Properties of Hereditary Substitution without Type Indices

Jason Reed

May 16, 2007

1 Introduction

One way of defining hereditary substitution in LF is to index the substitution operations by a type or simple type, so that they are manifestly terminating. One can alternatively omit the types from the *definition* of substitution, and prove that substitution terminates successfully on all well-typed terms. We show that there is no need to compromise between these approaches: even when such type decorations are absent, a reasonable definition of substitution still terminates (possibly in failure) on *all* inputs, whether well-typed or not. Moreover a suitable form of the associativity property

$$[M/x][N/y] = [[M/x]N/y][M/x] \qquad (y \notin FV(M))$$

for undecorated substitution can be similarly shown to hold for all terms, as long as all inner substitutions are defined.

2 Language

2.1 Syntax

Although we believe these results to be applicable to λ-terms defined in terms of canonical and atomic terms, the definitions are more convenient if we work in spine form. The syntax of terms is

$$\begin{array}{llll} \text{Terms} & M & ::= & x \cdot S \mid \lambda x.S \\ \text{Spines} & S & ::= & () \mid (M; S) \end{array}$$

We use V to stand uniformly for any expression, be it term or spine.

2.2 Typing

To get the inductive proofs to work, we do in fact engage in a certain sort of type discipline, but only a degenerately weak one, in the sense that *every* term

in the above language has a type. It is possible that it is equivalent to an intersection-typing system.

This notion of typing resembles the simple or 'skeletal' types that simply indicate the functional shape of a type, in that they are trees all of whose leaves are the single base type o. The grammar is as follows:

$$\begin{array}{lrcl} \text{Positive Types} & t & ::= & \{j_1, \ldots, j_n\} \to t \mid o \\ \text{Negative Types} & j & ::= & t \to j \mid o \end{array}$$

Note that these types involve finite sets $\{j_1, \ldots, j_n\}$ of 'negative' types j to be used to form 'positive' types t. We use the variable i to stand for such sets. These types are called 'positive' and 'negative' following the usual terminology of positive and negative positions in nested function types.

3 Syntactic Operations

3.1 Substitution

These definitions are essentially the standard ones with all type indices stripped off. Substitution $[M/y]V$ and reduction $[M|S]$ are partial functions, defined by the following clauses: (abbreviating $\sigma = [M/y]$)

$$\begin{array}{rcll} \sigma(\lambda x.N) & = & \lambda x.\sigma N & \\ \sigma(x \cdot S) & = & x \cdot \sigma S & (x \neq y) \\ \sigma(y \cdot S) & = & [M \mid \sigma S] & \\ \sigma() & = & () & \\ \sigma(N; S) & = & (\sigma N; \sigma S) & \\ [\lambda x.N \mid (M; S)] & = & [[M/x]N \mid S] & \\ [x \cdot S \mid ()] & = & x \cdot S & \end{array}$$

For any inputs that do not match the above patterns, these functions are undefined. One particularly important thing is that $[\lambda x.M \mid ()]$ fails to return, say, $\lambda x.M$, for otherwise Lemma 4.3 part 2 below is certainly false.

We write $X\downarrow$ to indicate that the computation implied by an expression X terminates. For instance, $[M/x]V\downarrow$ means 'either $[M/x]V$ exists, or finitely fails'. We write $X\Downarrow$ to indicate that X terminates successfully and outputs an answer. We write $X = Y\Downarrow$ to indicate that X and Y both terminate, and with the same answer.

3.2 Typing

Since every term is to have a type, we simply define three mutually recursive functions, $\mathbf{tp}(M), \mathbf{tp}(S), \mathbf{\{}x \in V\mathbf{\}}$, to directly compute the type of terms, spines, and variables.

The function $\mathbf{tp}(M)$ returns the positive type of a term M.

$$\begin{array}{rcl} \mathbf{tp}(x \cdot S) & = & o \\ \mathbf{tp}(\lambda x.M) & = & \mathbf{\{}x \in M\mathbf{\}} \to \mathbf{tp}(M) \end{array}$$

Given a spine S, the function $\mathbf{tp}(S)$ returns the 'type of S', that is, the negative type of a head that could be applied to S.

$$\begin{aligned}\mathbf{tp}(()) &= o\\ \mathbf{tp}((M;S)) &= \mathbf{tp}(M) \to \mathbf{tp}(S)\end{aligned}$$

Given an expression V, the function $\{x \in V\}$ returns the set of negative types that x 'needs to have' in V. For each occurrence of x in V, we look at the spine S it is applied to, and include the type of S in the set. We abbreviate $\{x \in V_1, \ldots, V_n\} = \{x \in V_1\} \cup \cdots \cup \{x \in V_n\}$.

$$\begin{aligned}\{x \in (x \cdot S)\} &= \{\mathbf{tp}(S)\} \cup \{x \in S\}\\ \{x \in (y \cdot S)\} &= \{x \in S\}\\ \{x \in (\lambda y.M)\} &= \{x \in M\}\\ \{x \in ()\} &= \{\}\\ \{x \in (M;S)\} &= \{x \in M, S\}\end{aligned}$$

We now define several relations and operations to express the induction measure for the proofs that follow. The relations $t \sqsubseteq t$ and $j \sqsubseteq j$ and $i \sqsubseteq i$ (all pronounced as 'hereditary subset of') are defined by

$$\frac{}{o \sqsubseteq o}$$

$$\frac{i \sqsubseteq i' \qquad t \sqsubseteq t'}{i \to t \sqsubseteq i' \to t'}$$

$$\frac{t \sqsubseteq t' \qquad j \sqsubseteq j'}{t \to j \sqsubseteq t' \to j'}$$

$$\frac{}{\{\} \sqsubseteq i}$$

$$\frac{i \sqsubseteq i' \qquad j \sqsubseteq j'}{i \cup \{j\} \sqsubseteq i' \cup \{j'\}}$$

Given a positive type t and a set i of negative types, consider a pair of these two items, written t/i. We use the variable p for these pairs generally. We define relations $\leq, <$ on these structures by

$$\frac{p \leq t/i}{p < (i \to t_0)/(\{t \to j\} \cup i_0)}$$

$$\frac{t \sqsubseteq t' \qquad i \sqsubseteq i'}{t/i \leq t'/i'} \qquad \frac{p < p'}{p \leq p'}$$

The notation $p_1 + p_2$ indicates a unordered simultaneous order on the structures p_1, p_2: $p_1 + p_2$ is considered equal to $p_2 + p_1$, and $p_1 + p_2$ is smaller than $p_1' + p_2'$ if either side of the former is smaller while the other remains the same, or if both get smaller. The operation $\cup$ binds more tightly than $/$, which binds more tightly than $+$.

4 Results

First some easy facts about $\sqsubseteq$.

Lemma 4.1 *$\sqsubseteq$ is a preorder, and the following rules are admissible:*

$$\frac{i_1 \sqsubseteq i_1' \qquad i_2 \sqsubseteq i_2'}{i_1 \cup i_2 \sqsubseteq i_1' \cup i_2'}$$

$$\frac{i \subseteq i'}{i \sqsubseteq i'}$$

Next is a result that formalizes what we need from ruling out $[\lambda x.M \mid ()] = \lambda x.M$.

Lemma 4.2 *Let x, y be two variables, possibly equal. If $[M/y](x \cdot S) = N$, then $\mathbf{tp}(N) = o$.*

Proof By induction on the derivation. ■

We then show that substitution and reduction are, in a suitable sense, non-increasing in the type of their arguments. This is arguably the most important (albeit also the most technical) lemma in this paper.

Lemma 4.3 *Abbreviate $\sigma = [M/y]$.*

1. *If $\sigma V \Downarrow$, then $\{x \in \sigma V\} \sqsubseteq \{x \in M, V\}$*
2. *If $\sigma V \Downarrow$, then $\mathbf{tp}(\sigma V) \sqsubseteq \mathbf{tp}(V)$.*
3. *If $[M|S] \Downarrow$, then $\{x \in [M|S]\} \sqsubseteq \{x \in M, S\}$.*

Proof By lexicographic induction. The measure that receives highest lexicographic priority for each branch is

1. $\mathbf{tp}(M)/\{y \in V\}$
2. $\mathbf{tp}(M)/\{y \in V\}$
3. $\mathbf{tp}(M)/\mathbf{tp}(S)$

Call this the *principal measure*. For the three branches of the lemma, say that branch 3 is considered smaller than 1 and 2, which are considered to be of the same size. Finally, if the principal measure and branch size both stay the same, we may proceed at lowest priority with smaller V.

1. Split cases on V.

   Case: $V = ()$. Immediate, since $\{\} \sqsubseteq \{x \in M\}$ by rule.

   Case: $V = (N; S)$. Compute

$$
\begin{aligned}
&\{x \in \sigma(N;S)\} \\
&= \{x \in (\sigma N; \sigma S)\} \\
&= \{x \in \sigma N, \sigma S\} \\
&\sqsubseteq \{x \in M, N, M, S\} && \text{by i.h. 1 twice} \\
&= \{x \in M, N, S\} && \text{properties of } \cup \\
&= \{x \in M, (N;S)\}
\end{aligned}
$$

The use of the induction hypothesis is licensed by the fact that the type $\{y \in N\}$ can be seen to be no larger than $\{y \in (N;S)\}$ just from inspecting definitions, and if it happens to be no smaller, then at least N is smaller than $(N;S)$.

Case: $V = \lambda z.N$. Compute

$$
\begin{aligned}
&\{x \in \sigma(\lambda z.N)\} \\
&= \{x \in \lambda z.\sigma N\} \\
&= \{x \in \sigma N\} \\
&\sqsubseteq \{x \in M, N\} && \text{by i.h. 1} \\
&= \{x \in M, \lambda z.N\}
\end{aligned}
$$

Case: $V = z \cdot S$ where $z \neq x$ and $z \neq y$. Compute

$$
\begin{aligned}
&\{x \in \sigma(z \cdot S)\} \\
&= \{x \in z \cdot \sigma S\} \\
&= \{x \in \sigma S\} \\
&\sqsubseteq \{x \in M, S\} && \text{by i.h. 1} \\
&= \{x \in M, z \cdot S\}
\end{aligned}
$$

Case: $V = x \cdot S$. Compute

$$
\begin{aligned}
&\{x \in \sigma(x \cdot S)\} \\
&= \{x \in x \cdot \sigma S\} \\
&= \{\mathbf{tp}(\sigma S)\} \cup \{x \in \sigma S\} \\
&\sqsubseteq \{\mathbf{tp}(\sigma S)\} \cup \{x \in M, S\} && \text{by i.h. 1} \\
&\sqsubseteq \{\mathbf{tp}(S)\} \cup \{x \in M, S\} && \text{by i.h. 2} \\
&= \{x \in M\} \cup (\mathbf{tp}(S) \cup \{x \in S\}) && \text{properties of } \cup \\
&= \{x \in M, x \cdot S\}
\end{aligned}
$$

For both appeals to the induction hypothesis, note that the principal measure may stays the same (at $\mathbf{tp}(M)/\{y \in x \cdot S\} = \mathbf{tp}(M)/\{y \in S\}$) and the branch size stays the same, but the size of the pertinent V nonetheless shrinks from $x \cdot S$ to S, and so the appeal is justified.

Case: $V = y \cdot S$. The principal measure for this case is

$$\mathbf{tp}(M)/\{y \in y \cdot S\} = \mathbf{tp}(M)/(\{\mathbf{tp}(S)\} \cup \{y \in S\})$$

We first invoke the induction hypothesis branch 2 to see that

$$\mathbf{tp}(\sigma S) \sqsubseteq \mathbf{tp}(S) \qquad (*)$$

The principal measure for this appeal is $\mathbf{tp}(M)/\{y \in S\}$, which is no larger, but may be equal to the one we started with if already

$\mathbf{tp}(S) \in \mathbf{\{}y \in S\mathbf{\}}$. However, if it is equal, then we are still able to proceed because S is smaller than $y \cdot S$. This same reasoning justifies the appeal to i.h. 1 below. From $(*)$ we infer easily that

$$\mathbf{tp}(M)/\mathbf{tp}(\sigma S) \leq \mathbf{tp}(M)/\{\mathbf{tp}(S)\} \cup \mathbf{\{}y \in S\mathbf{\}} \qquad (**)$$

Now compute

$$\begin{array}{ll}
\mathbf{\{}x \in \sigma(y \cdot S)\mathbf{\}} & \\
= \mathbf{\{}x \in [M \mid \sigma S]\mathbf{\}} & \\
\sqsubseteq \mathbf{\{}x \in M, \sigma S\mathbf{\}} & \text{by i.h. 3, licensed by } (**) \\
\sqsubseteq \mathbf{\{}x \in M, M, S\mathbf{\}} & \text{by i.h. 1} \\
= \mathbf{\{}x \in M, S\mathbf{\}} & \text{properties of } \cup \\
= \mathbf{\{}x \in M, y \cdot S\mathbf{\}} &
\end{array}$$

2. Split cases on V.

Case: $V = ()$. Immediate.

Case: $V = (N; S)$.

$$\begin{array}{ll}
\mathbf{tp}(\sigma(N; S)) & \\
= \mathbf{tp}((\sigma N; \sigma S)) & \\
= \mathbf{tp}(\sigma N) \to \mathbf{tp}(\sigma S) & \\
\sqsubseteq \mathbf{tp}(N) \to \mathbf{tp}(S) & \text{i.h. 2 twice} \\
= \mathbf{tp}(N; S) &
\end{array}$$

Case: $V = \lambda x.N$. Compute

$$\begin{array}{ll}
\mathbf{tp}(\sigma(\lambda x.N)) & \\
= \mathbf{tp}(\lambda x.\sigma N) & \\
= \mathbf{\{}x \in \sigma N\mathbf{\}} \to \mathbf{tp}(\sigma N) & \\
\sqsubseteq \mathbf{\{}x \in \sigma N\mathbf{\}} \to \mathbf{tp}(N) & \text{i.h. 2} \\
\sqsubseteq \mathbf{\{}x \in N\mathbf{\}} \to \mathbf{tp}(N) & \text{i.h. 1} \\
= \mathbf{tp}(\lambda x.N) &
\end{array}$$

Both appeals to the induction hypothesis keep the principal measure and the branch size constant, and decrease the size of the expression V.

Case: $V = x \cdot S$. (regardless of whether $x = y$ or $x \neq y$) Use Lemma 4.2, and note that $o \sqsubseteq o$.

3. Split cases on $\mathbf{tp}(M)$.

Case: $\mathbf{tp}(M) = o$. Then M is of the form $y \cdot S'$ for some variable y (which may in fact be x) and S must be $()$ for $[M|S]$ to be defined. All that remains to notice is

$$\begin{array}{l}
\mathbf{\{}x \in [M|S]\mathbf{\}} \\
= \mathbf{\{}x \in [y \cdot S' \mid ()]\mathbf{\}} \\
= \mathbf{\{}x \in y \cdot S'\mathbf{\}}
\end{array}$$

$= \{x \in y \cdot S', ()\}$
$= \{x \in M, S\}$

Case: $\mathbf{tp}(M) = i \to t$. Then M is of the form $\lambda y.M_0$ such that $\{y \in M_0\} = i$ and $\mathbf{tp}(M_0) = t$. Moreoever S must be of the form $(M'; S')$ for $[M|S]$ to be defined. The principal measure at this case is

$\mathbf{tp}(\lambda y.M_0)/\mathbf{tp}((M'; S'))$
$= (\{y \in M_0\} \to \mathbf{tp}(M_0))/(\mathbf{tp}(M') \to \mathbf{tp}(S'))$
$> \{y \in M_0\}/\mathbf{tp}(M')$
$= \mathbf{tp}(M')/\{y \in M_0\}$
$\therefore \mathbf{tp}(M')/\{y \in M_0\} < \mathbf{tp}(\lambda y.M_0)/\mathbf{tp}((M'; S'))$ (*)

so we are justified in using the induction hypothesis branch 2 to conclude

$$\mathbf{tp}([M'/y]M_0) \sqsubseteq \mathbf{tp}(M_0)$$

From this we can deduce

$\mathbf{tp}([M'/y]M_0)/\mathbf{tp}(S')$
$\leq \mathbf{tp}(M_0)/\mathbf{tp}(S')$
$< (\{y \in M_0\} \to \mathbf{tp}(M_0))/(\mathbf{tp}(M') \to \mathbf{tp}(S'))$
$= \mathbf{tp}(\lambda y.M_0)/\mathbf{tp}((M'; S'))$
$\therefore \mathbf{tp}([M'/y]M_0)/\mathbf{tp}(S') < \mathbf{tp}(\lambda y.M_0)/\mathbf{tp}((M'; S'))$ (**)

Now compute

$\{x \in [M|S]\}$
$= \{x \in [\lambda y.M_0 \mid (M'; S')]\}$
$= \{x \in [[M'/y]M_0 \mid S']\}$
$\sqsubseteq \{x \in [M'/y]M_0, S'\}$ i.h. 3, licensed by (**)
$\sqsubseteq \{x \in M', M_0, S'\}$ i.h. 1, licensed by (*)
$= \{x \in M, S\}$

■

Theorem 4.4 (Termination)

1. $[M/x]V\downarrow$
2. $[M|S]\downarrow$

Proof By lexicographic induction. The principal measure is

1. $\mathbf{tp}(M)/\{x \in V\}$
2. $\mathbf{tp}(M)/\mathbf{tp}(S)$

For equal values of this measure, branch 2 is considered smaller. For equal principal measure and branch size, we may proceed with smaller V.

1. Split cases on V.

   Case: $V = ()$. Immediate.

   Case: $V = (N; S)$. Apply induction hypothesis to N and S, at the same (or possibly smaller) measure but smaller terms.

   Case: $V = \lambda y.N$. Apply induction hypothesis to N, at the same measure but a smaller term.

   Case: $V = y \cdot S$. Apply induction hypothesis to S, at the same measure but a smaller expression.

   Case: $V = x \cdot S$. Immediately we can see that

   $$[M/x]S\downarrow \qquad \text{i.h. 1}$$

   by applying the induction at the same (or possibly smaller) principal measure for the smaller expression S. If $[M/x]S$ fails, then we are already done, for $[M/x](x \cdot S) = [M \mid [M/x]S]$ has already failed. Otherwise, reason as follows:

   $$\begin{array}{ll} \mathbf{tp}([M/x]S) \sqsubseteq \mathbf{tp}(S) & \text{Lemma 4.3} \\ \mathbf{tp}(M)/\mathbf{tp}([M/x]S) \leq \mathbf{tp}(M)/\mathbf{tp}(S) & \\ \leq \mathbf{tp}(M)/\{\mathbf{tp}(S)\} \cup \mathbf{\{}x \in S\mathbf{\}} & \\ = \mathbf{tp}(M)/\mathbf{\{}x \in x \cdot S\mathbf{\}} & \\ \therefore \mathbf{tp}(M)/\mathbf{tp}([M/x]S) \leq \mathbf{tp}(M)/\mathbf{\{}x \in x \cdot S\mathbf{\}} & \end{array}$$

   Thus we may appeal to i.h. 2 to see that $[M \mid [M/x]S]$ either exists or finitely fails.

2. Split cases on $\mathbf{tp}(M)$.

   Case: $\mathbf{tp}(M) = o$. Then M is of the form $y \cdot S'$. If $S = ()$, then $[M \mid S] = M$. Otherwise, reduction immediately fails.

   Case: $\mathbf{tp}(M) = i \to t$. Then M is of the form $\lambda y.N$. Consider whether S is of the form $(M_0; S_0)$. If it is not, then reduction immediately fails. If it is, note that the principal measure for this case is

   $$\mathbf{tp}(\lambda y.N)/\mathbf{tp}(M_0; S_0) = (\mathbf{\{}y \in N\mathbf{\}} \to \mathbf{tp}(N))/(\mathbf{tp}(M_0) \to \mathbf{tp}(S_0))$$

   Observe also that

   $$\mathbf{\{}y \in N\mathbf{\}}/\mathbf{tp}(M_0) < (\mathbf{\{}y \in N\mathbf{\}} \to \mathbf{tp}(N))/(\mathbf{tp}(M_0) \to \mathbf{tp}(S_0))$$

   which licenses using i.h. 1 to conclude $[M_0/y]N\downarrow$. It it fails, then $[\lambda y.N \mid (M_0; S_0)] = [[M_0/y]N \mid S_0]$ also fails, and we are done. Otherwise, reason that

$\mathbf{tp}([M_0/y]N) \sqsubseteq \mathbf{tp}(N)$ Lemma 4.3
$\mathbf{tp}([M_0/y]N)/\mathbf{tp}(S_0) \leq \mathbf{tp}(N)/\mathbf{tp}(S_0)$
$< (\{y \in N\} \to \mathbf{tp}(N))/(\mathbf{tp}(M_0) \to \mathbf{tp}(S_0))$
$= \mathbf{tp}(\lambda y.N)/\mathbf{tp}(M_0; S_0)$
$\therefore \mathbf{tp}([M_0/y]N)/\mathbf{tp}(S_0) < \mathbf{tp}(\lambda y.N)/\mathbf{tp}(M_0; S_0)$ $(*)$
$[[M_0/y]N \mid S_0]\downarrow$ i.h. 2 using $(*)$

■

Lemma 4.5 *If $x \notin FV(N)$, then $[M/x]N = N$ and $\{x \in N\} = \{\}$.*

Proof By induction on N. ■

Theorem 4.6 (Substitution Interchange) *Let M, N, V and S be given such that $x \notin FV(N)$. Abbreviate $\sigma = [N/y]$.*

1. *If $\sigma M\Downarrow$, $\sigma V\Downarrow$, and $[M/x]V\Downarrow$, then $\sigma[M/x]V = [\sigma M/x]\sigma V\Downarrow$.*
2. *If $\sigma M\Downarrow$, $\sigma S\Downarrow$, and $[M \mid S]\Downarrow$, then $\sigma[M \mid S] = [\sigma M \mid \sigma S]\Downarrow$.*

Proof By lexicographic induction. The principal measure is

1. $\mathbf{tp}(M)/\{x \in V\} + \mathbf{tp}(N)/\{y \in M, V\}$
2. $\mathbf{tp}(M)/\mathbf{tp}(S) + \mathbf{tp}(N)/\{y \in M, S\}$

and for equal values of it, case 2 is considered smaller, and we may proceed with smaller V within case 1. We show the most interesting cases.

1. Split cases on V.

   Case: $V = x \cdot S$. The principal measure here, call it μ, is

   $$\mu = \mathbf{tp}(M)/\{x \in x \cdot S\} + \mathbf{tp}(N)/\{y \in M, x \cdot S\}$$
   $$= \mathbf{tp}(M)/(\{\mathbf{tp}(S)\} \cup \{x \in S\}) + \mathbf{tp}(N)/\{y \in M, x \cdot S\}$$

   By assumption, we know $\sigma M\Downarrow, \sigma S\Downarrow, [M/x]S\Downarrow, [M \mid [M/x]S]\Downarrow$. By i.h. 1 at measure

   $$\mathbf{tp}(M)/\{x \in S\} + \mathbf{tp}(N)/\{y \in M, S\} \leq \mu$$

   (and smaller term S) we see

   $$\sigma[M/x]S = [\sigma M/x]\sigma S\Downarrow \qquad (*)$$

   We can reason that

   $\mathbf{tp}([M/x]S) \sqsubseteq \mathbf{tp}(S)$ Lemma 4.3
   $\{y \in [M/x]S\} \sqsubseteq \{y \in M, S\}$ Lemma 4.3
   $\therefore \mathbf{tp}(M)/\{\mathbf{tp}([M/x]S)\} + \mathbf{tp}(N)/\{y \in M, [M/x]S\} \leq \mu$

which licenses using i.h. 2 to conclude

$$\sigma[M \mid [M/x]S] = [\sigma M \mid \sigma[M/x]S]\Downarrow \qquad (**)$$

We can now calculate

$\sigma[M/x](x \cdot S) = \sigma[M \mid [M/x]S]$
$= [\sigma M \mid \sigma[M/x]S]$ by $(**)$
$= [\sigma M \mid [\sigma M/x]\sigma S]$ by $(*)$
$= [\sigma M/x](x \cdot \sigma S)$
$= [\sigma M/x]\sigma(x \cdot S)$

Case: $V = y \cdot S$. The principal measure μ here is

$$\mu = \mathbf{tp}(M)/\{\!|x \in y \cdot S|\!\} + \mathbf{tp}(N)/\{\!|y \in M, y \cdot S|\!\}$$

$$= \mathbf{tp}(M)/\{\!|x \in S|\!\} + \mathbf{tp}(N)/(\{\mathbf{tp}(S)\} \cup \{\!|y \in M, S|\!\})$$

By assumption, we know $\sigma M\Downarrow, [N/y]S\Downarrow, [N \mid [N/y]S]\Downarrow, [M/x]S\Downarrow$. By i.h. 1 at measure

$$\mathbf{tp}(M)/\{\!|x \in S|\!\} + \mathbf{tp}(N)/\{\!|y \in M, S|\!\} \leq \mu$$

(and smaller term S) we see

$$\sigma[M/x]S = [\sigma M/x]\sigma S\Downarrow \qquad (*)$$

We can reason that

$\mathbf{tp}(\sigma M) \sqsubseteq \mathbf{tp}(M)$ Lemma 4.3
$\mathbf{tp}(\sigma S) \sqsubseteq \mathbf{tp}(S)$ Lemma 4.3
$\{\!|x \in N|\!\} = \{\}$ Lemma 4.5
$\{\!|x \in \sigma S|\!\} \sqsubseteq \{\!|x \in N|\!\} \cup \{\!|x \in S|\!\}$ Lemma 4.3
$= \{\!|x \in S|\!\}$
$\therefore \mathbf{tp}(N)/\{\mathbf{tp}(\sigma S)\} + \mathbf{tp}(\sigma M)/\{\!|x \in N, \sigma S|\!\} \leq \mu$

which licenses using i.h. 2 to conclude

$$[\sigma M/x][N \mid \sigma S] = [[\sigma M/x]N \mid [\sigma M/x]\sigma S]\Downarrow \qquad (**)$$

We can now calculate

$\sigma[M/x](y \cdot S) = \sigma(y \cdot [M/x]S)$
$= [N \mid \sigma[M/x]S]$
$= [N \mid [\sigma M/x]\sigma S]$ by (*)
$= [[\sigma M/x]N \mid [\sigma M/x]\sigma S]$ Lemma 4.5
$= [\sigma M/x][N \mid \sigma S]$ by (**)
$= [\sigma M/x]\sigma(y \cdot S)$

2. Split cases on $\mathbf{tp}(M)$.

Case: $\mathbf{tp}(M) = o$. M must be of the form $x \cdot S$, and S must be of the form $()$ because $[M \mid S]\Downarrow$. On the one hand, $\sigma[M|S] = \sigma M\Downarrow$. But by Lemma 4.2, σM is not a lambda, so $[\sigma M \mid \sigma S] = [\sigma M \mid ()] = \sigma M\Downarrow$.

Case: $\mathbf{tp}(M) = i \to t$. We know that $[M \mid S]\Downarrow$, so M must be of the form $\lambda x.M_0$ and S must be of the form $(M'; S')$. The principal measure μ here is

$$\begin{aligned}
&\mu = \mathbf{tp}(M)/\{\mathbf{tp}(S)\} + \mathbf{tp}(N)/\mathbf{\{}y \in M, S\mathbf{\}} \\
&= \mathbf{tp}(\lambda x.M_0)/\{\mathbf{tp}((M'; S'))\} + \mathbf{tp}(N)/\mathbf{\{}y \in \lambda x.M_0, (M'; S')\mathbf{\}} \\
&= (\mathbf{\{}x \in M_0\mathbf{\}} \to \mathbf{tp}(M_0))/\{\mathbf{tp}(M') \to \mathbf{tp}(S')\} + \mathbf{tp}(N)/\mathbf{\{}y \in M_0, M', S'\mathbf{\}}
\end{aligned}$$

We can reason that

$$\begin{aligned}
&\mathbf{tp}([M'/x]M_0) \sqsubseteq \mathbf{tp}(M_0) && \text{Lemma 4.3} \\
&\mathbf{tp}([M'/x]M_0) < (\mathbf{\{}x \in M_0\mathbf{\}} \to \mathbf{tp}(M_0)) \\
&\mathbf{tp}(S') < \mathbf{tp}(M') \to \mathbf{tp}(S') \\
&\mathbf{\{}y \in [M'/x]M_0\mathbf{\}} \sqsubseteq \mathbf{\{}y \in M', M_0\mathbf{\}} && \text{Lemma 4.3} \\
&\therefore \mathbf{tp}([M'/x]M_0)/\{\mathbf{tp}(S')\} + \mathbf{tp}(N)/\mathbf{\{}y \in [M'/x]M_0, S'\mathbf{\}} < \mu
\end{aligned}$$

which licenses using i.h. 2 to conclude

$$\sigma[[M'/x]M_0 \mid S'] = [\sigma[M'/x]M_0 \mid \sigma S']\Downarrow \qquad (*)$$

And we can see that

$$\begin{aligned}
&\mathbf{tp}(M') < \mathbf{tp}(M') \to \mathbf{tp}(S') \\
&\mathbf{\{}x \in M_0\mathbf{\}} < \mathbf{\{}x \in M_0\mathbf{\}} \to \mathbf{tp}(M_0) \\
&\therefore \mathbf{tp}(M')/\mathbf{\{}x \in M_0\mathbf{\}} + \mathbf{tp}(N)/\mathbf{\{}y \in M', M_0\mathbf{\}} < \mu
\end{aligned}$$

which licenses using i.h. 1 to conclude

$$\sigma[M'/x]M_0 = [\sigma M'/x]\sigma M_0\Downarrow \qquad (**)$$

We can now calculate

$$\begin{aligned}
&\sigma[\lambda x.M_0 \mid (M'; S')] \\
&= \sigma[[M'/x]M_0 \mid S'] \\
&= [\sigma[M'/x]M_0 \mid \sigma S'] && \text{by } (*) \\
&= [[\sigma M'/x]\sigma M_0 \mid \sigma S'] && \text{by } (**) \\
&= [\lambda x.\sigma M_0 \mid (\sigma M'; \sigma S')] \\
&= [\sigma(\lambda x.M_0) \mid \sigma(M'; S')]
\end{aligned}$$

■