify $Bn_1' n_2 \xrightarrow{*} \mathcal{N}^{-S}$ first:

$$Bn_1'n_2 \;\xrightarrow{2}\; S(n_1'n_2)(n_2(n_1'n_2)) \;\xrightarrow{*}\; S\mathcal{L}_0^{-S}\mathcal{N}^{-S}{+}S\mathcal{H}_0^{-S}\mathcal{L}_0^{-S}{+}S\mathcal{L}_1^{-S}\mathcal{H}_0^{-S}{+}S\mathcal{L}_2^{-S}\mathcal{L}_1^{-S} \;\subseteq\; \mathcal{N}^{-S}.$$

Therefore, either $n_1'n_2 \in \mathcal{L}_0^{-S}$ (and $n_2(n_1'n_2) \xrightarrow{*} \mathcal{N}^{-S}$) or $n_2(n_1'n_2) \xrightarrow{*} \mathcal{H}_0^{-S}$.

– Suppose $n_1'n_2 \in \mathcal{L}_0^{-S}$. Then, either $n_1' \in \mathcal{L}_0^{-S}$ and $n_2 = S$, or $n_1' = SS$ and $n_2 \in \mathcal{L}_0^{-S}$. For these alternatives we compute:

  $$B\mathcal{L}_0^{-S}S \;\xrightarrow{2}\; S(\mathcal{L}_0^{-S}S)(S(\mathcal{L}_0^{-S}S)) \;\longrightarrow\; S\mathcal{L}_0^{-S}(S\mathcal{L}_0^{-S}) \;\xrightarrow{*}\; (S\mathcal{L}_0^{-S})^*[S\mathcal{N}] \subseteq \mathcal{N}^{-S} \quad \text{and}$$

  $$B(SS)\mathcal{L}_0^{-S} \;\xrightarrow{2}\; S(SS\mathcal{L}_0^{-S})(\mathcal{L}_0^{-S}(SS\mathcal{L}_0^{-S})) \subseteq S\mathcal{L}_0^{-S}(\mathcal{L}_0^{-S}\mathcal{L}_0^{-S}) \;\xrightarrow{*}\; (S\mathcal{L}_0^{-S})^*[\mathcal{N}^{-S}] \subseteq \mathcal{N}^{-S}.$$

  (Note: $\mathcal{L}_0^{-S}\mathcal{L}_0^{-S} \xrightarrow{*} (S\mathcal{L}_0^{-S})^*[\mathcal{N}^{-S}]$ was verified in (i) and (ii) above.) This way, we may verify our choice we have with:

  $$n_1\, n_2 \;\in\; (SS)^*[B\mathcal{L}_0^{-S}]\, S \;\xrightarrow{*}\; (SS)^*[B\mathcal{L}_0^{-S}S] \;\xrightarrow{*}\; (S\mathcal{L}_0^{-S})^*[\mathcal{N}^{-S}] \;\subseteq\; \mathcal{N}^{-S} \quad \text{and}$$

  $$n_1\, n_2 \;\in\; (SS)^*[B(SS)]\, \mathcal{L}_0^{-S} \;\xrightarrow{*}\; (S\mathcal{L}_0^{-S})^*[B(SS)\mathcal{L}_0^{-S}] \;\xrightarrow{*}\; (S\mathcal{L}_0^{-S})^*[\mathcal{N}^{-S}] \;\subseteq\; \mathcal{N}^{-S}.$$

– Suppose $n_2(n_1' n_2) \xrightarrow{*} \mathcal{H}_0^{-S}$ and $n_1' n_2 \notin \mathcal{L}_0^{-S}$. Then, $n_1' = S$ and $n_2 = SS$. In this case, we may verify:

$$n_1\, n_2 \;\in\; (SS)^*[BS]\,(SS) \;\xrightarrow{*}\; (S(SS))^*[SB(SSB)] \;\in\; (S\mathcal{L}_0^{-S})^*[S\mathcal{H}_0^{-S}\mathcal{L}_1^{-S}] \;\subseteq\; \mathcal{N}^{-S}.$$

In summary, $n_1 n_2 \xrightarrow{*} \mathcal{N}^{-S}$ for $n_1 \in \mathcal{H}_0 - S - S\mathcal{N}$, only if:

$$\begin{array}{ll}
n_1\, n_2 \;\in\; (SS)^*[S]\;\mathcal{L}_0^{-S}\,, & n_1\, n_2 \;\in\; (SS)^*[SBS]\;S\,, \\
n_1\, n_2 \;=\; (SSS)\;B\,, & n_1\, n_2 \;\in\; (SS)^*[SB(SS)]\;(SS)\,, \\
n_1\, n_2 \;\in\; (SS)^*[S\mathcal{H}_0^{-S}]\;\mathcal{L}_0^{-S}\,, & n_1\, n_2 \;\in\; (SS)^*[B\mathcal{L}_0^{-S}]\;S\,, \\
n_1\, n_2 \;\in\; (SS)^*[S\mathcal{L}_4^{-S}]\;S\,, & n_1\, n_2 \;\in\; (SS)^*[B(SS)]\;\mathcal{L}_0^{-S}\,, \qquad \text{or} \\
n_1\, n_2 \;\in\; SS(SS)\;\mathcal{L}_1^{-S}\,, & n_1\, n_2 \;\in\; (SS)^*[BS]\;(SS)\,.
\end{array}$$

#2. Suppose $n_1 \in \mathcal{H}_1 \cap \mathcal{N}^{-SS}$ and $n_2 = SS$. Then, $n_1 \in (SS + B)^*[n_1']$ for some $n_1' \in S(SB + SS\mathcal{H}_0^{-SS})(S + SS)$. We will first show $n_1 n_2 \xrightarrow{*} (S\mathcal{N})^*[n_1' n_2]$. Then, we will show that $n_1' n_2 \xrightarrow{*} \mathcal{N}^{-S}$ is not possible, i.e., $n_1' n_2 S\!\uparrow$. ¿From these results and Proposition R we can verify $n_1 n_2 \xrightarrow{*} \mathcal{N}^{-S}$ is impossible.
With Proposition R, we compute:

$$n_1\, n_2 \;=\; (SS + B)^*[n_1']\,(SS) \;\xrightarrow{*}\; (B + SS(SS))^*[n_1'(SS)]\,.$$

Clearly $n_1'(SS) \in (S\mathcal{N})^*[n_1' n_2]$. Suppose $n \in (S\mathcal{N})^*[n_1' n_2]$. Then,

$$B\,n \;=\; B\big((S\mathcal{N})^*[n_1' n_2]\big) \;\subseteq\; (S\mathcal{N})^*[n_1' n_2] \qquad\qquad \text{and}$$

$$SS(SS)\,n \;\longrightarrow\; Sn(SSn) \;\subseteq\; (S\mathcal{N})^*[n] \;\subseteq\; (S\mathcal{N})^*[n_1' n_2]\,.$$

Therefore, $n_1 n_2 \xrightarrow{*} (S\mathcal{N})^*[n_1' n_2]$.
We show in no case $n_1' n_2 \xrightarrow{*} \mathcal{N}^{-S}$ with the following:

$$\begin{aligned}
n_1'\, n_2 \;\in\; & S(SB + SS\mathcal{H}_0^{-SS})(S + SS)\,(SS) \;\xrightarrow{*} \\
& S\big((SS)^*[SSS + SS(SS) + SSB + SB]\big)\big(S + SS\big)\big(SS\big) \;\xrightarrow{*} \\
& \quad \big((SS)^*[SSS + SS(SS) + SSB + SB](SS)\big)\big((S + SS)(SS)\big) \;\xrightarrow{*} \\
& \qquad \big((B)^*[BB + B(SS(SS)) + B(B(SS)) + SB(SS)]\big)\big(B + SS(SS)\big)\,.
\end{aligned}$$

The left component in the final expression (not yet in normal form!) is a subset of $\mathcal{H}_0$. However, every application of a term in this left component with a term in the right component, B or $SS(SS)$, fails to match any $n_3 n_4 \in \mathcal{H}_0\mathcal{N}$ such that $n_3 n_4 \xrightarrow{*} \mathcal{N}^{-S}$ discussed before.

#3. Suppose $n_1 \in \mathcal{H}_1 \cap \mathcal{N}^{-S}$ and $n_2 = S$. Then, $n_1 = S n_3 n_4$ for some $n_3, n_4 \in \mathcal{N}$. Naturally, $n_1 n_2 \;=\; S n_3 n_4 S \;\longrightarrow\; (n_3 S)(n_4 S)$ and $n_3 S \xrightarrow{*} \mathcal{H}_0 + \mathcal{H}_1$. We reject $n_3 S \xrightarrow{*} \mathcal{H}_1$ because if so, we would have to accept $n_1 n_2 \xrightarrow{*} \mathcal{H}_1(SS)$, but this resulting set was proven disjoint from $\langle \mathcal{N}^{-S} \rangle$ just above. Then, $n_3 S \xrightarrow{*} \mathcal{H}_0$. Let $n_3', n_4' \in \mathcal{N}$ be such that $n_3S \xrightarrow{*} n'_3$ and $n_4S \xrightarrow{*} n'_4$. Then, $n'_3n'_4 \in \mathcal{H}_0\mathcal{N}$. Therefore, either $n'_3n'_4$ matches one of the choices found in **#1**, when supposing "$n_1 \in \mathcal{H}_0$," or else $n'_3 \in S\mathcal{N}$. For the first alternative, we extract $n_1 \in S\mathcal{L}_0^{-S}\mathcal{L}_0^{-S} + S\mathcal{K}_4S$ from the result at the end of **#1** ($\mathcal{K}_4 = (\mathcal{L}_2^{-S})^{-S}$ justifies the $S\mathcal{K}_4S$ part). This is because n'_3 cannot be $B(SS)$, $SB(SS)$, nor in $\mathcal{L}_0 - \mathcal{L}_0^{-S}$, e.g., not in $(SS)^*[S\mathcal{L}_4^{-S}]$, and n'_4 cannot be S, B, nor in $\mathcal{L}_1^{-S}$. The second alternative forces $n'_3 = SS$, so $n_3 = S$ and $n_1n_2 = SSn_4S \to SS(n_4S)$, which still requires $n_4S \xrightarrow{*} \mathcal{N}^{-S}$. Therefore, the second alternative only introduces the possibility of having any number of prefixes SS. However, the base expression must be an S-term given from the first alternative. Therefore, $n_1n_2 \xrightarrow{*} \mathcal{N}^{-S}$ for $n_1 \in \mathcal{N}^{-S}$ and $n_2 = S$, only if:

$$n_1 \in (SS)^*[S\mathcal{L}_0^{-S}\mathcal{L}_0^{-S}] \qquad \text{or} \qquad n_1 \in (SS)^*[S\mathcal{K}_4S]\,.$$

Let

$$\mathcal{J}_1 \stackrel{\text{def}}{=} SSS\,, \qquad \mathcal{J}_6 \stackrel{\text{def}}{=} (SS)^*[B(SS)]\,,$$
$$\mathcal{J}_2 \stackrel{\text{def}}{=} (SS)^*[S\mathcal{H}_0^{-S}]\,, \qquad \mathcal{J}_7 \stackrel{\text{def}}{=} (SS)^*[B\mathcal{L}_0^{-S}]\,,$$
$$\mathcal{J}_3 \stackrel{\text{def}}{=} (SS)^*[S\mathcal{L}_4^{-S}]\,, \qquad \mathcal{J}_8 \stackrel{\text{def}}{=} (SS)^*[S\mathcal{L}_0^{-S}]\,,$$
$$\mathcal{J}_4 \stackrel{\text{def}}{=} SS(SS)\,, \qquad \mathcal{J}_9 \stackrel{\text{def}}{=} (SS)^*[S\mathcal{L}_0^{-S}S]\,, \qquad \text{and}$$
$$\mathcal{J}_5 \stackrel{\text{def}}{=} (SS)^*[SB(SS)]\,, \qquad \mathcal{J}_{10} \stackrel{\text{def}}{=} (SS)^*[S\mathcal{L}_0^{-S}\mathcal{L}_0^{-S} + S\mathcal{K}_4S]\,.$$

We complete the grammar as follows:

$$\begin{aligned}
&\text{(xxi)} & \langle\mathcal{N}^{-S}\rangle &::= S \mid S\,\langle\mathcal{N}\rangle \mid S\,\langle\mathcal{H}_0^{-S}\rangle\,\langle\mathcal{L}_0^{-S}\rangle \mid S\,\langle\mathcal{L}_1^{-S}\rangle\,\langle\mathcal{H}_0^{-S}\rangle \mid S\,\langle\mathcal{L}_2^{-S}\rangle\,\langle\mathcal{L}_1^{-S}\rangle \mid \\
&&& \quad S\,\langle\mathcal{L}_4^{-S}\rangle\,S \mid \langle\mathcal{K}_0\rangle\,\langle\mathcal{L}_0^{-S}\rangle \mid \langle\mathcal{J}_1\rangle\,B \mid \langle\mathcal{J}_2\rangle\,\langle\mathcal{L}_0^{-S}\rangle \mid \langle\mathcal{J}_3\rangle\,S \mid \\
&&& \quad \langle\mathcal{J}_4\rangle\,\langle\mathcal{L}_1^{-S}\rangle \mid \langle\mathcal{L}_1\rangle\,S \mid \langle\mathcal{J}_5\rangle\,(\,S\,S\,) \mid \langle\mathcal{J}_6\rangle\,\langle\mathcal{L}_0^{-S}\rangle \mid \langle\mathcal{J}_7\rangle\,S \mid \\
&&& \quad \langle\mathcal{J}_9\rangle\,S \mid \langle\mathcal{J}_{10}\rangle\,S \mid \langle\mathcal{L}_2^{-S}\rangle\,(\,S\,S\,) \mid S\,\langle\mathcal{L}_0^{-S}\rangle\,\langle\mathcal{N}^{-S}\rangle \\
&\text{(xxii)} & \langle\mathcal{J}_1\rangle &::= S\,S\,S \\
&\text{(xxiii)} & \langle\mathcal{J}_2\rangle &::= \langle\mathcal{K}_0\rangle\,S \mid S\,\langle\mathcal{H}_0^{-S}\rangle \mid S\,S\,\langle\mathcal{J}_2\rangle \\
&\text{(xxiv)} & \langle\mathcal{J}_3\rangle &::= \langle\mathcal{K}_0\rangle\,S \mid S\,\langle\mathcal{L}_4^{-S}\rangle \mid S\,S\,\langle\mathcal{J}_3\rangle \\
&\text{(xxv)} & \langle\mathcal{J}_4\rangle &::= S\,S\,(\,S\,S\,) \mid \langle\mathcal{J}_1\rangle\,S \\
&\text{(xxvi)} & \langle\mathcal{J}_5\rangle &::= S\,B\,(\,S\,S\,) \mid S\,S\,\langle\mathcal{J}_5\rangle \\
&\text{(xxvii)} & \langle\mathcal{J}_6\rangle &::= B\,(\,S\,S\,) \mid S\,S\,\langle\mathcal{J}_6\rangle \\
&\text{(xxviii)} & \langle\mathcal{J}_7\rangle &::= \langle\mathcal{K}_4\rangle\,S \mid B\,\langle\mathcal{L}_0^{-S}\rangle \mid S\,S\,\langle\mathcal{J}_7\rangle \\
&\text{(xxix)} & \langle\mathcal{J}_8\rangle &::= S\,\langle\mathcal{L}_0^{-S}\rangle \mid S\,S\,\langle\mathcal{J}_8\rangle \\
&\text{(xxx)} & \langle\mathcal{J}_9\rangle &::= \langle\mathcal{J}_8\rangle\,S \mid S\,S\,\langle\mathcal{J}_9\rangle
\end{aligned}$$

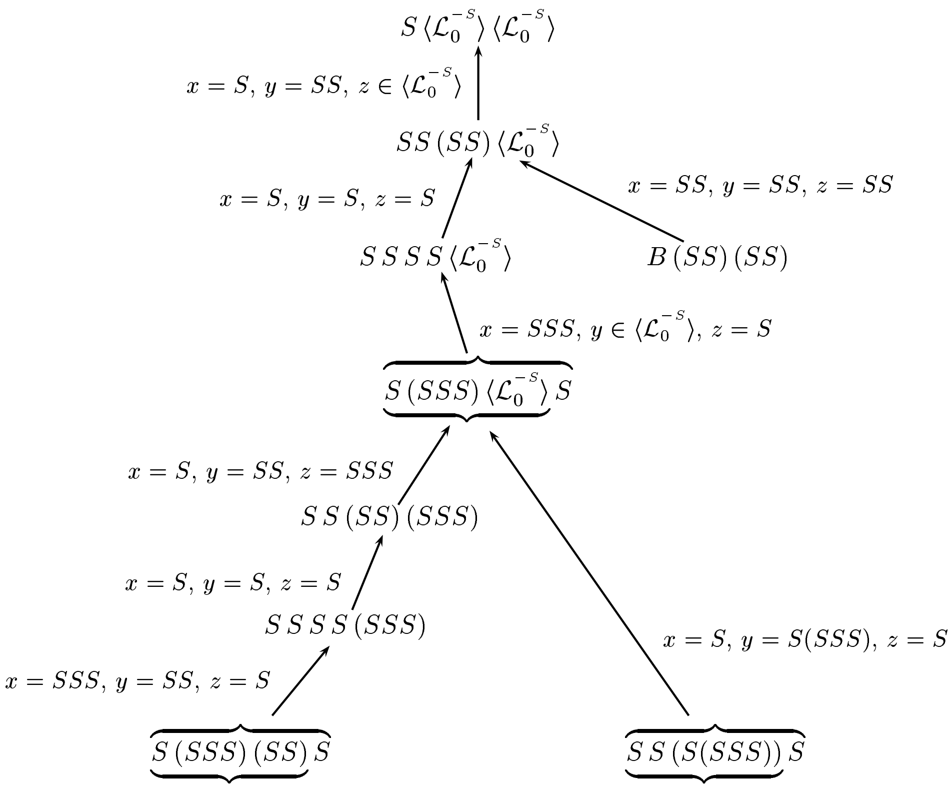
(xxxi) $\langle \mathcal{J}_{10} \rangle$ ::= $S\,S\,\langle \mathcal{J}_{10} \rangle \mid S\,\langle \mathcal{L}_0^{-S} \rangle\,\langle \mathcal{L}_0^{-S} \rangle \mid$
$SS\,(SS)\,\langle \mathcal{L}_0^{-S} \rangle \mid S\,S\,S\,S\,\langle \mathcal{L}_0^{-S} \rangle \mid B\,(SS)\,(SS) \mid \langle \mathcal{J}_{11} \rangle\,S$

(xxxii) $\langle \mathcal{J}_{11} \rangle$ ::= $S\,S\,\langle \mathcal{J}_{11} \rangle \mid S\,\langle \mathcal{K}_4 \rangle \mid$
$S\,(SSS)\,\langle \mathcal{L}_0^{-S} \rangle \mid S\,S\,(SS)\,(SSS) \mid S\,S\,S\,S\,(SSS)$

(xxxiii) $\langle \mathcal{J}_{12} \rangle$ ::= $S\,S\,\langle \mathcal{J}_{12} \rangle \mid S\,S\,(S(SSS)) \mid S\,(SSS)\,(SS)$

As we can see, for every right rule part of the form $X\,S$, where X is a non-terminal, the rule $X ::= SS\,X$ must be included in the set of rules for X.

The list of predecessors for $\mathcal{J}_{10}$ was obtained with an analysis like the one shown in the following diagram (the $(SS)^*[S\mathcal{K}_4S]$ case is not shown in the diagram; such terms are immediately put in $\langle \mathcal{J}_{11} \rangle$). Terms shown overbraced and underbraced in the diagram are at the boundaries between $\langle \mathcal{J}_{10} \rangle$ and $\langle \mathcal{J}_{11} \rangle$ ($S\,(SSS)\,\langle \mathcal{L}_0^{-S} \rangle\,S$ and $S\,(SSS)\,\langle \mathcal{L}_0^{-S} \rangle$ respectively) or between $\langle \mathcal{J}_{11} \rangle$ and $\langle \mathcal{J}_{12} \rangle$ ($S\,(SSS)\,(SS)\,S$ and $S\,(SSS)\,(SS)$ respectively, $S\,S\,(S(SSS))\,S$ and $S\,S\,(S(SSS))$ respectively).



9 Conclusion

We made a further step in lowering the complexity of the decision algorithm, by presenting a context-free grammar which fully characterizes all normalizing S-terms. Thus the complexity of deciding whether an S-term X has a normal form is $O(|X|^3)$, as given by the CYK algorithm.

Several proofs have been omitted for brevity; they will appear in a full version of this paper.

References

1. F. Baader and T. Nipkow. *Term Rewriting and All That.* Cambridge Univ. Press, 1998.
2. Henk P. Barendregt. *The Lambda Calculus, Its Syntax and Semantics.* North-Holland, Amsterdam, 1984.
3. Jan A. Bergstra and Jan Willem Klop. Conditional rewrite rules, confluence and termination. *JCSS*, 32(3):323–362, 1986.
4. Alonzo Church. The calculi of lambda conversion. *Annals of Math. Studies*, 6, 1941.
5. Haskell B. Curry and R. Feys. *Combinatory Logic.* North-Holland, Amsterdam, 1958.
6. Haskell B. Curry, J.Roger Hindley, and Jonathan P. Seldin. *Combinatory Logic II.* North-Holland, Amsterdam, 1972.
7. Nachum Dershowitz and Jean-Pierre Jouannaud. Rewrite systems. In *Handbook of Theoretical Computer Science*, volume B, pages 243–320. Elsevier, Amsterdam, 1990.
8. Gerard Huet and D. C. Oppen. Equations and rewrite rules: a survey. In *Formal Language Theory: Perspectives and Open Problems*, pages 349–405. Academic Press, London, 1990.
9. Jan Willem Klop. Term rewriting systems. In *Handbook of Logic in Computer Science*, volume 2, pages 2–116. Clarendon, 1992.
10. M. Schönfinkel. Über die Bausteine der mathematischen Logik. *Math. Annalen*, 92:305–316, 1924.
11. Dana S. Scott. Some philosophical issues concerning theories of combinators. *LNCS*, 37:346–366, 1975.
12. Raymond Smullyan. *To Mock a Mockingbird and Other Logic Puzzles Including an Amazing Adventure in Combinatory Logic.* Knopf, New York, 1985.
13. M. Sprenger and M. Wymann-Boeni. How to decide the lark. *Theoretical Computer Science*, 110:419–432, 1993.
14. Richard Statman. The word problem for Smullyan's lark combinator is decidable. *Journal of Symbolic Computation*, 7:103–112, 1989.
15. Johannes Waldmann. The combinator *S*. *Information and Computation*, 159:2–21, 2000.
16. Stathis Zachos. A decision algorithm for S-term normalization. In *PLS3:111-129*, Anogia, Crete, 2001.
17. Stathis Zachos, Stavros Routzounis, and Panos Hilaris. Deciding normalization and computing normal forms for S-terms. In *Proceedings 8th Panhellenic Conference in Informatics*, volume 1, pages 187–196, Nicosia, Cyprus, 2001.
18. Stathis (Efstathios) Zachos. *Kombinatorische Logik und S-Terme.* Dissertation, ETH Zürich, 1978.