UPPSALA
UNIVERSITET

U.U.D.M. Project Report 2013:12

A Syntax of the Simple Theory of Types

Mattias Granberg Olsson

Examensarbete i matematik, 15 hp
Handledare och examinator: Inger Sigstam
Juni 2013

Department of Mathematics
Uppsala University

A formal deductive system, the version of the simple theory of types formulated by Church in [2], is presented, formalised within a theory of strings powerful enough to define recursive structures of strings. These are shown to be indeed recursive, and are applied to define the notions of free and bound variables and substitutions of such. The behaviour of the language and inference relation under these substitutions is investigated, and used to generalise the rules of inference. Finally, the Deduction Theorem is proven.

Contents

Introduction

The origins of this thesis was a desire to better understand how the recursive structure that is mathematics is grounded, that is, the ultimately philosophical question whether anything can be seen as the foundation of everything. In set theory, for instance, seemingly innocent statements like “for every set A and every unary definite condition φ there is a set B containing exactly those elements of A that satisfy φ” abound, and makes one wonder what is a “unary definite condition”, and how such a thing as unary can have any meaning when (if we strive to use set theory as the foundation of mathematics) we have yet to define the natural numbers. This can of course be remedied by working within a meta theory, but said meta theory would then give rise to the same kinds of issues, marking but another recursive call.

However, set theory has not been the only attempt at a foundations of all of mathematics, the various incarnations of the equally various theories of types have also played this role, for instance. This thesis concerns the simple theory of types as it was formulated by Church in [2], though not as a foundational system but as a purely formal theory. This system (and, as far as I am aware, most type theories), makes precise the intuitive idea that objects, functions and propositions constitute essentially distinct entities, while at the same time sharing many properties. In his paper Church established the fundamental properties of this system, most notably the Deduction Theorem, and showed that elementary number theory could be developed within it. Later, Henkin (in [3]) showed that this system, with a specific notion of model, is actually complete.

Originally, the intention was that of summarising these results, but due to time, space and administrative constraints, this thesis now concerns itself with clarifying the work of Church ([2]). Likewise, while the strengths and weaknesses of this theory relative other theories would most certainly make an interesting study, unfortunately these as well fall outside the scope of this thesis. That being said, the presentation will be rather technical, with much attention to detail. Furthermore, for convenience and in the hopes that the results obtained herein will rest upon solid foundations, the meta theory will be ZFC.

The text is divided into four sections, two of which concern the simple theory of types as a formal theory. These sections, 2 and 3, are principally based upon Church’s formulation of the theory in [2], with a few additions and alterations introduced by (and thus due to) Henkin in [3]. In these sections, any reference to “type theory” will refer to the simple theory of types as defined therein. The remaining sections are the current, containing an introduction in the form of motivation, background and acknowledgements, and section 1 hereafter, which contain some preliminary results about strings. The latter is loosely based on the material on strings found in [4].

Some Notational Remarks

For any $n, m \in \mathbb{N} \cup \{\omega\}$ we will use the following notations:

$$[n, m] = \{k \in \mathbb{N} \cup \{\omega\} \mid n \leq k \leq m\}$$
$$[n, m[= \{k \in \mathbb{N} \cup \{\omega\} \mid n \leq k < m\}$$
$$]n, m[= \{k \in \mathbb{N} \cup \{\omega\} \mid n < k < m\}$$
$$]n, m] = \{k \in \mathbb{N} \cup \{\omega\} \mid n < k \leq m\}.$$

We will also use the abbreviations $\neg$, $\bigwedge$, $\bigvee$, $\forall$, $\exists$, $\Rightarrow$, $\Leftarrow$ and $\Leftrightarrow$ in the meta language, corresponding to the intended meaning of the formal symbols $\neg$, $\wedge$, $\vee$, $\forall$, $\exists$, $\rightarrow$, $\leftarrow$ and $\leftrightarrow$ in the obvious and usual fashion. Furthermore we will use many of the usual conventions, that most readers will probably be familiar with, as well as the extra conventions that strings will be denoted by a, b, c, d, e, f, recursive strings and formulas by u, v, w, q, type symbols by lowercase Greek letters at the beginning of the alphabet, well formed formulas by lowercase Greek letters and sets of well formed formulas by uppercase Greek letters. Furthermore $\subset$ will mean strict and $\subseteq$ non-strict set containment. A statement like $a, b \in A$ (such as the one in the beginning of this section) is to be interpreted as $a \in A$ and $b \in A$. Concerning sequences and indexed families (finite or not), we will use the convention (as opposed to strings) of treating them as sets, that is we will write $a \in \{a_k\}_{k \in I} \subseteq A$ when we wish to convey that $a = a_k$ for some $k \in I$ and $a_k \in A$ for all $k \in I$. However, should the a_k come from different sets A_k, we will write $\{a_k\}_{k \in I} \in \prod_{k \in I} A_k$ as usual. For finite sequences (with a meta theoretically well-defined arity) we will not distinguish between sequence and tuple notation, e.g. $\{k\}_{k=1}^4 = (1, 2, 3, 4)$.

Acknowledgements

I wish to express my gratitude to my supervisor Inger Sigstam, for her involvement, suggestions and patience. I also wish to thank my fellow student Tilo Wiklund, who read and commented on an early and a late draft of this paper.

1 Preliminaries: Strings

Throughout this section, let A be a set, called an *alphabet*, whose elements will be called *symbols*.

1.1 Definition (Strings)**.** Let $n \in \mathbb{N}$. A *string of length* n of symbols of A, or a string of length n with entries in A, is a function $a : [0, n[\longrightarrow A$. The set of all such strings of length n is denoted A^n. Furthermore we call $A^* = \bigcup_{n \in \mathbb{N}} A^n$ the set of all *strings* of symbols of A.

Note that all strings have finite length.

Remark 1. The only element of A^0 is called the *empty string* and is denoted by $\diamond$. For any symbol s, the string $a \in A^1$ such that

$$a(0) = s$$

is called a *single character string* and is often denoted by s. Hence there is a canonical injection of A into A^*.

1.2 Definition. Let $\mathrm{lh} : A^* \longrightarrow \mathbb{N}$ be the *length function*, so that for all $a \in A^*$ we have $a \in A^{\mathrm{lh}(a)}$.

1.3 Definition (Concatenation)**.** Let $\mathcal{C} : A^* \times A^* \longrightarrow A^*$ be the function such that for all $a, b \in A^*$ we have that

$$\mathcal{C}(a, b) \in A^{\mathrm{lh}(a)+\mathrm{lh}(b)}$$

and for all $k \in [0, \mathrm{lh}(a) + \mathrm{lh}(b)[$

$$\mathcal{C}(a, b)(k) = \begin{cases} a(k) & \text{if } k < \mathrm{lh}(a) \\ b(k - \mathrm{lh}(a)) & \text{if } k \geq \mathrm{lh}(a) \end{cases}$$

1.4 Proposition (Associativity of concatenation)**.** *For all* $a, b, c \in A^*$ *we have that*

$$\mathcal{C}(a, \mathcal{C}(b, c)) = \mathcal{C}(\mathcal{C}(a, b), c)$$

Proof. Let $a, b, c \in A^*$. We then have that

$$\begin{aligned} \mathrm{lh}(\mathcal{C}(a, \mathcal{C}(b, c))) &= \mathrm{lh}(a) + \mathrm{lh}(\mathcal{C}(b, c)) \\ &= \mathrm{lh}(a) + \mathrm{lh}(b) + \mathrm{lh}(c) \\ &= \mathrm{lh}(\mathcal{C}(a, b)) + \mathrm{lh}(c) \\ &= \mathrm{lh}(\mathcal{C}(\mathcal{C}(a, b), c)). \end{aligned}$$

Now let $k < \mathrm{lh}(a) + \mathrm{lh}(b) + \mathrm{lh}(c)$. If $k < \mathrm{lh}(a)$ then $k < \mathrm{lh}(a) + \mathrm{lh}(b) = \mathrm{lh}(\mathcal{C}(a, b))$, and hence

$$\mathcal{C}(\mathcal{C}(a, b), c)(k) = \mathcal{C}(a, b)(k) = a(k) = \mathcal{C}(a, \mathcal{C}(b, c))(k)$$

If instead $\mathrm{lh}(\mathfrak{a}) \le \mathsf{k} < \mathrm{lh}(\mathfrak{a}) + \mathrm{lh}(\mathfrak{b}) = \mathrm{lh}(\mathcal{C}(\mathfrak{a},\mathfrak{b}))$, then $\mathsf{k} - \mathrm{lh}(\mathfrak{a}) < \mathrm{lh}(\mathfrak{b})$ and

$$\begin{aligned}\mathcal{C}(\mathfrak{a},\mathcal{C}(\mathfrak{b},\mathfrak{c}))(\mathsf{k}) &= \mathcal{C}(\mathfrak{b},\mathfrak{c})(\mathsf{k} - \mathrm{lh}(\mathfrak{a}))\\ &= \mathfrak{b}(\mathsf{k} - \mathrm{lh}(\mathfrak{a}))\\ &= \mathcal{C}(\mathfrak{a},\mathfrak{b})(\mathsf{k})\\ &= \mathcal{C}(\mathcal{C}(\mathfrak{a},\mathfrak{b}),\mathfrak{c})(\mathsf{k}).\end{aligned}$$

Finally, if $\mathrm{lh}(\mathfrak{a})+\mathrm{lh}(\mathfrak{b}) \le \mathsf{k} < \mathrm{lh}(\mathfrak{a})+\mathrm{lh}(\mathfrak{b})+\mathrm{lh}(\mathfrak{c})$ then $\mathrm{lh}(\mathfrak{b}) \le \mathsf{k}-\mathrm{lh}(\mathfrak{a}) < \mathrm{lh}(\mathfrak{b})+\mathrm{lh}(\mathfrak{c}) = \mathrm{lh}(\mathcal{C}(\mathfrak{b},\mathfrak{c}))$ and $\mathsf{k} - \mathrm{lh}(\mathcal{C}(\mathfrak{a},\mathfrak{b})) = \mathsf{k} - (\mathrm{lh}(\mathfrak{a}) + \mathrm{lh}(\mathfrak{b})) < \mathrm{lh}(\mathfrak{c})$. Thus

$$\begin{aligned}\mathcal{C}(\mathfrak{a},\mathcal{C}(\mathfrak{b},\mathfrak{c}))(\mathsf{k}) &= \mathcal{C}(\mathfrak{b},\mathfrak{c})(\mathsf{k} - \mathrm{lh}(\mathfrak{a}))\\ &= \mathfrak{c}(\mathsf{k} - \mathrm{lh}(\mathfrak{a}) - \mathrm{lh}(\mathfrak{b}))\\ &= \mathfrak{c}(\mathsf{k} - (\mathrm{lh}(\mathfrak{a}) + \mathrm{lh}(\mathfrak{b})))\\ &= \mathfrak{c}(\mathsf{k} - \mathrm{lh}(\mathcal{C}(\mathfrak{a},\mathfrak{b})))\\ &= \mathcal{C}(\mathcal{C}(\mathfrak{a},\mathfrak{b}),\mathfrak{c})(\mathsf{k}).\end{aligned}$$

Hence $\mathcal{C}(\mathcal{C}(\mathfrak{a},\mathfrak{b}),\mathfrak{c})(\mathsf{k}) = \mathcal{C}(\mathfrak{a},\mathcal{C}(\mathfrak{b},\mathfrak{c}))$ for every $\mathsf{k} \in [0, \mathrm{lh}(\mathfrak{a}) + \mathrm{lh}(\mathfrak{b}) + \mathrm{lh}(\mathfrak{c})[$. □

Remark 2. As concatenation is associative, we will write $\mathfrak{ab}$ instead of $\mathcal{C}(\mathfrak{a},\mathfrak{b})$ for $\mathfrak{a},\mathfrak{b} \in \mathsf{A}^*$. Hence, if $\mathfrak{n} \in \mathbb{Z}^+$ and $\mathsf{s}_1, \mathsf{s}_2, \dots, \mathsf{s}_\mathfrak{n} \in \mathsf{A}$ is a sequence of symbols, we will by $\mathsf{s}_1\mathsf{s}_2 \dots \mathsf{s}_\mathfrak{n}$ mean the string $\mathfrak{a} \in \mathsf{A}^\mathfrak{n}$ such that $\mathfrak{a}(\mathsf{k}) = \mathsf{s}_\mathsf{k}$. This is also motivated by the fact that it is possible to prove string properties inductively, and define string operations recursively, by considering strings as being built up from shorter strings and symbols. The latter is the claim of the following theorem, the proof of which can be found in [4].

1.5 Theorem (String recursion)**.** *Given an alphabet* A*, let* E *be a set,* $\mathsf{e} \in \mathsf{E}$ *and* $\mathsf{h} : \mathsf{E} \times \mathsf{A} \longrightarrow \mathsf{E}$*. There is a unique function* $\mathsf{f} : \mathsf{A}^* \longrightarrow \mathsf{E}$ *such that*

$$\begin{aligned}\mathsf{f}(\diamond) &= \mathsf{e}\\ \mathsf{f}(\mathfrak{a}\mathsf{t}) &= \mathsf{h}(\mathsf{f}(\mathfrak{a}), \mathsf{t})\end{aligned}$$

for all $\mathfrak{a} \in \mathsf{A}^*$*,* $\mathsf{t} \in \mathsf{A}$*.*

1.6 Proposition (Cancellation)**.** *Let* $\mathfrak{a},\mathfrak{b},\mathfrak{c} \in \mathsf{A}^*$*. If* $\mathfrak{ab} = \mathfrak{ac}$ *then* $\mathfrak{b} = \mathfrak{c}$*. If* $\mathfrak{ab} = \mathfrak{cb}$*, then* $\mathfrak{a} = \mathfrak{c}$*.*

Proof. If $\mathfrak{ab} = \mathfrak{ac}$ then $\mathrm{lh}(\mathfrak{a}) + \mathrm{lh}(\mathfrak{b}) = \mathrm{lh}(\mathfrak{a}) + \mathrm{lh}(\mathfrak{c})$, i.e. $\mathrm{lh}(\mathfrak{b}) = \mathrm{lh}(\mathfrak{c})$. Thus

$$\mathfrak{b}(\mathsf{k}) = \mathfrak{ab}(\mathsf{k} + \mathrm{lh}(\mathfrak{a})) = \mathfrak{ac}(\mathsf{k} + \mathrm{lh}(\mathfrak{a})) = \mathfrak{c}(\mathsf{k})$$

for all $\mathsf{k} \in [0, \mathrm{lh}(\mathfrak{b})[= [0, \mathrm{lh}(\mathfrak{c})[$. Hence $\mathfrak{b} = \mathfrak{c}$.

If $\mathfrak{ab} = \mathfrak{cb}$, $\mathrm{lh}(\mathfrak{a}) = \mathrm{lh}(\mathfrak{c})$ like above. Additionally

$$\mathfrak{a}(\mathsf{k}) = \mathfrak{ab}(\mathsf{k}) = \mathfrak{cb}(\mathsf{k}) = \mathfrak{c}(\mathsf{k})$$

for all $\mathsf{k} \in [0, \mathrm{lh}(\mathfrak{a})[= [0, \mathrm{lh}(\mathfrak{c})[$. Thus $\mathfrak{a} = \mathfrak{c}$. □

Cancellation will be used without mention throughout this paper, as will most "obvious" properties.

Remark 3. Note that $\diamond a = a\diamond = a$ for all $a \in A^*$, making A^* into a monoid. It is in fact isomorphic to the free monoid on A. We will exploit this and use the familiar notation $\prod$, defined by

$$\left(\prod_{i=n}^{m} a_i\right) = \begin{cases} \diamond & \text{if } n = m+1 \\ \left(\prod_{i=n}^{m-1} a_i\right) a_m & \text{if } n \leq m \end{cases}$$

where $n, m \in \mathbb{Z}$, $n \leq m+1$, for the composition of any sequence $\{a_i\}_{i=n}^{m}$ such that $a_i \in A^*$ for all $n \leq i \leq m$.

1.7 Definition (Substrings and initial segments)**.** Let $a, b \in A^*$. If there are $c, d \in A^*$ such that $b = cad$, then a is a *substring* of b, and we write

$$a \preceq b.$$

In the special case when $c = \diamond$ above, a is called an *initial segment* of b, denoted

$$a \sqsubseteq b.$$

1.8 Proposition. *Both* $\preceq$ *and* $\sqsubseteq$ *are partial orders of* A^**.*

Proof. Let $a, b, c \in A^*$.

Since $a = \diamond a \diamond$, we have $a \preceq a$ and $a \sqsubseteq a$.

Suppose that $a \preceq b$ and $b \preceq a$. Then there are $d, e, f, g \in A^*$ such that $b = dae$ and $a = fbg$. Hence $a = fdaeg$. Consequently

$$\begin{aligned} \mathrm{lh}(a) &= \mathrm{lh}(fdaeg) \\ &= \mathrm{lh}(f) + \mathrm{lh}(d) + \mathrm{lh}(a) + \mathrm{lh}(e) + \mathrm{lh}(g). \end{aligned}$$

It follows that $\mathrm{lh}(f) + \mathrm{lh}(d) + \mathrm{lh}(e) + \mathrm{lh}(g) = 0$. As $\mathrm{lh}(f), \mathrm{lh}(d), \mathrm{lh}(e), \mathrm{lh}(g) \in \mathbb{N}$, we conclude that $\mathrm{lh}(f) = \mathrm{lh}(d) = \mathrm{lh}(e) = \mathrm{lh}(g) = 0$. Hence $d = e = f = g = \diamond$ and $a = \diamond b \diamond = b$.

If $a \sqsubseteq b$ and $b \sqsubseteq a$, the above situation applies.

Now suppose that $a \preceq b$ and $b \preceq c$. Then there exists $d, e, f, g \in A^*$ such that $b = dae$ and $c = fbg$. Hence $c = fdaeg$ and $a \preceq c$.

Finally, observe that if $a \sqsubseteq b$ and $b \sqsubseteq c$, then $d = \diamond = f$ in the preceding paragraph. Thus $c = aeg$ and $a \sqsubseteq c$. □

1.9 Lemma. *Let* $a, b \in A^*$. $a \sqsubseteq b$ *if and only if* $a = b \upharpoonright [0, \mathrm{lh}(a)[$.

Proof. If $a \sqsubseteq b$, then there is a $c \in A^*$ such that $ac = b$. Hence

$$b(k) = ac(k) = a(k)$$

for all $k \in [0, \mathrm{lh}(a)[$. Thus $b \upharpoonright [0, \mathrm{lh}(a)[= a$.

Conversely, if $a = b \restriction [0, \mathrm{lh}(a)[$, let $c \in A^{\mathrm{lh}(b)-\mathrm{lh}(a)}$ be defined by

$$c(k) = b(k + \mathrm{lh}(a))$$

for all $k \in [0, \mathrm{lh}(b) - \mathrm{lh}(a)[$. Then

$$\begin{aligned} ac(k) &= \begin{cases} a(k) & \text{if } k < \mathrm{lh}(a) \\ c(k - \mathrm{lh}(a)) & \text{if } k \geq \mathrm{lh}(a) \end{cases} \\ &= \begin{cases} a(k) & \text{if } k < \mathrm{lh}(a) \\ b(k - \mathrm{lh}(a) + \mathrm{lh}(a)) & \text{if } k \geq \mathrm{lh}(a) \end{cases} \\ &= \begin{cases} b(k) & \text{if } k < \mathrm{lh}(a) \\ b(k) & \text{if } k \geq \mathrm{lh}(a) \end{cases} \\ &= b(k) \end{aligned}$$

for all $k \in [0, \mathrm{lh}(a) + \mathrm{lh}(b) - \mathrm{lh}(a)[= [0, \mathrm{lh}(b)[$. Thus $ac = b$ and $a \sqsubseteq b$. □

1.10 Lemma. *Let $a, b, c \in A^*$ be such that $a \sqsubseteq bc$. If $\mathrm{lh}(a) \leq \mathrm{lh}(b)$, then $a \sqsubseteq b$. If $\mathrm{lh}(a) \geq \mathrm{lh}(b)$, there is a $d \sqsubseteq c$ such that $a = bd$ and $\mathrm{lh}(d) = \mathrm{lh}(a) - \mathrm{lh}(b)$.*

Proof. By definition there is an $e \in A^*$ such that $ae = bc$.

If $\mathrm{lh}(a) \leq \mathrm{lh}(b)$ then

$$a(k) = ae(k) = bc(k) = b(k)$$

for all $k \in [0, \mathrm{lh}(a)[$, so that $a = b \restriction [0, \mathrm{lh}(a)[$. Hence $a \sqsubseteq b$ by above.

If on the other hand $\mathrm{lh}(a) \geq \mathrm{lh}(b)$, we have that

$$a(k) = bc(k) = c(k - \mathrm{lh}(b))$$

for all $k \in [\mathrm{lh}(b), \mathrm{lh}(a)[$ (this statement is empty if $\mathrm{lh}(a) = \mathrm{lh}(b)$). Thus define $d \in A^{\mathrm{lh}(a)-\mathrm{lh}(b)}$ by

$$d(k) = a(k + \mathrm{lh}(b))$$

for all $k \in [0, \mathrm{lh}(a) - \mathrm{lh}(b)[$, so that $d \sqsubseteq c$. Furthermore

$$\begin{aligned} ((a \restriction [0, \mathrm{lh}(b)[)d)(k) &= \begin{cases} (a \restriction [0, \mathrm{lh}(b)[)(k) & \text{if } k < \mathrm{lh}(a \restriction [0, \mathrm{lh}(b)[) \\ d(k - \mathrm{lh}(b)) & \text{if } k \geq \mathrm{lh}(a \restriction [0, \mathrm{lh}(b)[) \end{cases} \\ &= \begin{cases} a(k) & \text{if } k < \mathrm{lh}(b) \\ a(k) & \text{if } k \geq \mathrm{lh}(b) \end{cases} \\ &= a(k) \end{aligned}$$

for all $k \in [0, \mathrm{lh}(a \restriction [0, \mathrm{lh}(b)[) + \mathrm{lh}(d)[= [0, \mathrm{lh}(b) + \mathrm{lh}(a) - \mathrm{lh}(b) = [0, \mathrm{lh}(a)[$, whereby $(a \restriction [0, \mathrm{lh}(b)[)d = a$. □

1.11 Corollary. *If* $\mathsf{a} \sqsubseteq \mathsf{b}$ *and* $\mathrm{lh}(\mathsf{a}) = \mathrm{lh}(\mathsf{b})$, *then* $\mathsf{a} = \mathsf{b}$.

1.12 Definition. A symbol s is said to *occur* in a string u if $\mathsf{s} \preceq \mathsf{u}$.

1.13 Proposition. *Let* $\mathsf{s} \in \mathsf{A}$ *and* $\mathsf{a}, \mathsf{b} \in \mathsf{A}^*$. *Then* s *occurs in* ab *if and only if* s *occurs in at least one of* a *and* b.

Proof. If $\mathsf{s} \preceq \mathsf{a}$ or $\mathsf{s} \preceq \mathsf{b}$ then $\mathsf{s} \preceq \mathsf{ab}$, since $\mathsf{a}, \mathsf{b} \preceq \mathsf{ab}$ and $\preceq$ is a partial order.

Conversely, if $\mathsf{s} \preceq \mathsf{ab}$ then there are $\mathsf{c}, \mathsf{d} \in \mathsf{A}^*$ such that

$$\mathsf{csd} = \mathsf{ab}.$$

We consider cases according as $\mathrm{lh}(\mathsf{cs}) \leq \mathrm{lh}(\mathsf{a})$ or not.

If $\mathrm{lh}(\mathsf{cs}) \leq \mathrm{lh}(\mathsf{a})$ then $\mathsf{cs} \sqsubseteq \mathsf{a}$ by Lemma 1.10, whereby $\mathsf{s} \preceq \mathsf{a}$.

Otherwise $\mathrm{lh}(\mathsf{cs}) > \mathrm{lh}(\mathsf{a})$, whence $\mathrm{lh}(\mathsf{a}) \leq \mathrm{lh}(\mathsf{c})$. By Lemma 1.10, $\mathsf{a} \sqsubseteq \mathsf{c}$ that is $\mathsf{c} = \mathsf{ae}$ for some $\mathsf{e} \in \mathsf{A}^*$. Thus $\mathsf{b} = \mathsf{esd}$, i.e. $\mathsf{s} \preceq \mathsf{b}$. □

1.14 Definition (Substitution)**.** The *substitution operator* $\mathrm{S} : \mathsf{A} \longrightarrow ((\mathsf{A}^*)^{\mathsf{A}^*})^{\mathsf{A}^*}$ is defined via recursion.

- Let $\mathrm{S}(\mathsf{s})(\mathsf{b})(\diamond) = \diamond$.
- If $\mathsf{a} \in \mathsf{A}^*$ and $\mathsf{t} \in \mathsf{A}$ we define

$$\mathrm{S}(\mathsf{s})(\mathsf{b})(\mathsf{at}) = \begin{cases} \mathrm{S}(\mathsf{s})(\mathsf{b})(\mathsf{a})\mathsf{b} & \text{if } \mathsf{t} = \mathsf{s} \\ \mathrm{S}(\mathsf{s})(\mathsf{b})(\mathsf{a})\mathsf{t} & \text{if } \mathsf{t} \neq \mathsf{s}. \end{cases}$$

1.15 Proposition. *For all* $\mathsf{s} \in \mathsf{A}$ *and* $\mathsf{a}, \mathsf{b}, \mathsf{c} \in \mathsf{A}^*$ *we have that*

$$\mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{bc}) = \mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{b})\,\mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{c}).$$

Proof. By induction on c.

- Consider $\diamond \in \mathsf{A}^*$. Then

$$\begin{aligned} \mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{b}\diamond) &= \mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{b}) \\ &= \mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{b})\diamond \\ &= \mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{b})\,\mathrm{S}(\mathsf{s})(\mathsf{a})(\diamond) \end{aligned}$$

  by Remark 3 on page 5.

- Now suppose $\mathsf{d} \in \mathsf{A}^*$ is such that

$$\mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{bd}) = \mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{b})\,\mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{d})$$

  and let $\mathsf{t} \in \mathsf{A}$. Then

$$\begin{aligned} \mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{bdt}) &= \begin{cases} \mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{bd})\mathsf{a} & \text{if } \mathsf{t} = \mathsf{s} \\ \mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{bd})\mathsf{t} & \text{if } \mathsf{t} \neq \mathsf{s} \end{cases} \\ &= \begin{cases} \mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{b})\,\mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{d})\mathsf{a} & \text{if } \mathsf{t} = \mathsf{s} \\ \mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{b})\,\mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{d})\mathsf{t} & \text{if } \mathsf{t} \neq \mathsf{s} \end{cases} \\ &= \mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{b})\,\mathrm{S}(\mathsf{s})(\mathsf{a})(\mathsf{dt}). \end{aligned}$$

$\square$

Stated otherwise, the above proposition claims that, given any $s \in A$, $b \in A^*$ the operator $S(s)(b) : A^* \longrightarrow A^*$ is a monoid homomorphism. Since intuitively, $S(s)(b)(a)$ is the string obtained from $a \in A^*$ by replacing every occurrence of the symbol s in a by the string b, it has some other useful properties as well.

1.16 Lemma. *Let $s \in A$ and $b \in A^*$. For all $a \in A^*$ we have that:*

1. *If s occurs in a then*

$$b \preceq S(s)(b)(a).$$

2. *If s does not occur in a then*

$$S(s)(b)(a) = a.$$

3. *If s does not occur in b, then s does not occur in $S(s)(b)(a)$.*

4. $S(s)(s)(a) = a$.

Proof. By induction on a:

1.
   - Consider $\diamond \in A^*$. That s occurs in $\diamond$ is absurd, and there is nothing to prove.
   - Now assume that $c \in A^*$ is such that if s occurs in c then $b \preceq S(s)(b)(c)$ and let $t \in A$. Suppose that s occurs in ct. Then s occurs in either of c or t, the latter case in which $s = t$. If s occurs in c then $b \preceq S(s)(b)(c)$ by assumption. Hence

$$b \preceq S(s)(b)(c)\, S(s)(b)(t) = S(s)(b)(ct).$$

   If on the other hand $s = t$ we have that $b = S(s)(b)(t)$. Hence

$$b \preceq S(s)(b)(c)\, S(s)(b)(t) = S(s)(b)(ct).$$

   Thus the claim holds for all $a \in A^*$.

2.
   - Consider $\diamond \in A^*$. Then s does not occur in $\diamond$ and

$$S(s)(b)(\diamond) = \diamond$$

   as desired.
   - Now assume that $c \in A^*$ is such that if s does not occur in c then $S(s)(b)(c) = c$ and let $t \in A$. Suppose that s does not occur in ct. Then s does not occur in w and $s \neq t$. Hence

$$S(s)(b)(ct) = S(s)(b)(c)t = ct.$$

   Thus the claim holds for all $a \in A^*$.

3. • First, s does not occur in $S(\mathsf{s})(\mathsf{b})(\diamond) = \diamond$.

• Furthermore, if $\mathsf{c} \in A^*$ is such that $\mathsf{s} \not\preceq S(\mathsf{s})(\mathsf{b})(\mathsf{c})$, and $\mathsf{t} \in A$, we have, since

$$S(\mathsf{s})(\mathsf{b})(\mathsf{ct}) = \begin{cases} S(\mathsf{s})(\mathsf{b})(\mathsf{c})\mathsf{b} & \text{if } \mathsf{t} = \mathsf{s} \\ S(\mathsf{s})(\mathsf{b})(\mathsf{c})\mathsf{t} & \text{if } \mathsf{t} \neq \mathsf{s} \end{cases}$$

and $\mathsf{s} \not\preceq \mathsf{b}$, that $\mathsf{s} \not\preceq S(\mathsf{s})(\mathsf{b})(\mathsf{ct})$.

4. • First $S(\mathsf{s})(\mathsf{s})(\diamond) = \diamond$ by definition.

• Consider some $\mathsf{c} \in A^*$ such that $S(\mathsf{s})(\mathsf{s})(\mathsf{c}) = \mathsf{c}$ and $\mathsf{t} \in A$. Then

$$\begin{aligned} S(\mathsf{s})(\mathsf{s})(\mathsf{ct}) &= \begin{cases} S(\mathsf{s})(\mathsf{s})(\mathsf{c})\mathsf{s} & \text{if } \mathsf{t} = \mathsf{s} \\ S(\mathsf{s})(\mathsf{s})(\mathsf{c})\mathsf{t} & \text{if } \mathsf{t} \neq \mathsf{s} \end{cases} \\ &= \begin{cases} \mathsf{ct} & \text{if } \mathsf{t} = \mathsf{s} \\ \mathsf{ct} & \text{if } \mathsf{t} \neq \mathsf{s} \end{cases} \\ &= \mathsf{ct}. \end{aligned}$$

Hence $S(\mathsf{s})(\mathsf{s})(\mathsf{a}) = \mathsf{a}$ for all $\mathsf{a} \in A^*$.

□

1.17 Proposition. *Let* $\mathsf{s}, \mathsf{t} \in A$*,* $\mathsf{a}, \mathsf{b} \in A^*$ *be such that* $\mathsf{t} \not\preceq \mathsf{a}$*. Then*

$$S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{t})(\mathsf{a})) = S(\mathsf{s})(\mathsf{b})(\mathsf{a}).$$

Proof. By induction on a:

• By definition

$$\begin{aligned} S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{t})(\diamond)) &= S(\mathsf{t})(\mathsf{b})(\diamond) \\ &= \diamond \\ &= S(\mathsf{s})(\mathsf{b})(\diamond). \end{aligned}$$

• Let $\mathsf{c} \in A^*$ be such that if $\mathsf{t} \not\preceq \mathsf{c}$ then $S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{t})(\mathsf{c})) = S(\mathsf{s})(\mathsf{b})(\mathsf{c})$, and $\mathsf{r} \in A$. Assume $\mathsf{t} \not\preceq \mathsf{cr}$, that is $\mathsf{t} \not\preceq \mathsf{c}$ and $\mathsf{t} \neq \mathsf{r}$. Then

$$\begin{aligned} S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{t})(\mathsf{cr})) &= S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{t})(\mathsf{c})\, S(\mathsf{s})(\mathsf{t})(\mathsf{r})) \\ &= S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{t})(\mathsf{c}))\, S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{t})(\mathsf{r})) \\ &= S(\mathsf{s})(\mathsf{b})(\mathsf{c})\, S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{t})(\mathsf{r})) \\ &= \begin{cases} S(\mathsf{s})(\mathsf{b})(\mathsf{c})\, S(\mathsf{t})(\mathsf{b})(\mathsf{t}) & \text{if } \mathsf{r} = \mathsf{s} \\ S(\mathsf{s})(\mathsf{b})(\mathsf{c})\, S(\mathsf{t})(\mathsf{b})(\mathsf{r}) & \text{if } \mathsf{r} \neq \mathsf{s} \end{cases} \\ &= \begin{cases} S(\mathsf{s})(\mathsf{b})(\mathsf{c})\mathsf{b} & \text{if } \mathsf{r} = \mathsf{s} \\ S(\mathsf{s})(\mathsf{b})(\mathsf{c})\mathsf{r} & \text{if } \mathsf{r} \neq \mathsf{s} \end{cases} \\ &= S(\mathsf{s})(\mathsf{b})(\mathsf{cr}). \end{aligned}$$

$\square$

In words, the above proposition guarantees the possibility of an intermediate, or "dummy", substitution, that is, to replace a symbol s by some string b in another string a we can take the detour of first substituting s for some fresh symbol t, which we then replace by b. The next proposition, on the other hand, gives a sufficient condition for when the order of two consecutive substitutions is irrelevant.

1.18 Proposition. *If* $\mathsf{s}, \mathsf{t} \in \mathsf{A}$ *and* $\mathsf{a}, \mathsf{b} \in \mathsf{A}^*$ *are such that* $\mathsf{s} \not\preceq \mathsf{b}$*,* $\mathsf{t} \not\preceq \mathsf{a}$ *and* $\mathsf{s} \neq \mathsf{t}$*, then*

$$S(\mathsf{s})(\mathsf{a})(S(\mathsf{t})(\mathsf{b})(\mathsf{c})) = S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{a})(\mathsf{c}))$$

for all $\mathsf{c} \in \mathsf{A}^*$*.*

Proof. By induction on c. First, observe that for any $\mathsf{r} \in \mathsf{A}$ we have

$$S(\mathsf{s})(\mathsf{a})(S(\mathsf{t})(\mathsf{b})(\mathsf{r})) = \begin{cases} S(\mathsf{s})(\mathsf{a})(\mathsf{b}) & \text{if } \mathsf{r} = \mathsf{t} \\ S(\mathsf{s})(\mathsf{a})(\mathsf{r}) & \text{if } \mathsf{r} \neq \mathsf{t} \end{cases} = \begin{cases} \mathsf{b} & \text{if } \mathsf{r} = \mathsf{t} \\ \mathsf{a} & \text{if } \mathsf{r} = \mathsf{s} \\ \mathsf{r} & \text{if } \mathsf{s} \neq \mathsf{r} \neq \mathsf{t} \end{cases}$$

and

$$S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{a})(\mathsf{r})) = \begin{cases} S(\mathsf{t})(\mathsf{b})(\mathsf{a}) & \text{if } \mathsf{r} = \mathsf{s} \\ S(\mathsf{t})(\mathsf{b})(\mathsf{r}) & \text{if } \mathsf{r} \neq \mathsf{s} \end{cases} = \begin{cases} \mathsf{a} & \text{if } \mathsf{r} = \mathsf{s} \\ \mathsf{b} & \text{if } \mathsf{r} = \mathsf{t} \\ \mathsf{r} & \text{if } \mathsf{t} \neq \mathsf{r} \neq \mathsf{s} \end{cases}$$

whereby $S(\mathsf{s})(\mathsf{a})(S(\mathsf{t})(\mathsf{b})(\mathsf{r})) = S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{a})(\mathsf{r}))$.

- Consider $\diamond \in \mathsf{A}^*$. Then

$$\begin{aligned} S(\mathsf{s})(\mathsf{a})(S(\mathsf{t})(\mathsf{b})(\diamond)) &= S(\mathsf{s})(\mathsf{a})(\diamond) \\ &= \diamond \\ &= S(\mathsf{t})(\mathsf{b})(\diamond) \\ &= S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{a})(\diamond)). \end{aligned}$$

- Now let $\mathsf{d} \in \mathsf{A}^*$ be such that $S(\mathsf{s})(\mathsf{a})(S(\mathsf{t})(\mathsf{b})(\mathsf{d})) = S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{a})(\mathsf{d}))$, and $\mathsf{r} \in \mathsf{A}$. Then

$$\begin{aligned} S(\mathsf{s})(\mathsf{a})(S(\mathsf{t})(\mathsf{b})(\mathsf{dr})) &= S(\mathsf{s})(\mathsf{a})(S(\mathsf{t})(\mathsf{b})(\mathsf{d})\, S(\mathsf{t})(\mathsf{b})(\mathsf{r})) \\ &= S(\mathsf{s})(\mathsf{a})(S(\mathsf{t})(\mathsf{b})(\mathsf{d}))\, S(\mathsf{s})(\mathsf{a})(S(\mathsf{t})(\mathsf{b})(\mathsf{r})) \\ &= S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{a})(\mathsf{d}))\, S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{a})(\mathsf{r})) \\ &= S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{a})(\mathsf{d})\, S(\mathsf{s})(\mathsf{a})(\mathsf{r})) \\ &= S(\mathsf{t})(\mathsf{b})(S(\mathsf{s})(\mathsf{a})(\mathsf{dr})) \end{aligned}$$

  where the third equality follows from the initial observation and the induction hypothesis.

□

1.19 Definition. The *character counting function* $\# : \mathsf{A} \longrightarrow (\mathbb{N})^{\mathsf{A}^*}$ is recursively defined as follows for all $\mathsf{s} \in \mathsf{A}$:

$$\#(\mathsf{s})(\diamond) = 0$$
$$\#(\mathsf{s})(\mathsf{at}) = \begin{cases} \#(\mathsf{s})(\mathsf{a}) + 1 & \text{if } \mathsf{s} = \mathsf{t} \\ \#(\mathsf{s})(\mathsf{a}) & \text{otherwise} \end{cases}$$

for all $\mathsf{a} \in \mathsf{A}^*$ and $\mathsf{t} \in \mathsf{A}$.

Similarly to the substitution operator, given any $\mathsf{s} \in \mathsf{A}$ the "s counting function" $\#(\mathsf{s}) : \mathsf{A}^* \longrightarrow \mathbb{N}$ is a monoid homomorphism.

1.20 Proposition. *For all* $\mathsf{s} \in \mathsf{A}$ *and all* $\mathsf{a}, \mathsf{b} \in \mathsf{A}^*$, $\#(\mathsf{s})(\mathsf{ab}) = \#(\mathsf{s})(\mathsf{a}) + \#(\mathsf{s})(\mathsf{b})$.

Proof. By induction on b:

- First note that $\#(\mathsf{s})(\mathsf{a}\diamond) = \#(\mathsf{s})(\mathsf{a}) = \#(\mathsf{s})(\mathsf{a}) + 0 = \#(\mathsf{s})(\mathsf{a}) + \#(\mathsf{s})(\diamond)$.
- Thus consider $\mathsf{c} \in \mathsf{A}^*$ such that

  $$\#(\mathsf{s})(\mathsf{ac}) = \#(\mathsf{s})(\mathsf{a}) + \#(\mathsf{s})(\mathsf{c})$$

  and let $\mathsf{t} \in \mathsf{A}$. Then

  $$\begin{aligned} \#(\mathsf{s})(\mathsf{act}) &= \begin{cases} \#(\mathsf{s})(\mathsf{ac}) + 1 & \text{if } \mathsf{s} = \mathsf{t} \\ \#(\mathsf{s})(\mathsf{ac}) & \text{otherwise} \end{cases} \\ &= \begin{cases} \#(\mathsf{s})(\mathsf{a}) + \#(\mathsf{s})(\mathsf{c}) + 1 & \text{if } \mathsf{s} = \mathsf{t} \\ \#(\mathsf{s})(\mathsf{a}) + \#(\mathsf{s})(\mathsf{c}) & \text{otherwise} \end{cases} \\ &= \#(\mathsf{s})(\mathsf{a}) + \#(\mathsf{s})(\mathsf{ct}). \end{aligned}$$

□

1.21 Lemma. *Let* $\mathsf{s} \in \mathsf{A}$ *and* $\mathsf{a} \in \mathsf{A}^*$. *If* $\mathsf{s} \not\preceq \mathsf{a}$ *then* $\#(\mathsf{s})(\mathsf{a}) = 0$.

Proof. By induction on a.

- Consider $\diamond$. By definition $\#(\mathsf{s})(\diamond) = 0$.
- Now consider bt, where $\mathsf{b} \in \mathsf{A}^*$ is such that

  $$\#(\mathsf{s})(\mathsf{b}) = 0$$

  if $\mathsf{s} \not\preceq \mathsf{b}$, and $\mathsf{t} \in \mathsf{A}$. Assume that $\mathsf{s} \not\preceq \mathsf{bt}$. Then $\mathsf{s} \not\preceq \mathsf{b}$ and $\mathsf{s} \not\preceq \mathsf{t}$. Thus $\mathsf{t} \neq \mathsf{s}$ and $\#(\mathsf{s})(\mathsf{b}) = 0$ by induction hypothesis. We conclude that

  $$\#(\mathsf{s})(\mathsf{bt}) = \#(\mathsf{s})(\mathsf{b}) = 0.$$

□

Of course $\#(s)(a) = |\{k \in \mathbb{N} \mid a(k) = s\}|$, that is $\#(s)$ counts the number of occurrences of s in strings. Regrettably, with the "occurrence free" approach which we have taken towards strings this fact will not avail us (other than helping our intuition).

1.22 Proposition. *For every* $s \in A$, $\#(s)$ *is a monotone function from* $(A^*, \preceq)$ *to* $(\mathbb{N}, \leq)$.

Proof. Let $a, b \in A^*$ be such that $a \preceq b$. Then there are $c, d \in A^*$ such that

$$cad = b.$$

Hence

$$\begin{aligned} \#(s)(b) &= \#(s)(cad) \\ &= \#(s)(c) + \#(s)(a) + \#(s)(d) \\ &\geq \#(s)(a) \end{aligned}$$

as claimed. □

1.23 Corollary. *For every* $s \in A$, $\#(s)$ *is a monotone function from* $(A^*, \sqsubseteq)$ *to* $(\mathbb{N}, \leq)$.

Recursive Structures of Strings

In this subsection we will study sets of strings with a certain recursive construction, the elements of which we will call recursive strings. Our main interest in these objects lies in the fact that not only the well-formed formulas of simple type theory, but the types themselves, constitute recursive structures of strings (for certain alphabets). The rest of this section will be devoted to establishing some of their properties, chiefly the Recursion Theorem (Theorem 1.30), to be exploited in the following sections.

1.24 Definition. A *sequence of recursion prefixes* is a heptuple

$$\mathcal{H} = (A, B, n, \{a_k\}_{k\in[1,n[}, \{m_k\}_{k\in[1,n[}, [,]),$$

where A is the alphabet, $B \subseteq A$, $n \in \mathbb{N} \cup \{\omega\}$, $[,] \in A \setminus B$ are two distinct symbols such that $[,] \npreceq a_k$ for all $k \in [1, n[$, $\{m_k\}_{k\in]1,n]} \subseteq \mathbb{Z}^+$, and $\{a_k\}_{k\in[1,n[} \subseteq A^* \setminus B^*$ is a sequence without repetitions, such that for all $i, j \in [1, n[$

$$a_i \sqsubseteq a_j \Rightarrow a_i = \diamond$$

if $i \neq j$.

1.25 Proposition. *Let* $\mathcal{H} = (A, B, n, \{a_k\}_{k\in[1,n[}, \{m_k\}_{k\in[1,n[}, [,])$ *be a sequence of recursion prefixes. Then there is a least set* $Z \subseteq A^*$ *among all sets* $D \subseteq A^*$ *with the following properties:*

- *For every* $s \in B$*, the single character string* $s \in D$.
- *For every* $k \in [1, n[$ *and* $\{u_i\}_{i=1}^{m_k} \subseteq D$*, the string* $[a_k \left(\prod_{i=1}^{m_k} u_i\right)] \in D$.

Proof. Let

$$Q = \{X \subseteq A^* \,|\, \forall s \in B : s \in X \wedge \forall k \in [1, n[: \forall \{u_i\}_{i=1}^{m_k} \subseteq X : [a_k \left(\prod_{i=1}^{m_k} u_i\right)] \in X\}.$$

Since $A^* \in Q$, $Q \neq \varnothing$. Define $Z = \bigcap_{X \in Q} X$. Then for every $s \in B$, $s \in X$ for every $X \in Q$, whence $s \in Z$. Similarly, let $k \in [1, n[$ and $\{u_i\}_{i=1}^{m_k} \subseteq Z$. Then $\{u_i\}_{i=1}^{m_k} \subseteq X$ for all $X \in Q$ by definition, whereby $[a_k \left(\prod_{i=1}^{m_k} u_i\right)] \in X$ for all $X \in Q$. Hence $[a_k \left(\prod_{i=1}^{m_k} u_i\right)] \in Z$. Furthermore, if $D \subseteq A^*$ has the closure properties above, then clearly $D \in Q$, whence $Z \subseteq D$. So Z is the least set with the above closure properties, and thereby unique. □

1.26 Definition. Let $\mathcal{H} = (A, B, n, \{a_k\}_{k \in [1,n[}, \{m_k\}_{k \in [1,n[}, [,])$ be a sequence of recursion prefixes. The *recursive structure of strings induced by* $\mathcal{H}$ is the pair $\mathcal{Z} = (Z, \mathcal{H})$, where $Z \subseteq A^*$ is the least set with the following properties:

- For every $s \in B$, the single character string $s \in Z$.
- For every $k \in [1, n[$ and $\{u_i\}_{i=1}^{m_k} \subseteq Z$, the string $[a_k \left(\prod_{i=1}^{m_k} u_i\right)] \in Z$.

By Proposition 1.25 above, the recursive structure of strings induced by some sequence of recursion prefixes is well-defined, and properties of its constituent elements, which we will call recursive strings, can be established by structural induction.

A familiar example of such a structure is given by well formed formulas of propositional logic[1], which can be defined as the recursive structure induced by

$$(P \cup \{\neg, \vee, \wedge, \rightarrow, \leftrightarrow, [,]\}, P, 5, (\neg, \vee, \wedge, \rightarrow, \leftrightarrow), (1, 2, 2, 2, 2), [,]),$$

where P is the set of atomic formulas. A yet more familiar example is given by

$$\mathcal{N} = (N, (\{0, S, [,]\}, \{0\}, 1, (S)_{k=1}^{1}, [,])),$$

a somewhat opaque definition of the natural numbers.

For the rest of this section, unless otherwise stated

$$\mathcal{H} = (A, B, n, \{a_k\}_{k \in [1,n[}, \{m_k\}_{k \in [1,n[}, [,])$$

will denote a fixed sequence of recursive prefixes, with $\mathcal{Z} = (Z, \mathcal{H})$ the induced recursive structure of strings.

Remark 4. Note that $\diamond \notin Z$.

The following lemma and theorem rely heavily on Lemma 1.10 and its corollary.

[1] in polish notation

1.27 Lemma. *Let* $u \in Z$. *Every* $a \in A^* \setminus \{\diamond\}$ *such that* $a \sqsubseteq u$ *satisfy the following two conditions*

$$\#([)(a) \geq \#(])(a)$$

and

$$\#([)(a) = \#(])(a) \Leftrightarrow a = u.$$

Proof. By induction on u.

- For every $s \in B$, every $a \in A^* \setminus \{\diamond\}$ such that $a \sqsubseteq s$ satisfies $a = s$, whence

$$\#([)(a) = 0 = \#(])(a).$$

- Let $k \in [1, n[$ and consider $[a_k \left(\prod_{i=1}^{m_k} u_i\right)]$ where for each $i \in [1, m_k]$, $u_i \in Z$ is such that

$$\#([)(a) \geq \#(])(a)$$

and

$$\#([)(a) = \#(])(a) \Leftrightarrow a = u_i$$

for all $a \in A^* \setminus \{\diamond\}$ such that $a \sqsubseteq u_i$. Now let $b \in A^* \setminus \{\diamond\}$ be such that $b \sqsubseteq [a_k \left(\prod_{i=1}^{m_k} u_i\right)]$.

If $\mathrm{lh}(b) = \mathrm{lh}([a_k \left(\prod_{i=1}^{m_k} u_i\right)])$ then $b = [a_k \left(\prod_{i=1}^{m_k} u_i\right)]$, in which case

$$\begin{aligned} \#([)(b) &= 1 + \#([)(a_k) + \left(\sum_{i=1}^{m_k} \#([)(u_i)\right) \\ &= 1 + \left(\sum_{i=1}^{m_k} \#(])(w)\right) \\ &= 1 + \#(])(a_k) + \left(\sum_{i=1}^{m_k} \#(])(w)\right) \\ &= \#(])(b) \end{aligned}$$

by induction hypothesis.

If on the other hand $\mathrm{lh}(b) \leq \mathrm{lh}([a_k) < \mathrm{lh}([a_k \left(\prod_{i=1}^{m_k} u_i\right)])$, then

$$b \sqsubseteq [a_k.$$

Since $\mathrm{lh}(b) \geq 1$, we also have $[\sqsubseteq b$. Thus

$$1 \leq \#([)(b) \leq \#([)([a_k) = 1 + \#([)(a_k) = 1$$

and

$$0 \leq \#(])(b) \leq \#(])([a_k) = 0$$

so that $\#([)(b) = 1 > 0 = \#(])(b)$.

If finally $\mathrm{lh}([a_k) < \mathrm{lh}(b) < \mathrm{lh}([a_k \left(\prod_{i=1}^{m_k} u_i\right)])$ then there is a $c \in A^{\mathrm{lh}(b)-\mathrm{lh}([a_k)}$ such that $c \sqsubseteq \left(\prod_{i=1}^{m_k} u_i\right)$ and $b = [a_k c$. Hence

$$\#([)(b) = 1 + \#([)(a_k) + \#([)(c) \geq 1 + \#(])(a_k) + \#(])(c) = 1 + \#(])(b) > \#(])(b)$$

by induction hypothesis.

Thus $\#([)(b) \geq \#(])(b)$, and $\#([)(b) > \#(])(b)$ if $\mathrm{lh}(b) < \mathrm{lh}([a_k \left(\prod_{i=1}^{m_k} u_i\right)])$. Hence

$$\mathrm{lh}(b) = \mathrm{lh}([a_k \left(\prod_{i=1}^{m_k} u_i\right)])$$

and

$$b = [a_k \left(\prod_{i=1}^{m_k} u_i\right)]$$

if $\#([)(b) = \#(])(b)$. Since conversely

$$\begin{aligned}\#([)(b) &= 1 + \#([)(a_k) + \left(\sum_{i=1}^{m_k} \#([)(u_i)\right) \\ &= 1 + \#(])(a_k) \left(\sum_{i=1}^{m_k} \#(])(u_i)\right) \\ &= \#(])(b)\end{aligned}$$

when $b = [a_k \left(\prod_{i=1}^{m_k} u_i\right)]$, we see that

$$\#([)(b) = \#(])(b) \Leftrightarrow b = [a_k \left(\prod_{i=1}^{m_k} u_i\right)].$$

□

1.28 Corollary. *Let* $m \in \mathbb{N}$ *and* $\{u_i\}_{i=1}^m, \{v_i\}_{i=1}^m \in Z$. *If* $(\prod_{i=1}^m u_i) \sqsubseteq (\prod_{i=1}^m v_i)$ *then* $u_i = v_i$ *for all* $i \in [1, m]$.

Proof. Let $a \in A^*$ be such that $(\prod_{i=1}^m u_i)\, a = (\prod_{i=1}^m v_i)$. Let $p \in [0, m]$ be greatest such that $u_i = v_i$ for all $i \in [1, p]$. But since $(\prod_{i=1}^m u_i)\, a = (\prod_{i=1}^m v_i)$, this entails $\left(\prod_{i=p+1}^m u_i\right) a = \left(\prod_{i=p+1}^m v_i\right)$. Assuming $p < m$, we get in particular

$$u_{p+1} \sqsubseteq \left(\prod_{i=p+1}^{m} v_i\right).$$

Depending on whether $\mathrm{lh}(u_{p+1}) \leq \mathrm{lh}(v_{p+1})$ or $\mathrm{lh}(u_{p+1}) \geq \mathrm{lh}(v_{p+1})$, we have $u_{p+1} \sqsubseteq v_{p+1}$ or $v_{p+1} \sqsubseteq u_{p+1}$. In either case $u_{p+1} = v_{p+1}$ since $\#([)(u_{p+1}) = \#(])(u_{p+1})$ and $\#([)(v_{p+1}) = \#(])(v_{p+1})$, which contradicts the choice of p. Hence $p = m$. □

The following theorem is the essential reason why recursive definitions of functions over recursive strings work.

1.29 Theorem. *Let $\mathcal{H} = (A, B, n, \{a_k\}_{k\in[1,n[}, \{m_k\}_{k\in[1,n[}, [,])$ be a sequence of recursion prefixes, with $\mathcal{Z} = (Z, \mathcal{H})$ the induced recursive structure of strings. For every $u \in Z$, the following conditions hold:*

1. *If $u = s$ for some $s \in B$ then s is unique with this property, and there is no $k \in [1, n[$ such that $u = [a_k \left(\prod_{i=1}^{m_k} u_i\right)]$ for some $\{u_i\}_{i=1}^{m_k} \subseteq Z$.*
2. *If $u \neq s$ for all $s \in B$ then $u = [a_k \left(\prod_{i=1}^{m_k} u_i\right)]$ for uniquely determined $k \in [1, n[$ and $\{u_i\}_{i=1}^{m_k} \subseteq Z$.*

Proof. By induction on u.

- Consider $s \in B$. (Clearly there is a unique $t \in B$ such that $s = t$.) Furthermore, if there are $k \in [1, n[$ and $\{u_i\}_{i=1}^{m_k} \in Z$ such that $s = [a_k \left(\prod_{i=1}^{m_k} u_i\right)]$, then

$$1 = \mathrm{lh}(s) = \mathrm{lh}([a_k \left(\prod_{i=1}^{m_k} u_i\right)]) = 2 + \mathrm{lh}(a_k) + \left(\sum_{i=1}^{m_k} \mathrm{lh}(u_i)\right) \geq 2$$

  which is absurd.

- Now consider $[a_k \left(\prod_{i=1}^{m_k} u_i\right)]$, where $k \in [1, n[$, and for each $i \in [1, m_k]$, $u_i \in Z$ is such that

  1. if $u_i = s$ for some $s \in B$ then s is unique with this property, and there is no $k \in [1, n[$ such that $u_i = [a_k \left(\prod_{j=1}^{m_k} w_j\right)]$ for some $\{u_i\}_{i=1}^{m_k} \subseteq Z$;
  2. if $u \neq s$ for all $s \in B$ then $u = [a_k \left(\prod_{i=1}^{m_k} u_i\right)]$ for uniquely determined $k \in [1, n[$ and $\{u_i\}_{i=1}^{m_k} \subseteq Z$.

  Like above $\mathrm{lh}([a_k \left(\prod_{i=1}^{m_k} u_i\right)]) \geq 2$, whence $[a_k \left(\prod_{i=1}^{m_k} u_i\right)] \neq s$ for all $s \in B$. Therefore assume that there are $l \in [1, n[$ and $\{v_j\}_{j=1}^{m_l} \in Z^*$ such that

$$[a_k \left(\prod_{i=1}^{m_k} u_i\right)] = [a_l \left(\prod_{j=1}^{m_l} v_j\right)].$$

  Then $a_k \left(\prod_{i=1}^{m_k} u_i\right) = a_l \left(\prod_{j=1}^{m_l} v_j\right)$. Depending on whether $\mathrm{lh}(a_k) \leq \mathrm{lh}(a_l)$ or $\mathrm{lh}(a_l) \leq \mathrm{lh}(a_k)$, $a_k \sqsubseteq a_l$ or $a_l \sqsubseteq a_k$. Thus $a_k = \diamond$, $a_l = \diamond$ or $l = k$.

  If $a_k = \diamond$, let $p \in [0, m_k]$ be least such that $\mathrm{lh}(\left(\prod_{i=1}^{p} u_i\right)) \geq \mathrm{lh}(a_l)$. By 1.10, there is some $b \sqsubseteq \left(\prod_{i=p+1}^{m_k} u_i\right)$ such that $a_l = \left(\prod_{i=1}^{p} u_i\right) b$. Since $b \preceq a_l$

$$\#([)(b) = 0 = \#(])(b),$$

whereby $b = \diamond$ or $b = \left(\prod_{i=p+1}^{m_k} u_i\right)$. In the latter case $a_l = \left(\prod_{i=1}^{m_k} u_i\right)$, a contradiction since $\left(\prod_{i=1}^{m_k} u_i\right) \in B^*$ or $[\preceq \left(\prod_{i=1}^{m_k} u_i\right)$ by induction hypothesis. So $b = \diamond$ and thus $a_l = \left(\prod_{i=1}^{p} u_i\right)$. If $p > 0$ then, like above, $a_l \in B^*$ or $[\preceq a_l$, which both contradict the definition of a_l. Hence $p = 0$, which implies that $a_l = \diamond = a_k$.

If $a_l = \diamond$, let $p \in [0, m_l]$ be least such that $\mathrm{lh}(\left(\prod_{j=1}^{p} v_j\right)) \geq \mathrm{lh}(a_k)$. By 1.10, there is some $b \sqsubseteq \left(\prod_{j=p+1}^{m_l} v_j\right)$ such that $a_k = \left(\prod_{j=1}^{p} v_j\right) b$. Then

$$\#([)(b) = 0 = \#(])(b)$$

like above, whereby $b = \diamond$ or $b = \left(\prod_{j=p+1}^{m_l} v_j\right)$. In the latter case $a_k = \left(\prod_{j=1}^{m_l} v_j\right)$, a contradiction since $\left(\prod_{j=1}^{m_l} v_j\right) \in B^*$ or $[\preceq \left(\prod_{j=1}^{m_l} v_j\right)$. In the former case $a_k = \left(\prod_{j=1}^{p} v_j\right)$, which, like above, implies that $p = 0$ and $a_k = \diamond = a_l$.

In any case $l = k$, since $\{a_i\}_{i \in [1,n[}$ does not allow repetitions. Thus $\left(\prod_{i=1}^{m_k} u_i\right) = \left(\prod_{i=1}^{m_k} v_i\right)$. By Corollary 1.28, $u_i = v_i$ for all $i \in [1, m_k]$.

□

As promised, this theorem has the following consequence, the proof of which is based upon the proof of the recursion theorem for natural numbers in [4]. It assures that functions on Z can be defined by structural recursion.

1.30 Theorem (Recursion Theorem). *Let $\mathcal{H} = (A, B, n, \{a_k\}_{k\in[1,n[}, \{m_k\}_{k\in[1,n[}, [,])$ be a sequence of recursion prefixes, with $\mathcal{Z} = (Z, \mathcal{H})$ the induced recursive structure of strings. For any nonempty set M and sequence $\{f_k\}_{k\in[0,n[}$ of functions $f_0 : B \longrightarrow M$, $f_k : M^{m_k} \longrightarrow M$, there is a unique function $F : Z \longrightarrow M$ satisfying*

- $F(s) = f_0(s)$ *for all* $s \in B$.
- $F([a_k \left(\prod_{i=1}^{m_k} u_i\right)]) = f_k(\{F(u_i)\}_{i=1}^{m_k})$ *for all* $k \in [1, n[$ *and* $\{u_i\}_{i=1}^{m_k} \subseteq Z^*$.

Proof. Define

$$\Xi = \{V \in \mathcal{P}(Z) \mid (\forall s \in B : s \in V) \wedge (\forall k \in [1,n[: \forall \{u_i\}_{i=1}^{m_k} \subseteq Z : [a_k \left(\prod_{i=1}^{m_k} u_i\right)] \in V \Rightarrow (\forall i \in [1, m_k] : u_i \in V))\}$$

and $\Psi : \Xi \longrightarrow \bigcup_{V\in\Xi} M^V$ by

$$\Psi(V) = \{p : V \longrightarrow M \mid (\forall s \in B : p(s) = f_0(s)) \wedge (\forall k \in [1,n[: \forall \{u_i\}_{i=1}^{m_k} \in Z^* : [a_k \left(\prod_{i=1}^{m_k} u_i\right)] \in V \Rightarrow p([a_k \left(\prod_{i=1}^{m_k} u_i\right)]) = f_k(\{p(u_i)\}_{i=1}^{m_k}))\}$$

for all $V \in \Xi$.

Sublemma 1.30.1. *For all* $u \in Z$*, if* $V, \mathcal{V} \in \Xi$*,* $p \in \Psi(V)$ *and* $r \in \Psi(\mathcal{V})$ *are such that* $u \in V \cap \mathcal{V}$*, then*

$$p(u) = r(u).$$

Proof. The proof will be by induction on $u \in Z$.

- Consider $s \in B$. We know that $s \in V \cap \mathcal{V}$ and $p(s) = f_0(s) = r(s)$ by definition.
- Consider now $[a_k\left(\prod_{i=1}^{m_k} v_i\right)] \in Z$, where $\{v_i\}_{i=1}^{m_k} \in Z$ is such that

  $$v_i \in V \cap \mathcal{V} \Rightarrow p(v_i) = r(v_i)$$

  for each $i \in [1, m_k]$. Assume that $[a_k\left(\prod_{i=1}^{m_k} v_i\right)] \in V \cap \mathcal{V}$. By definition of Ξ, $v_i \in V \cap \mathcal{V}$ for each $i \in [1, m_k]$. Thus $p(v_i) = r(v_i)$ for each $i \in [1, m_k]$ by hypothesis, and hence

  $$\begin{aligned} p([a_k\left(\prod_{i=1}^{m_k} v_i\right)]) &= f_k(\{p(v_i)\}_{i=1}^{m_k}) \\ &= f_k(\{r(v_i)\}_{i=1}^{m_k}) \\ &= r([a_k\left(\prod_{i=1}^{m_k} v_i\right)]) \end{aligned}$$

  again by definition.

□

Furthermore define

$$\Omega = \bigcup_{V \in \Xi} \Psi(V).$$

As is the common practise, we identify functions with their graphs. Hence the definition

$$F = \bigcup_{p \in \Omega} p$$

is meaningful, and by above is a well-defined function.

Sublemma 1.30.2. *For every* $u \in Z$ *there is a* $V \in \Xi$ *such that* $u \in V$ *and* $\Psi(V) \neq \varnothing$*.*

Proof. By induction on u.

- Consider $s \in B$. Let $V = \{v \in Z \mid \exists t \in B : v = t\}$, so that in particular $s \in V$. Thus $t \in V$ for all $t \in B$, and if $[a_l\left(\prod_{i=1}^{m_l} w_i\right)] \in V$ for some $l \in [1, n[$ and $\{w_i\}_{i=1}^{m_l} \subseteq Z$ then $w_i \in V$ for all $i \in [1, m_l]$ vacuously by 1.29. Hence $V \in \Xi$. Furthermore $p : V \longrightarrow M$ defined by

  $$p(t) = f_0(t)$$

  for all $t \in B$ is such that $p \in \Psi(V)$.

- Now let $k \in [1, n[$ and consider $[a_k(\prod_{i=1}^{m_k} v_i)] \in Z$ where $\{v_i\}_{i=1}^{m_k} \subseteq Z$ is such that there for each $i \in [1, m_k]$ is a $\mathcal{V}_i \in \Xi$ for which $v_i \in \mathcal{V}_i$ and $\Psi(\mathcal{V}_i) \neq \varnothing$. Let $\{\mathcal{V}_i\}_{i=1}^{m_k} \subseteq \Xi$ be this sequence, and $\{r_i\}_{i=1}^{m_k} \in \prod_{i=1}^{m_k} \Psi(\mathcal{V}_i)$. Furthermore, let $\mathcal{V} = \bigcup_{i=1}^{m_k} \mathcal{V}_i$ and define $r : \mathcal{V} \longrightarrow M$ via

$$r = \bigcup_{i=1}^{m_k} r_i.$$

This is a well-defined function, by Sublemma 1.30.1. Define $V = \mathcal{V} \cup \{[a_k(\prod_{i=1}^{m_k} v_i)]\}$ and $p = r \cup \{([a_k(\prod_{i=1}^{m_k} v_i)], f_k(\{r(v_i)\}_{i=1}^{m_k}))\}$. If $[a_k(\prod_{i=1}^{m_k} v_i)] \notin \mathcal{V}$ then p is a well-defined function. Otherwise $V = \mathcal{V}$ and there is a $j \in [1, m_k]$ such that $[a_k(\prod_{i=1}^{m_k} v_i)] \in \mathcal{V}_j$, whence

$$\begin{aligned} r([a_k\left(\prod_{i=1}^{m_k} v_i\right)]) &= r_j([a_k\left(\prod_{i=1}^{m_k} v_i\right)]) \\ &= f_k(\{r_j(v_i)\}_{i=1}^{m_k}) \\ &= f_k(\{r(v_i)\}_{i=1}^{m_k}) \end{aligned}$$

and $p = r$, a well-defined function. Furthermore, for all $s \in B$, $s \in \mathcal{V}_i$ for all $i \in [1, m_k]$, whereby $s \in V$ and $p(s) = r(s) = f_0(s)$. Now assume $l \in [1, n[$ and $\{w_i\}_{i=1}^{m_l} \subseteq Z$ are such that $[a_l(\prod_{i=1}^{m_l} w_i)] \in V$. Then either there is a $j \in [1, m_k]$ such that $[a_l(\prod_{i=1}^{m_l} w_i)] \in \mathcal{V}_j$, or $[a_l(\prod_{i=1}^{m_l} w_i)] = [a_k(\prod_{i=1}^{m_k} v_i)]$. In the former case, $w_i \in \mathcal{V}_j$ for all $i \in [1, m_l]$ and

$$\begin{aligned} p([a_l\left(\prod_{i=1}^{m_l} w_i\right)]) &= r_j([a_l\left(\prod_{i=1}^{m_l} w_i\right)]) \\ &= f_l(\{r_j(w_i)\}_{i=1}^{m_l}) \\ &= f_l(\{p(w_i)\}_{i=1}^{m_l}) \end{aligned}$$

by induction hypothesis. In the latter case $l = k$ and $w_i = v_i \in V$ for all $i \in [1, m_k]$ by Theorem 1.29, whence

$$\begin{aligned} p([a_l\left(\prod_{i=1}^{m_l} w_i\right)]) &= p([a_k\left(\prod_{i=1}^{m_k} v_i\right)]) \\ &= f_k(\{r(v_i)\}_{i=1}^{m_k}) \\ &= f_k(\{p(w_i)\}_{i=1}^{m_l}) \end{aligned}$$

by definition. In either case $w_i \in V$ for all $i \in [1, m_l]$, so that $V \in \Xi$, and $p \in \Psi(V)$.

□

Thus $F : Z \longrightarrow M$, that is F is a function defined on all of Z. Note that for all $s \in B$,

$$F(s) = p(s) = f_o(s)$$

for all $p \in \Omega$. Furthermore, for all $k \in [1, n[$ and $\{u_i\}_{i=1}^{m_k} \subseteq Z$ there is a $V \in \Xi$ with $\Psi(V) \neq \varnothing$ such that $[a_k \left(\prod_{i=1}^{m_k} u_i\right)] \in V$, by above. Hence

$$\begin{aligned} F([a_k \left(\prod_{i=1}^{m_k} u_i\right)]) &= p([a_k \left(\prod_{i=1}^{m_k} u_i\right)]) \\ &= f_k(\{p(u_i)\}_{i=1}^{m_k}) \\ &= f_k(\{F(u_i)\}_{i=1}^{m_k}) \end{aligned}$$

for any (i.e. some) $p \in \Psi(V)$. Thus $F \in \Psi(Z)$, whereby it is unique with the above properties, by Sublemma 1.30.1. □

This concludes our discussion of strings. The work we have thus carried out will pay off in the following section, when we use these concepts to define the formal language of the simple theory of types, the principal construction of this paper.

2 The Formal Language of the Simple Theory of Types

This section is divided into two parts. In the first we define the basic constructions of the language; in the second we define free and bound variables and substitution with respect to such, and derive some relationships between, and closure properties of the language under, these.

Symbols and Formulas

2.1 Definition (Type symbols). Let i, o, $\langle$ and $\rangle$ be distinct symbols[2]. Let $\mathcal{T} \subseteq \{\mathsf{i}, \mathsf{o}, \langle, \rangle\}^*$ be the smallest subset of $\{\mathsf{i}, \mathsf{o}, \langle, \rangle\}^*$ such that:

- $\mathsf{i}, \mathsf{o} \in \mathcal{T}$
- $\forall \alpha, \beta \in \mathcal{T}(\langle \alpha\beta \rangle \in \mathcal{T})$

The members of $\mathcal{T}$ are called *type symbols*.

Remark 5. We will follow Church and abbreviate $\langle \alpha\beta \rangle$ by $\alpha\beta$ (that is, remove the outermost pair of parentheses), and by α' denote $\langle\langle\alpha\alpha\rangle\langle\alpha\alpha\rangle\rangle$.

The type denoted by i is the type of *individuals*, which are thought of as constituting the most primitive objects of our interest, and o is the type of *propositions*. Furthermore, $\alpha\beta$ is intended to be viewed as the type of functions from β to α, as will be seen later. Thus, for example, oi denotes the type of unary propositional functions, that is propositions which may be false for certain individuals while true for others. Likewise $\langle \mathsf{oi} \rangle \mathsf{i}$ will denote the type of binary propositional functions. Since function application will be denoted by juxtaposition (enclosed in brackets), this has the notational advantage of displaying the type of the argument of a function at the same place as said argument are to be placed in the concatenation, i.e. last. Thus if φ is of type $\alpha\beta$ and ψ is of type β, $[\varphi\psi]$ will be of type α.

Remark 6. Of course $(\mathcal{T}, (\{\mathsf{i}, \mathsf{o}, \langle, \rangle\}, \{\mathsf{i}, \mathsf{o}\}, \{\diamond\}_{k=1}^{1}, \{2\}_{k=1}^{1}, \langle, \rangle))$ is the recursive structure of strings induced by $(\{\mathsf{i}, \mathsf{o}, \langle, \rangle\}, \{\mathsf{i}, \mathsf{o}\}, \{\diamond\}_{k=1}^{1}, \{2\}_{k=1}^{1}, \langle, \rangle)$. Hence we can define functions on $\mathcal{T}$ by recursion.

2.2 Definition (Primitive symbols). Let λ, $[$, $]$, N and A be distinct symbols and $\Pi : \mathcal{T} \longrightarrow \mathsf{P}$, $\iota : \mathcal{T} \longrightarrow \mathrm{I}$ be bijective functions, where P and I are two disjoint sets not containing any of these symbols. The set $\mathcal{K} = \{\lambda, [,]\}$ is called the set of *improper symbols*, and $\mathcal{L} = \{\mathsf{N}, \mathsf{A}\} \cup \mathsf{P} \cup \mathrm{I}$ the set of *logical constants*.

Furthermore, let $\{X_\alpha\}_{\alpha\in\mathcal{T}}$ be a family of pairwise disjoint countable sets such that $X_\alpha \cap (\mathcal{K} \cup \mathcal{L}) = \emptyset$ for every $\alpha \in \mathcal{T}$. The set $\mathcal{X} = \bigcup_{\alpha\in\mathcal{T}} X_\alpha$ will be called the set of *variables*. Let $\mathfrak{a} : \mathbb{N} \times \mathcal{T} \longrightarrow \mathcal{X}$ be an injective function such that $\mathfrak{a}((n, \alpha)) \in X_\alpha$ for all $n \in \mathbb{N}$.

Finally, let $\{C_\alpha\}_{\alpha\in\mathcal{T}}$ be a family of pairwise disjoint sets such that for all $\alpha \in \mathcal{T}$ we have $C_\alpha \cap (\mathcal{K} \cup \mathcal{L} \cup \mathcal{X}) = \emptyset$. Denote by $\mathcal{C} = \bigcup_{\alpha\in\mathcal{T}} C_\alpha$ the set of *constant symbols* (or

[2]i.e. objects

non-logical symbols). Then $\mathcal{B} = \mathcal{L} \cup \mathcal{X} \cup \mathcal{C}$ is the set of *proper symbols* and $\mathcal{S} = \mathcal{K} \cup \mathcal{B}$ is the set of *primitive symbols*.

The primitive symbols will be the alphabet from which we build the formulas of type theory. While the definition above might look unnecessarily subdivided, this is necessary to distinguish the different sorts of symbols that we need.

In the deductive system to be defined, the improper symbols will have no meaning of their own (hence their name), but their appearance will determine the interpretation of the other symbols. Thus λ will be used to denote function abstraction (thus including concepts from λ-calculus in our formal system), and $[,]$ much like in the definition of recursive strings (1.26). Among the logical constants, N will serve to denote negation, and A will be the formal or-symbol, as will be seen, while for every type α, $\Pi(\alpha)$ and $\iota(\alpha)$ will be the universal quantifier and the description operator (picks an element from a non-empty set), respectively.

The variables hopefully needs no further explanation, while the function $\mathfrak{a}$ will only serve to pick out specific such.

The C_α, finally, are completely arbitrary; they might even all be empty. In fact, in [2], Church does not include the possibility of such symbols in the language, while Henkin (in [3]) merely remarks that their inclusion does not alter the results obtained without them.

However, most of these things are yet to be made precise, and this will to a large extent be our aim henceforth. To begin, we have the following definition, motivated by the above.

2.3 Definition (Types of symbols)**.** Let $t : \mathcal{B} \longrightarrow \mathcal{T}$ be such that for all $\alpha \in \mathcal{T}$

$$t(s) = \begin{cases} oo & \text{if } s = \mathsf{N} \\ \langle oo \rangle o & \text{if } s = \mathsf{A} \\ o\langle o\alpha \rangle & \text{if } s = \Pi(\alpha) \\ \alpha\langle o\alpha \rangle & \text{if } s = \iota(\alpha) \\ \alpha & \text{if } s \in X_\alpha \\ \alpha & \text{if } s \in C_\alpha \end{cases}$$

for all $s \in \mathcal{B}$. Note that the constraints placed on the sets in 2.2 ensures that the cases above are mutually exclusive, making t well-defined.

For every $s \in \mathcal{B}$, the type symbol $t(s)$ is called the *type* of s.

2.4 Definition (Formulas)**.** The set $U \subseteq \mathcal{S}^*$ of *formulas*[3] is the least subset of $\mathcal{S}^*$ such that:

- For every proper symbol $s \in \mathcal{B}$, the single character string s is a formula.
- For every formula $u \in U$ and every variable $x \in \mathcal{X}$, the string $[\lambda x u]$ is a formula.

[3]the letter U should suggest that the formulas are untyped

- For all formulas $u, v \in U$, the string $[uv]$ is a formula.

Remark 7. Given any enumeration $\pi : \mathbb{N} \longrightarrow \mathcal{X}$, let $\{a_k\}_{k\in[1,\omega[}$ be the sequence $a_1 = \diamond$ and $a_k = \lambda\pi(k-2)$ for $k \in [2, \omega[$. Let $\{m_k\}_{k\in[1,\omega[}$ be the sequence $m_1 = 2$, $m_k = 1$ for $k \in [2, \omega[$. Then $(\mathcal{S}, \mathcal{B}, \{a_k\}_{k\in[1,\omega[}, \{m_k\}_{k\in[1,\omega[}, [,])$ is a sequence of recursion prefixes, and $(U, (\mathcal{S}, \mathcal{B}, \{a_k\}_{k\in[1,\omega[}, \{m_k\}_{k\in[1,\omega[}, [,]))$ is the recursive structure of strings induced by these. Henceforth, let this structure be fixed.

2.5 Theorem. *For any nonempty set* M *and functions* $f : \mathcal{B} \longrightarrow M$, $g : \mathcal{X} \times M \longrightarrow M$ *and* $h : M \times M \longrightarrow M$, *there is a unique function* $F : U \longrightarrow M$ *satisfying*

- $F(s) = f(s)$ *for all* $s \in \mathcal{B}$.
- $F([\lambda x u]) = g((x, F(u)))$ *for all* $x \in \mathcal{X}$ *and* $u \in U$.
- $F([uv]) = h((F(u), F(v)))$ *for all* $u, v \in U$.

Proof. Let $f_0 = f$, $f_1 = h$ and $f_{k+2}(\{p\}_{i=1}^{1}) = g((\pi(k), p))$. Define F by the Recursion Theorem 1.30. To prove that F is the unique function satisfying the above recurrence relation (i.e. that the above construction is independent from the choice of π), assume that $G : U \longrightarrow M$ is such that

- $G(s) = f(s)$ for all $s \in \mathcal{B}$.
- $G([\lambda x u]) = g((x, G(u)))$ for all $x \in \mathcal{X}$ and $u \in U$.
- $G([uv]) = h((G(u), G(v)))$ for all $u, v \in U$;

we prove that F satisfies the recurrence relations and that $F(u) = G(u)$ for all $u \in U$ by structural induction on u:

- For all $s \in \mathcal{B}$ we have $G(s) = f(s) = f_0(s) = F(s)$.
- Consider $[\lambda x v]$, where $x \in \mathcal{X}$ and $v \in U$ is such that $G(v) = F(v)$. Let $k \in \mathbb{N}$ be such that $x = \pi(k)$. Then

$$\begin{aligned} G([\lambda x v]) &= g((x, G(v))) \\ &= f_{k+2}(\{G(v)\}_{i=1}^{1}) \\ &= f_{k+2}(\{F(v)\}_{i=1}^{1}) \\ &= F([\lambda x v]) \end{aligned}$$

- Finally consider $[vw]$, where $v, w \in U$ are such that $G(q) = F(q)$ for all $q \in \{v, w\}$. Then

$$\begin{aligned} G([vw]) &= h((G(v), G(w))) \\ &= f_1((G(v), G(w))) \\ &= f_1((F(v), F(w))) \\ &= F([vw]) \end{aligned}$$

Thus $\mathsf{F} = \mathsf{G}$ and satisfies the recurrence relations. $\square$

2.6 Definition (Well-formed formulas)**.** Let $\infty \notin \mathcal{T}$ be an arbitrary (though of course new) symbol, let $\bar{\mathcal{T}} = \mathcal{T} \cup \{\infty\}$ and define $\bar{\mathsf{T}} : \mathsf{U} \longrightarrow \bar{\mathcal{T}}$ by structural recursion:

- For every proper symbol $\mathsf{s} \in \mathcal{B}$ we define

$$\mathsf{T}(\mathsf{s}) = \mathsf{t}(\mathsf{s}).$$

- For every formula $u \in \mathsf{U}$ and every variable $\mathbf{x} \in \mathcal{X}$,

$$\bar{\mathsf{T}}([\lambda \mathbf{x} u]) = \begin{cases} \langle \bar{\mathsf{T}}(u)\bar{\mathsf{T}}(\mathbf{x})\rangle & \text{if } \bar{\mathsf{T}}(u) \in \mathcal{T} \\ \infty & \text{otherwise.} \end{cases}$$

- For all formulas $u, v \in \mathsf{U}$, and any $\alpha, \beta \in \mathcal{T}$

$$\bar{\mathsf{T}}([uv]) = \begin{cases} \alpha & \text{if } \bar{\mathsf{T}}(u) = \alpha\beta \text{ and } \bar{\mathsf{T}}(v) = \beta \\ \infty & \text{otherwise.} \end{cases}$$

A formula $u \in \mathsf{U}$ is *well-formed* (or well-typed) if $\bar{\mathsf{T}}(u) \in \mathcal{T}$. The set of all well-formed formulas is denoted by W. The function $\mathsf{T} = \bar{\mathsf{T}} \upharpoonright W$ is called the *type assignment*, which to every $\varphi \in W$ assigns its *type*, $\mathsf{T}(\varphi) \in \mathcal{T}$.

Remark 8. If $\varphi \in W$ is such that $\mathsf{T}(\varphi) = \alpha$, we will use φ_α and φ synonymously. Thus writing φ_α for some $\varphi \in W$ will mean simply that $\varphi \in W_\alpha$.

We can verify properties of well-formed formulas by induction and define such properties by recursion. The latter is included mostly for completeness; it will not prove useful in the ... to come, since we will be working mostly with (untyped) formulas.

2.7 Proposition (Well-formed induction)**.** *Let* $\mathsf{E} \subseteq W$ *be such that:*

- *For all* $\mathsf{s} \in \mathcal{B}$, $\mathsf{s} \in \mathsf{E}$.
- *For all* $\alpha, \beta \in \mathcal{T}$, $\mathbf{x} \in X_\beta$ *and* $\varphi_\alpha \in \mathsf{E}$, $[\lambda \mathbf{x} \varphi] \in \mathsf{E}$.
- *For all* $\alpha, \beta \in \mathcal{T}$, $\varphi_{\alpha\beta}, \psi_\beta \in \mathsf{E}$, $[\varphi\psi] \in \mathsf{E}$.

Then $\mathsf{E} = W$.

Proof. We will show that for every $u \in \mathsf{U}$ such that $\bar{\mathsf{T}}(u) \in \mathcal{T}$, $u \in \mathsf{E}$, by induction on u.

- For all $\mathsf{s} \in \mathcal{B}$, $\mathsf{s} \in \mathsf{E}$ by assumption.
- Consider $[\lambda \mathbf{x} v] \in \mathsf{U}$, where $\mathbf{x} \in \mathcal{X}$ and $v \in \mathsf{U}$ is such that if $\bar{\mathsf{T}}(v) \in \mathcal{T}$ then $v \in \mathsf{E}$. Assume that $\bar{\mathsf{T}}([\lambda \mathbf{x} v]) \in \mathcal{T}$. By definition we must have $\bar{\mathsf{T}}(v) \in \mathcal{T}$, whence $v \in \mathsf{E}$ by induction hypothesis. Since $\mathbf{x} \in X_{\mathsf{T}(\mathbf{x})}$, $[\lambda \mathbf{x} v] \in \mathsf{E}$ by assumption.

- Finally consider $[vw]$, where $v, w \in U$ are such that if $\bar{T}(q) \in \mathcal{T}$ then $q \in E$, for $q \in \{v, w\}$. Assume that $\bar{T}([vw]) \in \mathcal{T}$. By definition we must have that there are $\alpha, \beta \in \mathcal{T}$ such that $\bar{T}(v) = \alpha\beta$, $\bar{T}(w) = \beta$. Since in particular $\bar{T}(v), \bar{T}(w) \in \mathcal{T}$, $v, w \in E$. By assumptions on E we have that $[vw] \in E$.

Hence $\varphi \in E$ for every $\varphi \in W$. Thus $E = W$. $\square$

2.8 Theorem (Well-formed recursion). *Let* $\{M_\alpha\}_{\alpha\in\mathcal{T}}$ *be a family of non-empty sets,* $\{g_{\alpha\beta}\}_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}} \in \prod_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}} M_{\alpha\beta}^{X_\alpha\times M_\beta}$, $\{h_{\alpha\beta}\}_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}} \in \prod_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}} M_\alpha^{M_{\alpha\beta}\times M_\beta}$ *and let* $f : \mathcal{B} \longrightarrow \bigcup_{\alpha\in\mathcal{T}} M_\alpha$ *be a function such that*

$$f(s_\alpha) \in M_\alpha$$

for all $\alpha \in T$ *and* $s_\alpha \in \mathcal{B}$. *Then there is a unique function* $F : W \longrightarrow \bigcup_{\alpha\in\mathcal{T}} M_\alpha$ *such that* $F(\varphi_\alpha) \in M_\alpha$ *for all* $\alpha \in \mathcal{T}$ *and* $\varphi_\alpha \in W$, *which moreover satisfies*

$$\begin{aligned} F(s) &= f(s) \\ F([\lambda x \psi_\beta]) &= g_{\alpha\beta}((x, F(\varphi))) \\ F([\varphi_{\alpha\beta}\psi_\beta]) &= h_{\alpha\beta}((F(\varphi), F(\psi))) \end{aligned}$$

for all $\alpha, \beta \in \mathcal{T}$, $s \in \mathcal{B}$, $x \in X_\alpha$ *and* $\varphi_{\alpha\beta}, \psi_\beta \in W$.

Proof. Let $M = \bigcup_{\alpha\in\mathcal{T}} M_\alpha$ and $M' = \bigcup_{\alpha\in\mathcal{T}}\{\alpha\} \times M_\alpha$. Define $f' : \mathcal{B} \longrightarrow M'$ by

$$f'(s) = (T(s), f(s))$$

for all $s \in \mathcal{B}$; this is well-defined since $f(s) \in M_{T(s)}$ for such s.

Moreover, for each $\alpha, \beta \in \mathcal{T}$ define $g'_{\alpha\beta} : X_\alpha \times (\{\beta\} \times M_\beta) \longrightarrow \{\alpha\beta\} \times M_{\alpha\beta}$ by

$$g'_{\alpha\beta}((x, (\beta, p))) = (\alpha\beta, g_{\alpha\beta}((x, p)))$$

for all $x \in X_\alpha$ and $p \in M_\beta$. Thus $\{g'_{\alpha\beta}\}_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}} \in \prod_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}}(\{\alpha\beta\}\times M_{\alpha\beta})^{X_\alpha\times(\{\beta\}\times M_\beta)}$. Since $\{X_\alpha\}_{\alpha\in\mathcal{T}}$ is a pairwise disjoint family of sets, so too is $\{X_\alpha \times (\{\beta\} \times M_\beta)\}_{(\alpha,\beta)}$. Thus $g' = \bigcup_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}} g'_{\alpha\beta}$ is a well-defined function, and since

$$\bigcup_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}} X_\alpha \times (\{\beta\} \times M_\beta) = (\bigcup_{\alpha\in\mathcal{T}} X_\alpha) \times (\bigcup_{\beta\in\mathcal{T}}\{\beta\} \times M_\beta) = \mathcal{X} \times M'$$

and $\bigcup_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}}\{\alpha\beta\} \times M_{\alpha\beta} \subset M'$, we get that $g' : \mathcal{X} \times M' \longrightarrow M'$; satisfying

$$g'((x, (\alpha, p))) = g'_{T(x)\alpha}((x, (\alpha, p))) = (T(x)\alpha, g_{T(x)\alpha}((x, p)))$$

for all $x \in \mathcal{X}$, $\alpha \in \mathcal{T}$ and $p \in M_\alpha$.

Finally, let $\{h'_{\alpha\beta}\}_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}} \in \prod_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}}(\{\alpha\} \times M_\alpha)^{(\{\alpha\beta\}\times M_{\alpha\beta})\times(\{\beta\}\times M_\beta)}$ be such that

$$h'_{\alpha\beta}(((\alpha\beta, p), (\beta, q))) = (\alpha, h_{\alpha\beta}((p, q)))$$

for all $\alpha, \beta \in \mathcal{T}$, $p \in M_{\alpha\beta}$ and $q \in M_\beta$. Similarly to above, the family $\{(\{\alpha\beta\} \times M_{\alpha\beta}) \times (\{\beta\} \times M_\beta)\}_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}}$ is pairwise disjoint, whence $h'' = \bigcup_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}} h'_{\alpha\beta}$ is a well-defined function $h'' : \bigcup_{(\alpha,\beta)\in\mathcal{T}\times\mathcal{T}}(\{\alpha\beta\} \times M_{\alpha\beta}) \times (\{\beta\} \times M_\beta) \longrightarrow M'$. Let $a \in M'$ (arbitrarily) and define $h' : M' \times M' \longrightarrow M'$ by

$$h'(((\alpha, p), (\beta, q))) = \begin{cases} h''(((\alpha, p), (\beta, q))) & \text{if } \exists\gamma \in T : \alpha = \gamma\beta \\ a & \text{otherwise} \end{cases}$$

for all $\alpha, \beta \in \mathcal{T}$, $p \in M_\alpha$ and $q \in M_\beta$. By construction we have that

$$h'(((\alpha\beta, p), (\beta, q))) = h''(((\alpha\beta, p), (\beta, q))) = h'_{\alpha\beta}(((\alpha\beta, p), (\beta, q))) = (\alpha, h_{\alpha\beta}((p, q)))$$

for all $\alpha, \beta \in \mathcal{T}$, $p \in M_{\alpha\beta}$ and $q \in M_\beta$.

Now define $F' : U \longrightarrow M'$ by recursion (with respect to f', g' and h'; Theorem 2.5). Let $d : M' \longrightarrow M$ be defined by

$$d((\alpha, p)) = p$$

for all $\alpha \in \mathcal{T}$ and $p \in M_\alpha$. The function $F : \mathcal{W} \longrightarrow \bigcup_{\alpha\in\mathcal{T}} M_\alpha$ given by $F = d \circ F' \restriction \mathcal{W}$ is the desired function.

To prove that F is the unique function such that $F(\varphi) \in M_{T(\varphi)}$ for all φ in $\mathcal{W}$ satisfying the recurrence relations, let $G : \mathcal{W} \longrightarrow M$ be another such. We will prove that $F'(\varphi) = (T(\varphi), F(\varphi))$ and $G(\varphi) = F(\varphi)$ for all $\varphi \in \mathcal{W}$, by induction.

- Consider $s \in \mathcal{B}$. Since $F'(s) = f'(s) = (T(s), f(s))$ we have that

  $$G(s) = f(s) = d((T(s), f(s))) = d(F'(s)) = F(s)$$

  and thus $F'(s) = (T(s), F(s))$.

- Now consider $[\lambda x\psi] \in \mathcal{W}$, where $\alpha, \beta \in \mathcal{T}$, $x \in \mathcal{W}_\alpha$ and $\psi \in \mathcal{W}_\beta$ is such that $F'(\psi) = (\beta, F(\psi))$ and $G(\psi) = F(\psi)$. Then

  $$\begin{aligned} F'([\lambda x\psi]) &= g'((x, F'(\psi))) \\ &= g'((x, (\beta, F(\psi)))) \\ &= (\alpha\beta, g_{\alpha\beta}((x, F(\psi)))) \end{aligned}$$

  and thus

  $$\begin{aligned} G([\lambda x\psi]) &= g_{\alpha\beta}((x, G(\psi))) \\ &= g_{\alpha\beta}((x, F(\psi))) \\ &= d((\alpha\beta, g_{\alpha\beta}((x, F(\psi))))) \\ &= d(F'([\lambda x\psi])) \\ &= F([\lambda x\psi]). \end{aligned}$$

  Hence $F'([\lambda x\psi]) = (\alpha\beta, F([\lambda x\psi]))$.

- Finally consider $[\psi\vartheta] \in W_\alpha$, where $\alpha, \beta \in \mathcal{T}$, $\psi_{\alpha\beta}, \vartheta_\beta \in W$ are such that $G(\eta) = F(\eta)$ and $F'(\eta) = (T(\eta), F(\eta))$ for all $\eta \in \{\psi, \vartheta\}$. Then

$$\begin{aligned} F'([\psi\vartheta]) &= h'((F'(\psi), F'(\vartheta))) \\ &= h'(((\alpha\beta, F(\psi)), (\beta, F(\vartheta)))) \\ &= (\alpha, h_{\alpha\beta}((F(\psi), F(\vartheta)))) \end{aligned}$$

and thus

$$\begin{aligned} G([\psi\vartheta]) &= h_{\alpha\beta}((G(\psi), G(\vartheta))) \\ &= h_{\alpha\beta}((F(\psi), F(\vartheta))) \\ &= d((\alpha, h_{\alpha\beta}((F(\psi), F(\vartheta))))) \\ &= d(F'([\psi\vartheta])) \\ &= F([\psi\vartheta]). \end{aligned}$$

Hence we also have $F'([\psi\vartheta]) = (\alpha, F([\psi\vartheta]))$.

Thus $G = F$ and $F'(\varphi) = (T(\varphi), F(\varphi))$ for all $\varphi \in W$. In particular this means that $F(\varphi) \in M_{T(\varphi)}$ for all $\varphi \in W$. □

Finally, more applicable are the following properties. These will be used in section 3 to prove that the rules of inference (Definition 3.2) preserve the property of being well-formed. The proof of the first makes heavy use of Lemma 1.10.

2.9 Lemma. *Let* $u, v, w \in U$.

1. *Let* $x \in \mathcal{X}$. *If* $x \neq u$ *and* $aub = [\lambda xv]$ *for some* $a, b \in \mathcal{S}^*$, *either* $u = [\lambda xv]$ *or there are* $c, d \in \mathcal{S}^*$ *such that* $cud = v$, $[\lambda xc = a$ *and* $d] = b$.

2. *If* $aub = [vw]$ *for some* $a, b \in \mathcal{S}^*$, *then* $u = [vw]$ *or there are* $c, d \in \mathcal{S}^*$ *such that either* $cud = v$, $[c = a$ *and* $dw] = b$, *or* $cud = w$, $[vc = a$ *and* $d] = b$.

Proof. 1. If $\mathrm{lh}(a) = 0$ then $u \sqsubseteq [\lambda xv]$, whence $u = [\lambda xv]$ by Corollary 1.28.

If $\mathrm{lh}(a) = 1$, then $a = [$, whereby $ub = \lambda xv]$. Since $\mathrm{lh}(u) \geq 1$, we have $\lambda \sqsubseteq u$, and $\#([)(\lambda) = 0\,\#(])(\lambda)$ as $[\, \neq \lambda \neq\,]$. By 1.27 $\lambda = u$, which is absurd.

If $\mathrm{lh}(a) = 2$, $a = [\lambda$, whence $ub = xv]$. Since $\mathrm{lh}(u) \geq 1$, $x \sqsubseteq u$. Hence $x = u$ by Corollary 1.28 (because $x \in U$), contradictory to the assumption.

Thus suppose $\mathrm{lh}(a) \geq 3$, whereby there is a $c \in \mathcal{S}^*$ such that $a = [\lambda xc$. Hence $cub = v]$. Furthermore, because $\mathrm{lh}(u) \geq 1$, we have

$$\begin{aligned} \mathrm{lh}(c) + \mathrm{lh}(u) &= \mathrm{lh}(cu) \\ &\leq \mathrm{lh}(v]) \\ &= \mathrm{lh}(v) + 1 \\ &\leq \mathrm{lh}(v) + \mathrm{lh}(u), \end{aligned}$$

so that $\mathrm{lh}(c) \leq \mathrm{lh}(v)$, whence $c \sqsubseteq v$. Since if $\mathrm{lh}(b) = 0$,

$$[\lambda x c u = [\lambda x v]$$

whereby

$$\begin{aligned}
\#([)([\lambda x c u) &= \#([)([\lambda x v]) \\
&= \#(])([\lambda x v]) \\
&= \#(])([\lambda x c u) \\
&= \#(])([\lambda x) + \#(])(c) + \#(])(u) \\
&= 0 + \#(])(c) + \#([)(u) \\
&\leq \#([)(c) + \#([)(u) \\
&< 1 + \#([)(c) + \#([)(u) \\
&= \#([)([\lambda x) + \#([)(c) + \#([)(u) \\
&= \#([)([\lambda x c u)
\end{aligned}$$

by Lemma 1.27, which is absurd, $\mathrm{lh}(b) > 0$. Thus

$$\begin{aligned}
\mathrm{lh}(cu) + \mathrm{lh}(b) &= \mathrm{lh}(cub) \\
&= \mathrm{lh}(v]) \\
&= \mathrm{lh}(v) + 1 \\
&\leq \mathrm{lh}(v) + \mathrm{lh}(b)
\end{aligned}$$

i.e. $\mathrm{lh}(cu) \leq \mathrm{lh}(v)$. Hence $cu \sqsubseteq v$, that is there is a $d \in \mathcal{S}^*$ such that

$$cud = v.$$

Since thereby $cud] = cub$, $d] = b$ as desired.

2. If $\mathrm{lh}(a) = 0$, $u \sqsubseteq [vw]$, whence $u = [vw]$ by Corollary 1.28.

   If $1 \leq \mathrm{lh}(a) \leq \mathrm{lh}(v)$, then $[\sqsubseteq a \sqsubseteq [v$ but $a \neq [v$. Let $c, e \in \mathcal{S}^*$ be such that $[c = a$ and $[v = ae$. Hence $ew] = ub$ and $ce = v$, whence $\#([)(c) \geq \#(])(c)$ by Lemma 1.27. Thus

$$\begin{aligned}
\#([)(c) + \#([)(e) &= \#([)(v) \\
&= \#(])(v) \\
&= \#(])(c) + \#(])(e) \\
&\leq \#([)(c) + \#(])(e)
\end{aligned}$$

   so that $\#([)(b) \leq \#(])(b)$. Since if $\mathrm{lh}(e) < \mathrm{lh}(u)$ we have that $u = ef$ for some $f \sqsubseteq w]$ but $e \neq u$, whence $\#(])(e) < \#([)(e)$ by 1.27, we conclude that $\mathrm{lh}(e) \geq \mathrm{lh}(u)$. Hence $u \sqsubseteq e$, so that $e = ud$ for some $d \in \mathcal{S}^*$. Now $v = ce = cud$, $a = [c$ and $b = dw]$.

If $\mathrm{lh}(v)+1 \leq \mathrm{lh}(a) \leq \mathrm{lh}(v)+\mathrm{lh}(w)+2$ then there is a $c \in \mathcal{S}^*$ such that $[vc = a$, $c \sqsubseteq w]$ and $\mathrm{lh}(c) \leq \mathrm{lh}(w)+1$. Thus $[vw] = aub = [vcub$ so that $cub = w]$. If $\mathrm{lh}(c) = \mathrm{lh}(w)+1$, $\mathrm{lh}(a) = \mathrm{lh}([vw])$, contradicting the fact that $\mathrm{lh}(u) > 0$. Hence $\mathrm{lh}(c) \leq \mathrm{lh}(w)$, so that $c \sqsubseteq w$. Let $e \in \mathcal{S}^*$ be such that $ce = w$, that is $ce] = w] = cub$, whence $ub = e]$. Since $\#([)(c) \geq \#(])(c)$, we have that $\#([)(e) \leq \#(])(e)$ like above. If $\mathrm{lh}(e) < \mathrm{lh}(u)$, so that $e \sqsubseteq u$ but $e \neq u$, $\#([)(e) > \#(])(e)$ which is a contradiction; hence $\mathrm{lh}(e) \geq \mathrm{lh}(u)$, i.e. $u \sqsubseteq e$. Let $d \in \mathcal{S}^*$ be such that $ud = e$. Then $cub = ce] = cud]$ so that $b = d]$.

□

2.10 Lemma. *For any $\alpha, \beta \in \mathcal{T}$, $\varphi_\alpha, \psi_\beta, \vartheta_\beta \in W$ and $a, b \in \mathcal{S}^*$ such that*

$$\varphi = a\psi b$$

and ψ is not a single character string, $a\vartheta b \in W_\alpha$.

Proof. By induction on φ.

- If $a\psi b = s$ for some $s \in \mathcal{B}$ and $a, b \in \mathcal{S}^*$, then $\psi = s$ is a single character string, and there is nothing to prove.
- Assume $a\psi b = [\lambda x\eta]$ for some $a, b \in \mathcal{S}^*$, $\gamma, \delta \in \mathcal{T}$, $x \in X_\delta$, and $\eta \in W_\gamma$ such that if

$$\eta = c\psi d$$

for some $c, d \in \mathcal{S}^*$, then $c\vartheta d \in W_\gamma$. Thus $\psi = [\lambda x\eta]$ or there are $c, d \in \mathcal{S}^*$ such that $c\psi d = \eta$ and $a = [\lambda xc$, $b = d]$ by 2.9 since $\psi \neq x$. In the first case $\beta = \gamma\delta$, $a = b = \diamond$ and $\vartheta \in W_{\gamma\delta}$. In the second case $c\vartheta d \in W_\gamma$ by the induction hypothesis, whereby

$$a\vartheta b = [\lambda x c\vartheta d] \in W_{\gamma\delta}$$

as desired.

- Suppose $a\psi b = [\eta\xi]$ for some $a, b \in \mathcal{S}^*$, $\gamma, \delta \in \mathcal{T}$ and $\eta_{\gamma\delta}, \xi_\delta \in W$ such that if

$$\eta = c\psi d$$

for some $c, d \in \mathcal{S}^*$, then $c\vartheta d \in W_{\gamma\delta}$, and if

$$\xi = e\psi f$$

for some $e, f \in \mathcal{S}^*$, then $e\vartheta f \in W_\delta$. By Lemma 2.9, $\psi = [\eta\xi]$ or there are $c, d \in \mathcal{S}^*$ such that $c\psi d = \eta$, $a = [c$ and $b = d\xi]$, or $c\psi d = \xi$, $a = [\eta c$ and $b = d]$.

If $\psi = [\eta\xi]$, $\beta = \gamma\delta$ and $\vartheta \in W_{\gamma\delta}$ by assumption.

If there are $\mathsf{c}, \mathsf{d} \in \mathcal{S}^*$ such that $\mathsf{c}\psi\mathsf{d} = \eta$, $\mathsf{a} = [\mathsf{c}$ and $\mathsf{b} = \mathsf{d}\xi]$, then

$$\mathsf{c}\vartheta\mathsf{d} \in W_{\gamma\delta}$$

by induction hypothesis, and

$$\mathsf{a}\vartheta\mathsf{b} = [\mathsf{c}\vartheta\mathsf{d}\xi] \in W_{\gamma}.$$

If on the other hand $\mathsf{c}\psi\mathsf{d} = \xi$, $\mathsf{a} = [\eta\mathsf{c}$ and $\mathsf{b} = \mathsf{d}]$ for some $\mathsf{c}, \mathsf{d} \in \mathcal{S}^*$, we have that

$$\mathsf{c}\vartheta\mathsf{d} \in W_{\delta}$$

by induction hypothesis. Hence

$$\mathsf{a}\vartheta\mathsf{b} = [\eta\mathsf{c}\vartheta\mathsf{d}] \in W_{\gamma}$$

as desired.

□

Variables and Substitutions

2.11 Definition (Free and bound variables)**.** The sets $\mathsf{VF}(\mathsf{u}) \subseteq \mathcal{X}$ and $\mathsf{VB}(\mathsf{u}) \subseteq \mathcal{X}$ of *free* and *bound* variables in $\mathsf{u} \in \mathsf{U}$, are defined by structural recursion as follows:

- For every $\mathsf{s} \in \mathcal{B}$,

$$\mathsf{VF}(\mathsf{s}) = \begin{cases} \{\mathsf{s}\} & \text{if } \mathsf{s} \in \mathcal{X} \\ \varnothing & \text{otherwise} \end{cases}, \qquad \mathsf{VB}(\mathsf{s}) = \varnothing.$$

- For every $\nu \in \mathsf{V}$ and $\boldsymbol{x} \in \mathcal{X}$

$$\mathsf{VF}([\lambda \boldsymbol{x} \nu]) = \mathsf{VF}(\nu) \setminus \{\boldsymbol{x}\}$$

and

$$\mathsf{VB}([\lambda \boldsymbol{x} \nu]) = \mathsf{VB}(\nu) \cup \{\boldsymbol{x}\}.$$

- For all $\nu, w \in \mathsf{U}$

$$\mathsf{VF}([\nu w]) = \mathsf{VF}(\nu) \cup \mathsf{VF}(w)$$

and

$$\mathsf{VB}([\nu w]) = \mathsf{VB}(\nu) \cup \mathsf{VB}(w).$$

Given $\mathsf{u} \in \mathsf{U}$, we say that $\boldsymbol{x}$ is *free* in u if $\boldsymbol{x} \in \mathsf{VF}(\mathsf{u})$, and that $\boldsymbol{x}$ is *bound* in u if $\boldsymbol{x} \in \mathsf{VB}(\mathsf{u})$.

2.12 Lemma. *For all* $u \in \mathsf{U}$*,* $\mathsf{VF}(u) \cup \mathsf{VB}(u) = \{x \in \mathcal{X} \,|\, x \preceq u\}$*.*

Proof. By induction on u.

- Let $s \in \mathcal{B}$. If $s \in \mathcal{X}$

$$\mathsf{VF}(s) \cup \mathsf{VB}(s) = \{s\} = \{x \in \mathcal{X} \,|\, x \preceq s\}$$

  and otherwise

$$\mathsf{VF}(s) \cup \mathsf{VB}(s) = \varnothing = \{x \in \mathcal{X} \,|\, x \preceq s\}.$$

- Consider $[\lambda y v] \in \mathsf{U}$, where $y \in \mathsf{X}$ and $v \in \mathsf{U}$ is such that $\mathsf{VF}(v) \cup \mathsf{VB}(v) = \{x \in \mathcal{X} \,|\, x \preceq v\}$. Then

$$\begin{aligned} \mathsf{VF}([\lambda y v]) \cup \mathsf{VB}([\lambda y v]) &= (\mathsf{VF}(v) \setminus \{y\}) \cup (\mathsf{VB}(v) \cup \{y\}) \\ &= (\mathsf{VF}(v) \cup \mathsf{VB}(v)) \cup \{y\} \\ &= \{x \in \mathcal{X} \,|\, x \preceq v\} \cup \{y\} \\ &= \{x \in \mathcal{X} \,|\, x \preceq yv\} \\ &= \{x \in \mathcal{X} \,|\, x \preceq [\lambda y v]\}. \end{aligned}$$

- Finally consider $[vw] \in \mathsf{U}$, where $v, w \in \mathsf{U}$ are such that $\mathsf{VF}(v) \cup \mathsf{VB}(v) = \{x \in \mathcal{X} \,|\, x \preceq v\}$ and $\mathsf{VF}(w) \cup \mathsf{VB}(w) = \{x \in \mathcal{X} \,|\, x \preceq w\}$.

$$\begin{aligned} \mathsf{VF}([vw]) \cup \mathsf{VB}([vw]) &= \mathsf{VF}(v) \cup \mathsf{VF}(w) \cup \mathsf{VB}(v) \cup \mathsf{VB}(w) \\ &= \{x \in \mathcal{X} \,|\, x \preceq v\} \cup \{x \in \mathcal{X} \,|\, x \preceq w\} \\ &= \{x \in \mathcal{X} \,|\, x \preceq vw\} \\ &= \{x \in \mathcal{X} \,|\, x \preceq [vw]\}. \end{aligned}$$

Hence $\mathsf{VF}(u) \cup \mathsf{VB}(u) = \{x \in \mathcal{X} \,|\, x \preceq u\}$ for all $u \in \mathsf{U}$, as desired. □

2.13 Definition (Closed formulas)**.** A formula is called *closed* if it has no free variables. The set of all closed well-formed formulas is denoted by $\overline{W}$. Furthermore the notation $\overline{W}_\alpha = \{\varphi \in \overline{W} \,|\, \mathsf{T}(\varphi) = \alpha\}$ will be used for any $\alpha \in \mathcal{T}$.

Note that a variable may be both free and bound in a specific formula u, i.e. there is no claim that $\mathsf{VF}(u) \cap \mathsf{VB}(u) = \varnothing$. However, in many cases this would be a preferable situation. We will therefore define substitution operators SF and SB, which replaces only free or bound variables, respectively.

Remark 9. Note that the substitution operators below are well-defined. The proofs of this would be exactly the recursive definitions with induction hypotheses added.

2.14 Definition (Substitution of free variables)**.** The function SF : $\mathcal{X} \longrightarrow (\mathsf{U}^{\mathsf{U}})^{\mathsf{U}}$ is defined via structural recursion as follows, for all $x \in \mathcal{X}$ and $u \in \mathsf{U}$:

- For any $s \in \mathcal{B}$ we define

$$\mathrm{SF}(\mathbf{x})(u)(s) = \begin{cases} u & \text{if } s = \mathbf{x} \\ s & \text{otherwise.} \end{cases}$$

- For any $v \in U$ and $\mathbf{y} \in \mathcal{X}$ we define

$$\mathrm{SF}(\mathbf{x})(u)([\lambda\mathbf{y}v]) = \begin{cases} [\lambda\mathbf{y}v] & \text{if } \mathbf{y} = \mathbf{x} \\ [\lambda\mathbf{y}\,\mathrm{SF}(\mathbf{x})(u)(v)] & \text{otherwise.} \end{cases}$$

- For any $v, w \in U$ define

$$\mathrm{SF}(\mathbf{x})(u)([vw]) = [\mathrm{SF}(\mathbf{x})(u)(v)\,\mathrm{SF}(\mathbf{x})(u)(w)].$$

2.15 Definition (Substitution of bound variables)**.** The function $\mathrm{SB} : \mathcal{X} \longrightarrow (U^U)^{\mathcal{X}}$ is defined via structural recursion as follows, for all $\mathbf{x}, \mathbf{y} \in \mathcal{X}$:

- For any $s \in \mathcal{B}$ we define

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})(s) = s.$$

- For any $u \in U$ and $\mathbf{z} \in \mathcal{X}$ we define

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda\mathbf{z}u]) = \begin{cases} [\lambda\mathbf{y}\,\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(u))] & \text{if } \mathbf{z} = \mathbf{x} \\ [\lambda\mathbf{z}\,\mathrm{SB}(\mathbf{x})(\mathbf{y})(u)] & \text{otherwise.} \end{cases}$$

- For any $u, v \in U$ define

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})([uv]) = [\mathrm{SB}(\mathbf{x})(\mathbf{y})(u)\,\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)].$$

2.16 Proposition. *Let $\alpha \in \mathcal{T}$, $\mathbf{x}, \mathbf{y} \in X_\alpha$ and $\psi \in W_\alpha$.*

1. $\mathrm{SF}(\mathbf{x})(\psi)(\varphi) \in W_{\mathsf{T}(\varphi)}$ *for every* $\varphi \in W$.
2. $\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi) \in W_{\mathsf{T}(\varphi)}$ *for every* $\varphi \in W$.

Proof. By induction on $\varphi \in W$.

1. - Let $\beta \in \mathcal{T}$ and $s \in \mathcal{B}$ be such that $\mathsf{T}(s) = \beta$. If $s = \mathbf{x}$, then $\alpha = \beta$ and

   $$\mathrm{SF}(\mathbf{x})(\psi)(s) = \psi \in W_\alpha = W_\beta.$$

   If on the other hand $s \neq \mathbf{x}$, then

   $$\mathrm{SF}(\mathbf{x})(\psi)(s) = s \in W_\beta.$$

   In both cases, $\mathrm{SF}(\mathbf{x})(\psi)(s) \in W_\beta$.

- Now let $\beta, \gamma \in \mathcal{T}$ and $\mathbf{y} \in X_\gamma$. Assume $\xi \in W_\beta$ is such that $\mathrm{SF}(\mathbf{x})(\psi)(\xi) \in W_\beta$ and consider $[\lambda \mathbf{y} \xi] \in W_{\beta\gamma}$. If $\mathbf{y} = \mathbf{x}$ then $\beta\gamma = \beta\alpha$ and

$$\mathrm{SF}(\mathbf{x})(\psi)([\lambda \mathbf{y} \xi]) = [\lambda \mathbf{y} \xi] \in W_{\beta\alpha},$$

while if $\mathbf{y} \neq \mathbf{x}$ then

$$\mathrm{SF}(\mathbf{x})(\psi)([\lambda \mathbf{y} \xi]) = [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\psi)(\xi)] \in W_{\beta\gamma}.$$

Hence $\mathrm{SF}(\mathbf{x})(\psi)([\lambda \mathbf{y} \xi]) \in W_{\beta\gamma}$.

- Let $\beta, \gamma \in \mathcal{T}$ and assume $\xi \in W_{\beta\gamma}$ and $\vartheta \in W_\gamma$ are such that $\mathrm{SF}(\mathbf{x})(\psi)(\xi) \in W_{\beta\gamma}$ and $\mathrm{SF}(\mathbf{x})(\psi)(\vartheta) \in W_\gamma$. Then $[\xi\vartheta] \in W_\beta$ and

$$\mathrm{SF}(\mathbf{x})(\psi)([\xi\vartheta]) = [\mathrm{SF}(\mathbf{x})(\psi)(\xi)\, \mathrm{SF}(\mathbf{x})(\psi)(\vartheta)] \in W_\beta,$$

as desired.

Hence $\mathrm{SF}(\mathbf{x})(\psi)(\varphi) \in W_\beta$ for all $\beta \in \mathcal{T}$ and $\varphi \in W_\beta$.

2. - Let $\beta \in \mathcal{T}$ and $s \in \mathcal{B}$ be such that $t(s) = \beta$. Then

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})(s) = s \in W_\beta.$$

- Now let $\beta, \gamma \in \mathcal{T}$ and $\mathbf{z} \in X_\gamma$. Assume $\xi \in W_\beta$ is such that $\mathrm{SB}(\mathbf{x})(\mathbf{y})(\xi) \in W_\beta$ and consider $[\lambda \mathbf{z} \xi] \in W_{\beta\gamma}$. If $\mathbf{z} = \mathbf{x}$ then $\beta\gamma = \beta\alpha$ and

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z} \xi]) = [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(\xi))] \in W_{\beta\alpha},$$

by above, while if $\mathbf{z} \neq \mathbf{x}$ then

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z} \xi]) = [\lambda \mathbf{z}\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(\xi)] \in W_{\beta\gamma}.$$

Hence $\mathrm{SB}(\mathbf{x})(\varphi)([\lambda \mathbf{y} \xi]) \in W_{\beta\gamma}$.

- Finally let $\beta, \gamma \in \mathcal{T}$ and assume $\xi \in W_{\beta\gamma}$ and $\vartheta \in W_\gamma$ are such that $\mathrm{SB}(\mathbf{x})(\mathbf{y})(\xi) \in W_{\beta\gamma}$ and $\mathrm{SB}(\mathbf{x})(\mathbf{y})(\vartheta) \in W_\gamma$. Then $[\xi\vartheta] \in W_\beta$ and

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})([\xi\vartheta]) = [\mathrm{SB}(\mathbf{x})(\mathbf{y})(\xi)\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(\vartheta)] \in W_\beta,$$

as desired.

Hence $\mathrm{SF}(\mathbf{x})(\mathbf{y})(\varphi) \in W_\beta$ for all $\beta \in \mathcal{T}$ and $\varphi \in W_\beta$.

□

2.17 Lemma. *Let $\alpha \in \mathcal{T}$.*

1. *For any finite set of variables $Y \subseteq X_\alpha$, there is an $\mathbf{x} \in X_\alpha$ such that $\mathbf{x} \notin Y$.*
2. *For any finite set R of formulas there is an $\mathbf{x} \in X_\alpha$ such that $\mathbf{x}$ does not occur in any formula of R.*

Proof. 1. Immediate since X_α is countable.

2. Let $\mathsf{Y} = \{\mathbf{y} \in \mathsf{X}_\alpha \,|\, \exists \mathsf{u} \in \mathsf{R} : \mathbf{y} \preceq \mathsf{u}\}$. Since R is finite and every $\mathsf{u} \in \mathsf{R}$ has finite length, Y is finite. Hence there is an $\mathbf{x} \notin \mathsf{Y}$, by above, whereby $\mathbf{x} \not\preceq \mathsf{u}$ for all $\mathsf{u} \in \mathsf{R}$. □

Below will follow a number of lemmas establishing the relationship between the different kinds of substitution, which will be seen to behave more or less as we expect. They will also be useful when proving properties of the syntax in section 3.

2.18 Lemma. *Let $\mathbf{x} \in \mathcal{X}$ and $\mathsf{u} \in \mathsf{U}$.*

1. *If $\mathbf{x}$ is free in u, then $\mathsf{VF}(\mathsf{u}) \setminus \{\mathbf{x}\} \subseteq \mathsf{VF}(\mathrm{SF}(\mathbf{x})(\nu)(\mathsf{u})) \subseteq (\mathsf{VF}(\mathsf{u}) \setminus \{\mathbf{x}\}) \cup \mathsf{VF}(\nu)$ for every $\nu \in \mathsf{U}$.*
2. *If $\mathbf{x}$ is not free in u, then $\mathrm{SF}(\mathbf{x})(\nu)(\mathsf{u}) = \mathsf{u}$ for every $\nu \in \mathsf{U}$.*
3. *Let $\mathbf{y} \in \mathcal{X}$ and $\nu \in \mathsf{U}$ be such that $\mathbf{x}$ is not bound in ν. Then $\mathbf{x}$ is bound in u if and only if $\mathbf{x}$ is bound in $\mathrm{SF}(\mathbf{y})(\nu)(\mathsf{u})$.*

Proof. The proofs are by structural induction on u.

1. Let $\nu \in \mathsf{U}$.
   - Let $s \in \mathcal{B}$. If $\mathbf{x}$ is free in s then $s = \mathbf{x}$. Thus we have that
   $$\mathrm{SF}(\mathbf{x})(\nu)(s) = \nu.$$
   Hence $\mathsf{VF}(\mathrm{SF}(\mathbf{x})(\nu)(s)) = \mathsf{VF}(\nu) = (\mathsf{VF}(s) \setminus \{\mathbf{x}\}) \cup \mathsf{VF}(\nu)$.
   - Now let $w \in \mathsf{U}$ be such that if $\mathbf{x}$ is free in w then
   $$\mathsf{VF}(w) \setminus \{\mathbf{x}\} \subseteq \mathsf{VF}(\mathrm{SF}(\mathbf{x})(\nu)(w)) \subseteq (\mathsf{VF}(w) \setminus \{\mathbf{x}\}) \cup \mathsf{VF}(\nu),$$
   and let $\mathbf{y} \in \mathcal{X}$. Additionally, assume that $\mathbf{x}$ is free in $[\lambda \mathbf{y} w]$, that is $\mathbf{x}$ is free in w and $\mathbf{y} \neq \mathbf{x}$. Hence
   $$\mathrm{SF}(\mathbf{x})(\nu)([\lambda \mathbf{y} w]) = [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\nu)(w)]$$
   whereby
   $$\begin{aligned}\mathsf{VF}(\mathrm{SF}(\mathbf{x})(\nu)([\lambda \mathbf{y} w])) &= \mathsf{VF}([\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\nu)(w)]) \\ &= \mathsf{VF}(\mathrm{SF}(\mathbf{x})(\nu)(w)) \setminus \{\mathbf{y}\} \\ &\subseteq ((\mathsf{VF}(w) \setminus \{\mathbf{x}\}) \cup \mathsf{VF}(\nu)) \setminus \{\mathbf{y}\} \\ &= ((\mathsf{VF}(w) \setminus \{\mathbf{y}\}) \setminus \{\mathbf{x}\}) \cup (\mathsf{VF}(\nu) \setminus \{\mathbf{y}\}) \\ &\subseteq (\mathsf{VF}([\lambda \mathbf{y} w]) \setminus \{\mathbf{x}\}) \cup \mathsf{VF}(\nu)\end{aligned}$$
   and
   $$\begin{aligned}\mathsf{VF}(\mathrm{SF}(\mathbf{x})(\nu)([\lambda \mathbf{y} w])) &= \mathsf{VF}(\mathrm{SF}(\mathbf{x})(\nu)(w)) \setminus \{\mathbf{y}\} \\ &\supseteq (\mathsf{VF}(w) \setminus \{\mathbf{x}\}) \setminus \{\mathbf{y}\} \\ &= (\mathsf{VF}(w) \setminus \{\mathbf{y}\}) \setminus \{\mathbf{x}\} \\ &= \mathsf{VF}([\lambda \mathbf{y} w]) \setminus \{\mathbf{x}\}\end{aligned}$$

- Finally let $w, q \in U$ be such that if x is free in a then

$$\mathsf{VF}(a) \setminus \{x\} \subseteq \mathsf{VF}(\mathrm{SF}(x)(v)(a)) \subseteq (\mathsf{VF}(a) \setminus \{x\}) \cup \mathsf{VF}(v)$$

for both $a \in \{w, q\}$. Then

$$\begin{aligned}
\mathsf{VF}(\mathrm{SF}(x)(v)([wq])) &= \mathsf{VF}([\mathrm{SF}(x)(v)(w)\,\mathrm{SF}(x)(v)(q)]) \\
&= \mathsf{VF}(\mathrm{SF}(x)(v)(w)) \cup \mathsf{VF}(\mathrm{SF}(x)(v)(q)) \\
&\subseteq ((\mathsf{VF}(w) \setminus \{x\}) \cup \mathsf{VF}(v)) \cup ((\mathsf{VF}(q) \setminus \{x\}) \cup \mathsf{VF}(v)) \\
&= (\mathsf{VF}(w) \setminus \{x\}) \cup (\mathsf{VF}(q) \setminus \{x\}) \cup \mathsf{VF}(v) \\
&= ((\mathsf{VF}(w) \cup \mathsf{VF}(q)) \setminus \{x\}) \cup \mathsf{VF}(v) \\
&= (\mathsf{VF}([wq]) \setminus \{x\}) \cup \mathsf{VF}(v)
\end{aligned}$$

and

$$\begin{aligned}
\mathsf{VF}(\mathrm{SF}(x)(v)([wq])) &= \mathsf{VF}(\mathrm{SF}(x)(v)(w)) \cup \mathsf{VF}(\mathrm{SF}(x)(v)(q)) \\
&\supseteq (\mathsf{VF}(w) \setminus \{x\}) \cup (\mathsf{VF}(q) \setminus \{x\}) \\
&= (\mathsf{VF}(w) \cup \mathsf{VF}(q)) \setminus \{x\} \\
&= \mathsf{VF}([wq]) \setminus \{x\}.
\end{aligned}$$

Hence $\mathsf{VF}(u) \setminus \{x\} \subseteq \mathsf{VF}(\mathsf{SF}(x)(v)(u)) \subseteq (\mathsf{VF}(u) \setminus \{x\}) \cup \mathsf{VF}(v)$.

2. Let $v \in U$.
   - Let $s \in \mathcal{B}$. If x is not free in s, then $x \neq s$. Hence

$$\mathrm{SF}(x)(v)(s) = s.$$

   - Let $z \in \mathcal{X}$ and $w \in U$ be such that if x is not free in w, then $\mathrm{SF}(x)(v)(w) = w$. Now, if x is not free in $[\lambda z w]$, either $x = z$ or x is not free in w, by 2.11. Hence

$$\begin{aligned}
\mathrm{SF}(x)(v)([\lambda z w]) &= \begin{cases} [\lambda z w] & \text{if } z = x \\ [\lambda z\, \mathrm{SF}(x)(v)(w)] & \text{otherwise} \end{cases} \\
&= \begin{cases} [\lambda z w] & \text{if } z = x \\ [\lambda z w] & \text{otherwise} \end{cases} \\
&= [\lambda z w].
\end{aligned}$$

   - Finally, let $w, q \in U$ be such that, if x is not free in a, then $\mathrm{SF}(x)(v)(a) = a$ for all $a \in \{w, q\}$. Moreover, assume that x is not free in $[wq]$. By 2.11, x is free in neither w nor q. Thus

$$\mathrm{SF}(x)(v)([wq]) = [\mathrm{SF}(x)(v)(w)\,\mathrm{SF}(y)(v)(q)] = [wq].$$

Hence $\mathrm{SF}(x)(v)(u) = u$.

3. Let $y \in \mathcal{X}$ and $v \in U$ be such that x is not bound in v.
   - Let $s \in \mathcal{B}$. Clearly, x is not bound in s. Furthermore

   $$\mathrm{SF}(y)(v)(s) = \begin{cases} v & \text{if } s = y \\ s & \text{otherwise} \end{cases}$$

   so that x is not bound in $\mathrm{SF}(y)(v)(s)$ either. Hence x is bound in s if and only if x is bound in $\mathrm{SF}(y)(v)(s)$
   - Let $z \in \mathcal{X}$ and $w \in U$ be such that x is bound in w if and only if x is bound in $\mathrm{SF}(y)(v)(w)$.

     Assume that x is bound in $[\lambda z w]$. Then $z = x$ or x is bound in w, so that $z = x$ or x is bound in $\mathrm{SF}(y)(v)(w)$, by hypothesis. Thus x is bound in $[\lambda z\, \mathrm{SF}(y)(v)(w)]$. Furthermore

   $$\mathrm{SF}(y)(v)([\lambda z w]) = \begin{cases} [\lambda z w] & \text{if } z = y \\ [\lambda z\, \mathrm{SF}(y)(v)(w)] & \text{otherwise.} \end{cases}$$

     Hence x is bound in $\mathrm{SF}(y)(v)([\lambda z w])$.

     Conversely, if x is not bound in $[\lambda z w]$, then $z \neq x$ and x is not bound in w. Thus x is not bound in $\mathrm{SF}(y)(v)(w)$, by hypothesis, and thereby not in $[\lambda z\, \mathrm{SF}(y)(v)(w)]$ either. Hence x is not bound in $\mathrm{SF}(y)(v)([\lambda z w])$.
   - Let $w, q \in U$ be such that x is bound in a if and only if x is bound in $\mathrm{SF}(y)(v)(a)$, for all $a \in \{w, q\}$. Then x is bound in $[wq]$ if and only if x is bound in w or q, if and only if x is bound in $\mathrm{SF}(y)(v)(a)$ for some $a \in \{w, q\}$, if and only if x is bound in $\mathrm{SF}(y)(v)([wq])$.

   Hence x is bound in u if and only if x is bound in $\mathrm{SF}(y)(v)(u)$.

□

2.19 Lemma. *Let* $x, y \in \mathcal{X}$. *For every* $u \in U$

$$\mathrm{S}(x)(y)(u) = \mathrm{SF}(x)(y)(\mathrm{SB}(x)(y)(u)),$$

that is $\mathrm{S}(x)(y) = \mathrm{SF}(x)(y) \circ \mathrm{SB}(x)(y)$.

Proof. The proof is by structural induction on $u \in U$.

- Let $s \in \mathcal{B}$. Now

$$\begin{aligned} \mathrm{SF}(x)(y)(\mathrm{SB}(x)(y)(s)) &= \mathrm{SF}(x)(y)(s) \\ &= \begin{cases} y & \text{if } s = x \\ s & \text{otherwise} \end{cases} \\ &= \mathrm{S}(x)(y)(s). \end{aligned}$$

- Let $v \in \mathsf{U}$ be such that

$$\mathrm{S}(\mathbf{x})(\mathbf{y})(v) = \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))$$

and $\mathbf{z} \in \mathcal{X}$.

$$\begin{aligned}
&\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z} v])) \\
&= \begin{cases} \mathrm{SF}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))]) & \text{if } \mathbf{z} = \mathbf{x} \\ \mathrm{SF}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z}\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(v)]) & \text{otherwise} \end{cases} \\
&= \begin{cases} [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))] & \text{if } \mathbf{z} = \mathbf{x} = \mathbf{y} \\ [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)))] & \text{if } \mathbf{z} = \mathbf{x} \neq \mathbf{y} \\ [\lambda \mathbf{z}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))] & \text{otherwise} \end{cases} \\
&= \begin{cases} [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))] & \text{if } \mathbf{z} = \mathbf{x} = \mathbf{y} \\ [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))] & \text{if } \mathbf{z} = \mathbf{x} \neq \mathbf{y} \\ [\lambda \mathbf{z}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))] & \text{otherwise} \end{cases} \\
&= \begin{cases} [\lambda \mathbf{y}\, \mathrm{S}(\mathbf{x})(\mathbf{y})(v)] & \text{if } \mathbf{z} = \mathbf{x} \\ [\lambda \mathbf{z}\, \mathrm{S}(\mathbf{x})(\mathbf{y})(v)] & \text{otherwise} \end{cases} \\
&= \begin{cases} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} v]) & \text{if } \mathbf{z} = \mathbf{x} \\ \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z} v]) & \text{otherwise} \end{cases} \\
&= \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z} v])
\end{aligned}$$

since $\mathbf{x}$ is not free in $\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))$ if $\mathbf{x} \neq \mathbf{y}$, by 2.18.

- Finally let $v, w \in \mathsf{U}$ be such that

$$\mathrm{S}(\mathbf{x})(\mathbf{y})(\mathsf{q}) = \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathsf{q}))$$

for all $\mathsf{q} \in \{v, w\}$. Then

$$\begin{aligned}
&\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})([vw])) \\
&= [\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w))] \\
&= [\mathrm{S}(\mathbf{x})(\mathbf{y})(v)\, \mathrm{S}(\mathbf{x})(\mathbf{y})(w)] \\
&= \mathrm{S}(\mathbf{x})(\mathbf{y})([vw]).
\end{aligned}$$

Hence

$$\mathrm{S}(\mathbf{x})(\mathbf{y})(\mathsf{u}) = \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathsf{u}))$$

for every $\mathsf{u} \in \mathsf{U}$. □

2.20 Corollary. *1. For all* $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ *and* $\mathsf{u} \in \mathsf{U}$, $\mathrm{S}(\mathbf{x})(\mathbf{y})(\mathsf{u}) \in \mathsf{U}$.

2. *For all* $\alpha, \beta \in \mathcal{T}$*,* $\mathbf{x}, \mathbf{y} \in X_\alpha$ *and* $\varphi \in W_\beta$*,* $S(\mathbf{x})(\mathbf{y})(\varphi) \in W_\beta$*.*

2.21 Corollary. *For all* $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ *and* $u \in U$*,*

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} u]) = [\lambda \mathbf{y}\, \mathrm{S}(\mathbf{x})(\mathbf{y})(u)].$$

2.22 Lemma. *Let* $\mathbf{x} \in \mathcal{X}$ *and* $u \in U$*.*

1. *For any* $\mathbf{y} \in \mathcal{X}$ *distinct from* $\mathbf{x}$*,*

$$\mathsf{VB}(u) \setminus \{\mathbf{x}\} \subseteq \mathsf{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(u)) \subseteq (\mathsf{VB}(u) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\}.$$

2. *For any* $\mathbf{y} \in \mathcal{X}$*,*

$$\mathsf{VF}(u) \setminus \{\mathbf{y}\} \subseteq \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(u)) \subseteq \mathsf{VF}(u)$$

3. $\mathrm{SF}(\mathbf{x})(\mathbf{x})(u) = u$ *and* $\mathrm{SB}(\mathbf{x})(\mathbf{x})(u) = u$*.*

4. $\mathrm{SF}(\mathbf{z})(\mathbf{y})(\mathrm{SF}(\mathbf{x})(\mathbf{z})(u)) = \mathrm{SF}(\mathbf{x})(\mathbf{y})(u)$ *and* $\mathrm{SB}(\mathbf{z})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{z})(u)) = \mathrm{SB}(\mathbf{x})(\mathbf{y})(u)$ *for all* $\mathbf{y}, \mathbf{z} \in \mathcal{X}$ *such that* $\mathbf{z}$ *does not occur in* u*.*

5. *If* $\mathbf{x}$ *is not bound in* u*, then*

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})(u) = u$$

for all $\mathbf{y} \in \mathcal{X}$*, and*

$$\mathrm{SF}(\mathbf{x})(v)(u) = \mathrm{S}(\mathbf{x})(v)(u)$$

for all $v \in U$*.*

Proof. The proofs are by induction on u.

1. Let $\mathbf{y} \in \mathcal{X} \setminus \{\mathbf{x}\}$.
   - Let $s \in \mathcal{B}$. By definition, $s = \mathrm{SB}(\mathbf{x})(\mathbf{y})(s)$, whereby

$$\mathsf{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(s)) = \varnothing = \mathsf{VB}(s) \setminus \{\mathbf{x}\} \subseteq (\mathsf{VB}(s) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\}.$$

   - Now consider $[\lambda \mathbf{z} v] \in U$, where $v \in U$ is such that

$$\mathsf{VB}(v) \setminus \{\mathbf{x}\} \subseteq \mathsf{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \subseteq (\mathsf{VB}(v) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\},$$

   and $\mathbf{z} \in \mathcal{X}$. Then

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z} v]) = \begin{cases} [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))] & \text{if } \mathbf{z} = \mathbf{x} \\ [\lambda \mathbf{z}\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(v)] & \text{otherwise.} \end{cases}$$

By 2.18 above, $\mathrm{VB}(\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))) = \mathrm{VB}((\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)))$. Thus, if $\mathbf{z} = \mathbf{x}$,

$$\begin{aligned}\mathrm{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z} v])) &= \mathrm{VB}([\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))]) \\ &= \mathrm{VB}(\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))) \cup \{\mathbf{y}\} \\ &= \mathrm{VB}((\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))) \cup \{\mathbf{y}\}\end{aligned}$$

and

$$\begin{aligned}\mathrm{VB}([\lambda \mathbf{z} v]) \setminus \{\mathbf{x}\} &= (\mathrm{VB}(v) \cup \{\mathbf{z}\}) \setminus \{\mathbf{x}\} \\ &= \mathrm{VB}(v) \setminus \{\mathbf{x}\} \\ &\subseteq \mathrm{VB}((\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))) \\ &\subseteq (\mathrm{VB}(v) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\} \\ &= ((\mathrm{VB}(v) \cup \{\mathbf{z}\}) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\} \\ &= (\mathrm{VB}([\lambda \mathbf{z} v]) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\}.\end{aligned}$$

Likewise, if $\mathbf{z} \neq \mathbf{x}$

$$\begin{aligned}\mathrm{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z} v])) &= \mathrm{VB}([\lambda \mathbf{z}\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(v)]) \\ &= \mathrm{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \cup \{\mathbf{z}\}\end{aligned}$$

and

$$\begin{aligned}\mathrm{VB}([\lambda \mathbf{z} v]) \setminus \{\mathbf{x}\} &= (\mathrm{VB}(v) \cup \{\mathbf{z}\}) \setminus \{\mathbf{x}\} \\ &= (\mathrm{VB}(v) \setminus \{\mathbf{x}\}) \cup \{\mathbf{z}\} \\ &\subseteq \mathrm{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \cup \{\mathbf{z}\} \\ &\subseteq ((\mathrm{VB}(v) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\}) \cup \{\mathbf{z}\} \\ &= ((\mathrm{VB}(v) \cup \{\mathbf{z}\}) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\} \\ &= (\mathrm{VB}([\lambda \mathbf{z} v]) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\}.\end{aligned}$$

Hence

$$\mathrm{VB}([\lambda \mathbf{z} v]) \setminus \{\mathbf{x}\} \subseteq \mathrm{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z} v])) \subseteq (\mathrm{VB}([\lambda \mathbf{z} v]) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\}$$

in either case, as desired.

- Finally consider $[vw] \in \mathfrak{U}$, where $v, w \in \mathfrak{U}$ are such that

$$\mathrm{VB}(q) \setminus \{\mathbf{x}\} \subseteq \mathrm{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(q)) \subseteq (\mathrm{VB}(q) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\},$$

for both $q \in \{v, w\}$. Then

$$\begin{aligned}\mathrm{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})([vw])) &= \mathrm{VB}([\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(w)]) \\ &= \mathrm{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \cup \mathrm{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w))\end{aligned}$$

and

$$\begin{aligned}
\mathsf{VB}([vw]) \setminus \{\mathbf{x}\} &= (\mathsf{VB}(v) \cup \mathsf{VB}(w)) \setminus \{\mathbf{x}\} \\
&= (\mathsf{VB}(v) \setminus \{\mathbf{x}\}) \cup (\mathsf{VB}(w) \setminus \{\mathbf{x}\}) \\
&\subseteq \mathsf{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \cup \mathsf{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w)) \\
&\subseteq ((\mathsf{VB}(v) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\}) \cup ((\mathsf{VB}(w) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\}) \\
&= ((\mathsf{VB}(v) \cup \mathsf{VB}(w)) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\} \\
&= (\mathsf{VB}([vw]) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\}
\end{aligned}$$

that is

$$\mathsf{VB}([vw]) \setminus \{\mathbf{x}\} \subseteq \mathsf{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})([vw])) \subseteq (\mathsf{VB}([vw]) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\}$$

as desired.

Hence

$$\mathsf{VB}(\mathsf{u}) \setminus \{\mathbf{x}\} \subseteq \mathsf{VB}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathsf{u})) \subseteq (\mathsf{VB}(\mathsf{u}) \setminus \{\mathbf{x}\}) \cup \{\mathbf{y}\}$$

for all $\mathsf{u} \in \mathsf{U}$, as claimed.

2. Let $\mathbf{y} \in \mathcal{X}$.
   - Let $\mathsf{s} \in \mathcal{B}$. We know that

     $$\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathsf{s}) = \mathsf{s}$$

     so that

     $$\mathsf{VF}(\mathsf{s}) \setminus \{\mathbf{y}\} \subseteq \mathsf{VF}(\mathsf{s}) = \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathsf{s})).$$

   - Let $\mathbf{z} \in \mathcal{X}$ and $v \in \mathsf{U}$ be such that

     $$\mathsf{VF}(v) \setminus \{\mathbf{y}\} \subseteq \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \subseteq \mathsf{VF}(v).$$

     If $\mathbf{z} = \mathbf{x}$ then

     $$\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z} v]) = [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))]$$

     whence

     $$\begin{aligned}
     \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z} v])) &= \mathsf{VF}([\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))]) \\
     &= \mathsf{VF}(\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))) \setminus \{\mathbf{y}\} \\
     &\subseteq ((\mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \setminus \{\mathbf{x}\}) \cup \mathsf{VF}(\mathbf{y})) \setminus \{\mathbf{y}\} \\
     &= \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \setminus \{\mathbf{x}\} \setminus \{\mathbf{y}\} \\
     &\subseteq \mathsf{VF}(v) \setminus \{\mathbf{x}\} \setminus \{\mathbf{y}\} \\
     &= \mathsf{VF}([\lambda \mathbf{z} v]) \setminus \{\mathbf{y}\} \\
     &\subseteq \mathsf{VF}([\lambda \mathbf{z} v])
     \end{aligned}$$

and

$$\begin{aligned}\mathsf{VF}([\lambda z v]) \setminus \{\mathbf{y}\} &= \mathsf{VF}(v) \setminus \{\mathbf{x}\} \setminus \{\mathbf{y}\} \\ &\subseteq \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \setminus \{\mathbf{x}\} \setminus \{\mathbf{y}\} \\ &\subseteq \mathsf{VF}(\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v))) \setminus \{\mathbf{y}\} \\ &= \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda z v]))\end{aligned}$$

by Lemma 2.18 and the induction hypothesis (note that this shows that $\mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda z v])) = \mathsf{VF}([\lambda z v]) \setminus \{\mathbf{y}\}$ in this case).

If on the other hand $z \neq \mathbf{x}$ then

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda z v]) = [\lambda z\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(v)]$$

so that

$$\begin{aligned}\mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda z v])) &= \mathsf{VF}([\lambda z\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(v)]) \\ &= \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \setminus \{z\}.\end{aligned}$$

Since

$$\begin{aligned}\mathsf{VF}([\lambda z v]) \setminus \{\mathbf{y}\} &= \mathsf{VF}(z) \setminus \{z\} \setminus \{\mathbf{y}\} \\ &\subseteq \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \setminus \{z\} \\ &\subseteq \mathsf{VF}(v) \setminus \{z\} \\ &= \mathsf{VF}([\lambda z v])\end{aligned}$$

we have that $\mathsf{VF}([\lambda z v]) \setminus \{\mathbf{y}\} \subseteq \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda z v])) \subseteq \mathsf{VF}([\lambda z v])$.

- Let $v, w \in \mathsf{U}$ be such that

$$\mathsf{VF}(\mathsf{q}) \setminus \{\mathbf{y}\} \subseteq \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathsf{q})) \subseteq \mathsf{VF}(\mathsf{q})$$

for both $\mathsf{q} \in \{v, w\}$. Then

$$\begin{aligned}\mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})([vw])) &= \mathsf{VF}([\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(w)]) \\ &= \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \cup \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w))\end{aligned}$$

and

$$\begin{aligned}\mathsf{VF}([vw]) \setminus \{\mathbf{y}\} &= (\mathsf{VF}(v) \cup \mathsf{VF}(w)) \setminus \{\mathbf{y}\} \\ &= (\mathsf{VF}(v) \setminus \{\mathbf{y}\}) \cup (\mathsf{VF}(w) \setminus \{\mathbf{y}\}) \\ &\subseteq \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)) \cup \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w)) \\ &\subseteq \mathsf{VF}(v) \cup \mathsf{VF}(w) \\ &= \mathsf{VF}([vw]).\end{aligned}$$

Hence $\mathsf{VF}(\mathsf{u}) \setminus \{\mathbf{y} \subseteq \mathsf{VF}(\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathsf{u})) \subseteq \mathsf{VF}(\mathsf{u})$ for all $\mathsf{u} \in \mathsf{U}$, as desired.

3. • Let $s \in \mathcal{B}$. Then

$$\mathrm{SF}(\mathbf{x})(\mathbf{x})(s) = \begin{cases} \mathbf{x} & \text{if } s = \mathbf{x} \\ s & \text{otherwise} \end{cases}$$
$$= s.$$

Furthermore

$$\mathrm{SB}(\mathbf{x})(\mathbf{x})(s) = s.$$

• Now let $\mathbf{y} \in \mathcal{X}$, let $v \in \mathrm{U}$ be such that $\mathrm{SF}(\mathbf{x})(\mathbf{x})(v) = v$ and $\mathrm{SB}(\mathbf{x})(\mathbf{x})(v) = v$ and consider $[\lambda \mathbf{y} v]$. We have that

$$\mathrm{SF}(\mathbf{x})(\mathbf{x})([\lambda \mathbf{y} v]) = \begin{cases} [\lambda \mathbf{y} v] & \text{if } \mathbf{y} = \mathbf{x} \\ [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{x})(v)] & \text{otherwise} \end{cases}$$
$$= \begin{cases} [\lambda \mathbf{y} v] & \text{if } \mathbf{y} = \mathbf{x} \\ [\lambda \mathbf{y} v] & \text{otherwise} \end{cases}$$
$$= [\lambda \mathbf{y} v].$$

as well as

$$\mathrm{SB}(\mathbf{x})(\mathbf{x})([\lambda \mathbf{y} v]) = \begin{cases} [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{x})(\mathrm{SB}(\mathbf{x})(\mathbf{x})(v))] & \text{if } \mathbf{y} = \mathbf{x} \\ [\lambda \mathbf{y}\, \mathrm{SB}(\mathbf{x})(\mathbf{x})(v)] & \text{otherwise} \end{cases}$$
$$= \begin{cases} [\lambda \mathbf{y} v] & \text{if } \mathbf{y} = \mathbf{x} \\ [\lambda \mathbf{y} v] & \text{otherwise} \end{cases}$$
$$= [\lambda \mathbf{y} v].$$

• Finally let $v, w \in \mathrm{U}$ be such that $\mathrm{SF}(\mathbf{x})(\mathbf{x})(v) = v = \mathrm{SB}(\mathbf{x})(\mathbf{x})(v)$ and $\mathrm{SF}(\mathbf{x})(\mathbf{x})(w) = w = \mathrm{SB}(\mathbf{x})(\mathbf{x})(w)$. Then

$$\mathrm{SF}(\mathbf{x})(\mathbf{x})([vw]) = [\mathrm{SF}(\mathbf{x})(\mathbf{x})(v)\, \mathrm{SF}(\mathbf{x})(\mathbf{x})(w)]$$
$$= [vw]$$

and

$$\mathrm{SB}(\mathbf{x})(\mathbf{x})([vw]) = [\mathrm{SB}(\mathbf{x})(\mathbf{x})(v)\, \mathrm{SB}(\mathbf{x})(\mathbf{x})(w)]$$
$$= [vw].$$

Hence $\mathrm{SF}(\mathbf{x})(\mathbf{x})(u) = u$ and $\mathrm{SB}(\mathbf{x})(\mathbf{x})(u) = u$ for all $u \in \mathrm{U}$, as claimed.

4. Let $\mathbf{y}, \mathbf{z} \in \mathcal{X}$.

- Let $s \in \mathcal{B} \setminus \{z\}$. First observe that

$$\mathrm{SB}(z)(y)(\mathrm{SB}(x)(z)(s)) = \mathrm{SB}(z)(y)(s) = s = \mathrm{SB}(x)(y)(s).$$

Furthermore

$$\begin{aligned}\mathrm{SF}(z)(y)(\mathrm{SF}(x)(z)(s)) &= \begin{cases}\mathrm{SF}(z)(y)(z) & \text{if } s = x\\ \mathrm{SF}(z)(y)(s) & \text{otherwise}\end{cases}\\ &= \begin{cases} y & \text{if } s = x\\ s & \text{otherwise}\end{cases}\\ &= \begin{cases}\mathrm{SF}(x)(y)(s) & \text{if } s = x\\ \mathrm{SF}(x)(y)(s) & \text{otherwise}\end{cases}\\ &= \mathrm{SF}(x)(y)(s).\end{aligned}$$

- Assume that $v \in \mathrm{U}$ is such that if z does not occur in v, then

$$\mathrm{SB}(z)(y)(\mathrm{SB}(x)(z)(v)) = \mathrm{SB}(x)(y)(v)$$

and

$$\mathrm{SF}(z)(y)(\mathrm{SF}(x)(z)(v)) = \mathrm{SF}(x)(y)(v).$$

Let $a \in \mathcal{X}$ and assume z does not occur in $[\lambda a v]$, i.e. that $a \neq z$ and z does not occur in v. We consider cases according as $a = x$ or not. If $a = x$ then

$$\begin{aligned}\mathrm{SB}(z)(y)(\mathrm{SB}(x)(z)([\lambda a v])) &= \mathrm{SB}(z)(y)([\lambda z\, \mathrm{S}(x)(z)(v)])\\ &= [\lambda y\, \mathrm{S}(z)(y)(\mathrm{S}(x)(z)(v))]\\ &= [\lambda y\, \mathrm{S}(x)(y)(v)]\\ &= \mathrm{SB}(x)(y)([\lambda a v])\end{aligned}$$

by 1.17, and

$$\begin{aligned}\mathrm{SF}(z)(y)(\mathrm{SF}(x)(z)([\lambda a v])) &= \mathrm{SF}(z)(y)([\lambda a v])\\ &= [\lambda a v]\\ &= \mathrm{SF}(x)(y)([\lambda a v]).\end{aligned}$$

If on the other hand $a \neq x$

$$\begin{aligned}\mathrm{SB}(z)(y)(\mathrm{SB}(x)(z)([\lambda a v])) &= \mathrm{SB}(z)(y)([\lambda a\, \mathrm{SB}(x)(z)(v)])\\ &= [\lambda a\, \mathrm{SB}(z)(y)(\mathrm{SB}(x)(z)(v))]\\ &= [\lambda a\, \mathrm{SB}(x)(y)(v)]\\ &= \mathrm{SB}(x)(y)([\lambda a v])\end{aligned}$$

and

$$\begin{aligned}\mathrm{SF}(z)(\mathbf{y})(\mathrm{SF}(\mathbf{x})(z)([\lambda \mathbf{a} v])) &= \mathrm{SF}(z)(\mathbf{y})([\lambda \mathbf{a}\, \mathrm{SF}(\mathbf{x})(z)(v)]) \\ &= [\lambda \mathbf{a}\, \mathrm{SF}(z)(\mathbf{y})(\mathrm{SF}(\mathbf{x})(z)(v))] \\ &= [\lambda \mathbf{a}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(v)] \\ &= \mathrm{SF}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{a} v])\end{aligned}$$

by induction hypothesis.

- Let $v, w \in \mathsf{U}$ be such that

$$\mathrm{SB}(z)(\mathbf{y})(\mathrm{SB}(\mathbf{x})(z)(q)) = \mathrm{SB}(\mathbf{x})(\mathbf{y})(q)$$

and

$$\mathrm{SF}(z)(\mathbf{y})(\mathrm{SF}(\mathbf{x})(z)(q)) = \mathrm{SF}(\mathbf{x})(\mathbf{y})(q)$$

for all $q \in \{v, w\}$ with $z \not\preceq q$. Thus, if z does not occur in $[vw]$ (i.e. $z \not\preceq v, w$),

$$\begin{aligned}\mathrm{SB}(z)(\mathbf{y})(\mathrm{SB}(\mathbf{x})(z)([vw])) &= \mathrm{SB}(z)(\mathbf{y})([\mathrm{SB}(\mathbf{x})(z)(v)\, \mathrm{SB}(\mathbf{x})(z)(w)]) \\ &= [\mathrm{SB}(z)(\mathbf{y})(\mathrm{SB}(\mathbf{x})(z)(v))\, \mathrm{SB}(z)(\mathbf{y})(\mathrm{SB}(\mathbf{x})(z)(w))] \\ &= [\mathrm{SB}(\mathbf{x})(\mathbf{y})(v)\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(w)] \\ &= \mathrm{SB}(\mathbf{x})(\mathbf{y})([vw])\end{aligned}$$

and

$$\begin{aligned}\mathrm{SF}(z)(\mathbf{y})(\mathrm{SF}(\mathbf{x})(z)([vw])) &= \mathrm{SF}(z)(\mathbf{y})([\mathrm{SF}(\mathbf{x})(z)(v)\, \mathrm{SF}(\mathbf{x})(z)(w)]) \\ &= [\mathrm{SF}(z)(\mathbf{y})(\mathrm{SF}(\mathbf{x})(z)(v))\, \mathrm{SF}(z)(\mathbf{y})(\mathrm{SF}(\mathbf{x})(z)(w))] \\ &= [\mathrm{SF}(\mathbf{x})(\mathbf{y})(v)\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(w)] \\ &= \mathrm{SF}(\mathbf{x})(\mathbf{y})([vw])\end{aligned}$$

by assumption.

Hence

$$\mathrm{SB}(z)(\mathbf{y})(\mathrm{SB}(\mathbf{x})(z)(u)) = \mathrm{SB}(\mathbf{x})(\mathbf{y})(u)$$

and

$$\mathrm{SF}(z)(\mathbf{y})(\mathrm{SF}(\mathbf{x})(z)(u)) = \mathrm{SF}(\mathbf{x})(\mathbf{y})(u)$$

for all u such that z does not occur in u.

5. Let $\mathbf{y} \in \mathcal{X}$ and $v \in \mathsf{U}$.

- Let $s \in \mathcal{B}$. $\mathbf{x}$ is never bound in s,

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})(s) = s$$

and

$$\begin{aligned}\mathrm{SF}(\mathbf{x})(v)(s) &= \begin{cases} v & \text{if } s = \mathbf{x} \\ s & \text{otherwise} \end{cases} \\ &= \mathrm{S}(\mathbf{x})(v)(s).\end{aligned}$$

- Let $z \in \mathcal{X}$ and $w \in \mathrm{U}$ be such that if $\mathbf{x}$ is not bound in w, then $\mathrm{SB}(\mathbf{x})(\mathbf{y})(w) = w$ and $\mathrm{SF}(\mathbf{x})(v)(w) = \mathrm{S}(\mathbf{x})(v)(w)$. Now, if $\mathbf{x}$ is not bound in $[\lambda z w]$, then $\mathbf{x} \neq z$ and $\mathbf{x}$ is not bound in w. Hence

$$\begin{aligned}\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda z w]) &= [\lambda z\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(w)] \\ &= [\lambda z w]\end{aligned}$$

as well as

$$\begin{aligned}\mathrm{SF}(\mathbf{x})(v)([\lambda z w]) &= [\lambda z\, \mathrm{SF}(\mathbf{x})(v)(w)] \\ &= [\lambda z\, \mathrm{S}(\mathbf{x})(v)(w)] \\ &= \mathrm{S}(\mathbf{x})(v)([\lambda z w]).\end{aligned}$$

- Finally, let $w, q \in \mathrm{U}$ be such that, if $\mathbf{x}$ is not bound in a, then $\mathrm{SB}(\mathbf{x})(\mathbf{y})(a) = a$ and $\mathrm{SF}(\mathbf{x})(v)(a) = \mathrm{S}(\mathbf{x})(v)(a)$, for all $a \in \{w, q\}$. Moreover, assume that $\mathbf{x}$ is not bound in $[wq]$, whence $\mathbf{x}$ is bound in neither w nor q. Thus

$$\begin{aligned}\mathrm{SB}(\mathbf{x})(\mathbf{y})([wq]) &= [\mathrm{SB}(\mathbf{x})(\mathbf{y})(w)\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(q)] \\ &= [wq]\end{aligned}$$

and

$$\begin{aligned}\mathrm{SF}(\mathbf{x})(v)([wq]) &= [\mathrm{SF}(\mathbf{x})(v)(w)\, \mathrm{SF}(\mathbf{y})(v)(q)] \\ &= [\mathrm{S}(\mathbf{x})(v)(w)\, \mathrm{S}(\mathbf{x})(v)(q)] \\ &= \mathrm{S}(\mathbf{x})(v)([wq]).\end{aligned}$$

Hence, if $\mathbf{x}$ is not bound in u, $\mathrm{SB}(\mathbf{x})(\mathbf{y})(u) = u$ and $\mathrm{SF}(\mathbf{x})(v)(u) = \mathrm{S}(\mathbf{x})(v)(u)$.

□

2.23 Lemma. *Let $\mathbf{x} \in \mathcal{X}$ and $u \in \mathrm{U}$. If $\mathbf{x}$ is not free in u, then*

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})(u) = \mathrm{S}(\mathbf{x})(\mathbf{y})(u)$$

for all $\mathbf{y} \in \mathcal{X}$.

Proof. Assume that x is not free in u, and let $y \in \mathcal{X}$. If $y = x$ then

$$S(x)(y)(u) = u = SB(x)(y)(u)$$

by 2.22; if $y \neq x$, x is not free in $SB(x)(y)(u)$ by the same, and hence

$$\begin{aligned} S(x)(y)(u) &= SF(x)(y)(SB(x)(y)(u)) \\ &= SB(x)(y)(u) \end{aligned}$$

by 2.19 and 2.18. □

2.24 Lemma. *For all $x, y, z \in \mathcal{X}$ and $u, v \in U$ such that x does not occur in v and $y \neq z$ we have that*

1. *If $x \neq z$ then*

$$SF(x)(y)(SF(z)(v)(u)) = SF(z)(v)(SF(x)(y)(u)).$$

2.

$$SB(x)(y)(SF(z)(v)(u)) = SF(z)(v)(SB(x)(y)(u))$$

Proof. The proofs are by induction on u.

1. • Let $s \in \mathcal{B}$.
   If $z = s$,

$$SF(z)(v)(s) = v$$

   and x is not free in $SF(z)(v)(s)$. Thus by 2.18

$$SF(x)(y)(SF(z)(v)(s)) = SF(z)(v)(s)$$

   and, since $x \neq s$,

$$SF(z)(v)(SF(x)(y)(s)) = SF(z)(v)(s)$$

   by definition.
   If $z \neq s$,

$$\begin{aligned} SF(z)(v)(SF(x)(y)(s)) &= \begin{cases} SF(z)(v)(y) & \text{if } x = s \\ SF(z)(v)(s) & \text{otherwise} \end{cases} \\ &= \begin{cases} y & \text{if } x = s \\ s & \text{otherwise} \end{cases} \\ &= SF(x)(y)(s) \\ &= SF(x)(y)(SF(z)(v)(s)). \end{aligned}$$

- Let $w \in \mathsf{U}$ be such that $\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(\mathrm{SF}(\boldsymbol{z})(\nu)(w)) = \mathrm{SF}(\boldsymbol{z})(\nu)(\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(w))$ and let $\boldsymbol{a} \in \mathcal{X}$.

  If $\boldsymbol{a} = \boldsymbol{x}$ we have that

$$\begin{aligned}\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(\mathrm{SF}(\boldsymbol{z})(\nu)([\lambda\boldsymbol{a}w])) &= \mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{a}\,\mathrm{SF}(\boldsymbol{z})(\nu)(w)]) \\ &= [\lambda\boldsymbol{a}\,\mathrm{SF}(\boldsymbol{z})(\nu)(w)] \\ &= \mathrm{SF}(\boldsymbol{z})(\nu)([\lambda\boldsymbol{a}w]) \\ &= \mathrm{SF}(\boldsymbol{z})(\nu)(\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{a}w]))\end{aligned}$$

  since $\boldsymbol{x} \neq \boldsymbol{z} \neq \boldsymbol{y}$.

  If $\boldsymbol{a} = \boldsymbol{z}$ we have that

$$\begin{aligned}\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(\mathrm{SF}(\boldsymbol{z})(\nu)([\lambda\boldsymbol{a}w])) &= \mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{a}w]) \\ &= [\lambda\boldsymbol{a}\,\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(w)] \\ &= \mathrm{SF}(\boldsymbol{z})(\nu)([\lambda\boldsymbol{a}\,\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(w)]) \\ &= \mathrm{SF}(\boldsymbol{z})(\nu)(\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{a}w])).\end{aligned}$$

  Finally, if $\boldsymbol{x} \neq \boldsymbol{a} \neq \boldsymbol{z}$ we have that

$$\begin{aligned}\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(\mathrm{SF}(\boldsymbol{z})(\nu)([\lambda\boldsymbol{a}w])) &= \mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{a}\,\mathrm{SF}(\boldsymbol{z})(\nu)(w)]) \\ &= [\lambda\boldsymbol{a}\,\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(\mathrm{SF}(\boldsymbol{z})(\nu)(w))] \\ &= [\lambda\boldsymbol{a}\,\mathrm{SF}(\boldsymbol{z})(\nu)(\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(w))] \\ &= \mathrm{SF}(\boldsymbol{z})(\nu)([\lambda\boldsymbol{a}\,\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(w)]) \\ &= \mathrm{SF}(\boldsymbol{z})(\nu)(\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{a}w])).\end{aligned}$$

  In either case $\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(\mathrm{SF}(\boldsymbol{z})(\nu)([\lambda\boldsymbol{a}w])) = \mathrm{SF}(\boldsymbol{z})(\nu)(\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{a}w]))$.

- Let $w, q \in \mathsf{U}$ be such that $\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(\mathrm{SF}(\boldsymbol{z})(\nu)(a)) = \mathrm{SF}(\boldsymbol{z})(\nu)(\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(a))$ for both $a \in \{w, q\}$. Then

$$\begin{aligned}\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(\mathrm{SF}(\boldsymbol{z})(\nu)([wq])) &= \mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})([\mathrm{SF}(\boldsymbol{z})(\nu)(w)\,\mathrm{SF}(\boldsymbol{z})(\nu)(q)]) \\ &= [\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(\mathrm{SF}(\boldsymbol{z})(\nu)(w))\,\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(\mathrm{SF}(\boldsymbol{z})(\nu)(q))] \\ &= [\mathrm{SF}(\boldsymbol{z})(\nu)(\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(w))\,\mathrm{SF}(\boldsymbol{z})(\nu)(\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(q))] \\ &= \mathrm{SF}(\boldsymbol{z})(\nu)([\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(w)\,\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(q)]) \\ &= \mathrm{SF}(\boldsymbol{z})(\nu)(\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})([wq])).\end{aligned}$$

Hence $\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(\mathrm{SF}(\boldsymbol{z})(\nu)(u)) = \mathrm{SF}(\boldsymbol{z})(\nu)(\mathrm{SF}(\boldsymbol{x})(\boldsymbol{y})(u))$ for all $u \in \mathsf{U}$ as desired.

2. - Let $s \in \mathcal{B}$. As

$$\mathrm{SF}(\boldsymbol{z})(\nu)(s) = \begin{cases} \nu & \text{if } s = \boldsymbol{z} \\ s & \text{otherwise} \end{cases}$$

$\mathbf{x}$ is not bound in $\mathrm{SF}(\mathbf{z})(v)(u)$. Thus

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathrm{SF}(\mathbf{z})(v)(s)) = \mathrm{SF}(\mathbf{z})(v)(s)$$

by Lemma 2.22. Furthermore

$$\mathrm{SF}(\mathbf{z})(v)(\mathrm{SB}(\mathbf{x})(\mathbf{y})(s)) = \mathrm{SF}(\mathbf{z})(v)(s)$$

by definition of SB.

- Let $w \in \mathcal{U}$ be such that $\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathrm{SF}(\mathbf{z})(v)(w)) = \mathrm{SF}(\mathbf{z})(v)(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w))$ and let $\mathbf{a} \in \mathcal{X}$.

  If $\mathbf{x} = \mathbf{a} = \mathbf{z}$, then

$$\begin{aligned}\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathrm{SF}(\mathbf{z})(v)([\lambda \mathbf{a} w])) &= \mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{a} w]) \\ &= [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w))]\end{aligned}$$

  and

$$\begin{aligned}\mathrm{SF}(\mathbf{z})(v)(\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{a} w])) &= \mathrm{SF}(\mathbf{z})(v)([\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w))]) \\ &= [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{z})(v)(\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w)))] \\ &= [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w))]\end{aligned}$$

  since $\mathbf{z}$ is not free in $\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w))$, by 2.18.

  If $\mathbf{x} = \mathbf{a} \neq \mathbf{z}$, then

$$\begin{aligned}\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathrm{SF}(\mathbf{z})(v)([\lambda \mathbf{a} w])) &= \mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{a}\, \mathrm{SF}(\mathbf{z})(v)(w)]) \\ &= [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathrm{SF}(\mathbf{z})(v)(w)))] \\ &= [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SF}(\mathbf{z})(v)(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w)))]\end{aligned}$$

  by induction hypothesis, while

$$\begin{aligned}\mathrm{SF}(\mathbf{z})(v)(\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{a} w])) &= \mathrm{SF}(\mathbf{z})(v)([\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w))]) \\ &= [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{z})(v)(\mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w)))] \\ &= [\lambda \mathbf{y}\, \mathrm{SF}(\mathbf{x})(\mathbf{y})(\mathrm{SF}(\mathbf{z})(v)(\mathrm{SB}(\mathbf{x})(\mathbf{y})(w)))]\end{aligned}$$

  by 1.

  If $\mathbf{x} \neq \mathbf{a} = \mathbf{z}$, then

$$\begin{aligned}\mathrm{SB}(\mathbf{x})(\mathbf{y})(\mathrm{SF}(\mathbf{z})(v)([\lambda \mathbf{a} w])) &= \mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{a} w]) \\ &= [\lambda \mathbf{a}\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(w)].\end{aligned}$$

  Furthermore

$$\begin{aligned}\mathrm{SF}(\mathbf{z})(v)(\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{a} w])) &= \mathrm{SF}(\mathbf{z})(v)([\lambda \mathbf{a}\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(w)]) \\ &= [\lambda \mathbf{a}\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(w)].\end{aligned}$$

If $x \neq a \neq z$, then

$$\begin{aligned}\mathrm{SB}(x)(y)(\mathrm{SF}(z)(v)([\lambda a w])) &= \mathrm{SB}(x)(y)([\lambda a\, \mathrm{SF}(z)(v)(w)])\\ &= [\lambda a\, \mathrm{SB}(x)(y)(\mathrm{SF}(z)(v)(w))]\\ &= [\lambda a\, \mathrm{SF}(z)(v)(\mathrm{SB}(x)(y)(w))]\\ &= \mathrm{SF}(z)(v)([\lambda a\, \mathrm{SB}(x)(y)(w)])\\ &= \mathrm{SF}(z)(v)(\mathrm{SB}(x)(y)([\lambda a w]))\end{aligned}$$

by hypothesis.

- Now let $w, q \in U$ be such that

$$\mathrm{SB}(x)(y)(\mathrm{SF}(z)(v)(a)) = \mathrm{SF}(z)(v)(\mathrm{SB}(x)(y)(a))$$

for all $a \in \{w, q\}$.

$$\begin{aligned}\mathrm{SB}(x)(y)(\mathrm{SF}(z)(v)([wq])) &= \mathrm{SB}(x)(y)([\mathrm{SF}(z)(v)(w)\, \mathrm{SF}(z)(v)(q)])\\ &= [\mathrm{SB}(x)(y)(\mathrm{SF}(z)(v)(w))\, \mathrm{SB}(x)(y)(\mathrm{SF}(z)(v)(q))]\\ &= [\mathrm{SF}(z)(v)(\mathrm{SB}(x)(y)(w))\, \mathrm{SF}(z)(v)(\mathrm{SB}(x)(y)(q))]\\ &= \mathrm{SF}(z)(v)([\mathrm{SB}(x)(y)(w)\, \mathrm{SB}(x)(y)(q)])\\ &= \mathrm{SF}(z)(v)(\mathrm{SB}(x)(y)([wq]))\end{aligned}$$

by hypothesis.

□

To conclude this section we prove the following theorem, which, though quite intuitive, will prove very useful in the next section (3).

2.25 Theorem. *For every* $u \in U$ *and* $x \in \mathcal{X}$ *there is an* $n \in \mathbb{N}$ *and sequences* $\{r_k\}_{k=0}^{n}$, $\{u_k\}_{k=1}^{n}$ *such that* $r_k \in \mathcal{S}^*$ *for all* $k \in [0, n]$, $u_k \in U$ *for all* $k \in [1, n]$ *and*

$$u = r_o \left(\prod_{k=1}^{n} [\lambda x u_k] r_k \right)$$

with the additional properties

$$\mathrm{SB}(x)(y)(u) = r_o \left(\prod_{k=1}^{n} \mathrm{S}(x)(y)([\lambda x u_k]) r_k \right)$$

and

$$\mathrm{SF}(x)(v)(u) = \mathrm{S}(x)(v)(r_o) \left(\prod_{k=1}^{n} [\lambda x u_k]\, \mathrm{S}(x)(v)(r_k) \right)$$

for every $v \in U$. *Furthermore, if* $u \in W$ *then* $u_k \in W$ *for every* $k \in [1, n]$.

Proof. Let $\mathbf{y} \in \mathcal{X}$ and $v \in U$. The proof is by induction on u:

- For any $s \in \mathcal{B}$, $s = s\left(\prod_{k=1}^{0}[\lambda \mathbf{x} s]\diamond\right)$ is such a representation, for which the statement holds trivially.
- Now consider $[\lambda z w]$, where $z \in \mathcal{X}$, and $w \in U$ is such that there is a representation

$$w = a_0 \left(\prod_{k=1}^{m} [\lambda \mathbf{x} w_k] a_k \right)$$

where $m \in \mathbb{N}$, $w_k \in U$ for all $k \in [1, m]$, and $w_k \in W$ for all $k \in [1, m]$ if w happens to be well-formed, which also satisfies

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})(w) = a_0 \left(\prod_{k=1}^{m} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} w_k]) a_k \right)$$

and

$$\mathrm{SF}(\mathbf{x})(v)(w) = \mathrm{S}(\mathbf{x})(v)(a_0) \left(\prod_{k=1}^{m} [\lambda \mathbf{x} w_k]\, \mathrm{S}(\mathbf{x})(v)(a_k) \right).$$

Note that $[\lambda z w]$ is well-formed if and only if w is well-formed.

If $z = \mathbf{x}$, then

$$[\lambda z w] = \diamond \left(\prod_{k=1}^{1} [\lambda z w] \diamond \right)$$

where $w \in U$ is well-formed if $[\lambda z w]$ is, whereas

$$\begin{aligned} \mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda z w]) &= [\lambda \mathbf{y}\, \mathrm{S}(\mathbf{x})(\mathbf{y})(w)] \\ &= \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda z w]) \\ &= \diamond \left(\prod_{k=1}^{1} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda z w]) \diamond \right) \end{aligned}$$

and

$$\begin{aligned} \mathrm{SF}(\mathbf{x})(v)([\lambda z w]) &= [\lambda z w] \\ &= \mathrm{S}(\mathbf{x})(v)(\diamond) \left(\prod_{k=1}^{1} [\lambda z w]\, \mathrm{S}(\mathbf{x})(v)(\diamond) \right). \end{aligned}$$

If on the other hand $z \neq \mathbf{x}$ and $m = 0$, then $w = a_0$, so that

$$[\lambda z w] = [\lambda z a_0] \left(\prod_{k=1}^{0} [\lambda \mathbf{x} w_k] a_k \right)$$

where $w_k \in U$ is well-formed for all $k \in [1, 0]$,

$$\begin{aligned}
\mathrm{SB}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{z} w]) &= [\lambda \mathbf{z}\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(w)] \\
&= [\lambda \mathbf{z} a_0 \left(\prod_{k=1}^{0} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} w_k]) a_k \right)] \\
&= [\lambda \mathbf{z} a_0] \\
&= [\lambda \mathbf{z} a_0] \left(\prod_{k=1}^{0} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} w_k]) a_k \right)
\end{aligned}$$

(note the moved bracket) and

$$\begin{aligned}
\mathrm{SF}(\mathbf{x})(v)([\lambda \mathbf{z} w]) &= [\lambda \mathbf{z}\, \mathrm{SF}(\mathbf{x})(v)(w)] \\
&= [\lambda \mathbf{z}\, \mathrm{S}(\mathbf{x})(v)(a_0) \left(\prod_{k=1}^{0} [\lambda \mathbf{x} w_k]\, \mathrm{S}(\mathbf{x})(v)(a_k) \right)] \\
&= [\lambda \mathbf{z}\, \mathrm{S}(\mathbf{x})(v)(a_0)] \\
&= [\lambda \mathbf{z}\, \mathrm{S}(\mathbf{x})(v)(a_0)] \left(\prod_{k=1}^{0} [\lambda \mathbf{x} w_k]\, \mathrm{S}(\mathbf{x})(v)(a_k) \right).
\end{aligned}$$

If finally $\mathbf{z} \neq \mathbf{x}$ and $m > 0$, set $r_0 = [\lambda \mathbf{z} a_0$, $u_k = w_k$ and $r_k = a_k$ for all $k \in [1, m[$, while $u_m = w_m$ and $r_m = a_m]$. Hence

$$\begin{aligned}
[\lambda \mathbf{z} w] &= [\lambda \mathbf{z} a_0 \left(\prod_{k=1}^{m} [\lambda \mathbf{x} w_k] \right)] \\
&= r_0 \left(\prod_{k=1}^{m-1} [\lambda \mathbf{x} w_k] a_k \right) [\lambda \mathbf{x} w_m] a_m] \\
&= r_0 \left(\prod_{k=1}^{m-1} [\lambda \mathbf{x} u_k] r_k \right) [\lambda \mathbf{x} u_m] r_m \\
&= r_0 \left(\prod_{k=1}^{m} [\lambda \mathbf{x} u_k] r_k \right).
\end{aligned}$$

Furthermore, $u_k = w_k \in U$ for all $k \in [1, m]$, and if $[\lambda \mathbf{z} w]$ is well-formed, then w

is as well, so that $u_k = w_k \in W$ for all $k \in [1, m]$. Finally

$$\begin{aligned}
\mathrm{SB}(\boldsymbol{x})(\boldsymbol{y})([\lambda z w]) &= [\lambda \boldsymbol{z}\, \mathrm{SB}(\boldsymbol{x})(\boldsymbol{y})(w)] \\
&= [\lambda z a_0 \left(\prod_{k=1}^{m} \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda \boldsymbol{x} w_k]) a_k \right)] \\
&= r_0 \left(\prod_{k=1}^{m-1} \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda \boldsymbol{x} w_k]) a_k \right) \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda \boldsymbol{x} w_m]) a_m] \\
&= r_0 \left(\prod_{k=1}^{m-1} \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda \boldsymbol{x} u_k]) r_k \right) \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda \boldsymbol{x} u_m]) r_m \\
&= r_0 \left(\prod_{k=1}^{m} \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda \boldsymbol{x} u_k]) r_k \right)
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{SF}(\boldsymbol{x})(v)([\lambda z w]) &= [\lambda \boldsymbol{z}\, \mathrm{SF}(\boldsymbol{x})(v)(w)] \\
&= [\lambda \boldsymbol{z}\, \mathrm{S}(\boldsymbol{x})(v)(a_0) \left(\prod_{k=1}^{m} [\lambda \boldsymbol{x} w_k]\, \mathrm{S}(\boldsymbol{x})(v)(a_k) \right)] \\
&= \mathrm{S}(\boldsymbol{x})(v)(r_0) \left(\prod_{k=1}^{m-1} [\lambda \boldsymbol{x} w_k]\, \mathrm{S}(\boldsymbol{x})(v)(a_k) \right) [\lambda \boldsymbol{x} w_m]\, \mathrm{S}(\boldsymbol{x})(v)(a_m]) \\
&= \mathrm{S}(\boldsymbol{x})(v)(r_0) \left(\prod_{k=1}^{m-1} [\lambda \boldsymbol{x} u_k]\, \mathrm{S}(\boldsymbol{x})(v)(r_k) \right) [\lambda \boldsymbol{x} u_m]\, \mathrm{S}(\boldsymbol{x})(v)(r_m) \\
&= \mathrm{S}(\boldsymbol{x})(v)(r_0) \left(\prod_{k=1}^{m} [\lambda \boldsymbol{x} u_k]\, \mathrm{S}(\boldsymbol{x})(v)(r_k) \right).
\end{aligned}$$

- Now consider $[wq]$, where $w, q \in U$ are such that there are representations

$$\begin{aligned}
w &= a_0 \left(\prod_{k=1}^{m} [\lambda \boldsymbol{x} w_k] a_k \right) \\
q &= b_0 \left(\prod_{j=1}^{l} [\lambda \boldsymbol{x} q_j] b_j \right)
\end{aligned}$$

where $m, l \in \mathbb{N}$, $w_k, q_j \in U$ for all $k \in [1, m]$ and $j \in [1, l]$, satisfying $w_k \in W$ for all $k \in [1, m]$ and $q_j \in W$ for all $j \in [1, l]$ if w or q happens to be well-formed, respectively, which have the additional properties that

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})(w) = a_0 \left(\prod_{k=1}^{m} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} w_k]) a_k \right)$$

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})(q) = b_0 \left(\prod_{j=1}^{l} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} q_k]) b_k \right)$$

and

$$\mathrm{SF}(\mathbf{x})(\nu)(w) = \mathrm{S}(\mathbf{x})(\nu)(a_0) \left(\prod_{k=1}^{m} [\lambda \mathbf{x} w_k]\, \mathrm{S}(\mathbf{x})(\nu)(a_k) \right)$$

$$\mathrm{SF}(\mathbf{x})(\nu)(q) = \mathrm{S}(\mathbf{x})(\nu)(b_0) \left(\prod_{j=1}^{l} [\lambda \mathbf{x} q_k]\, \mathrm{S}(\mathbf{x})(\nu)(b_k) \right).$$

Note that $[wq] \in W$ if and only if there are $\alpha, \beta \in \mathcal{T}$ such that $w \in W_{\alpha\beta}$ and $q \in W_{\beta}$.

If $m = 0 = l$ then

$$\begin{aligned} [wq] &= [a_0 b_0] \\ &= [a_0 b_o] \left(\prod_{k=1}^{0} [\lambda \mathbf{x} \varphi_k] c_k \right) \end{aligned}$$

where $\varphi_k \in W$ for all $k \in [1, 0]$. Thus

$$\begin{aligned} \mathrm{SB}(\mathbf{x})(\mathbf{y})([wq]) &= [\mathrm{SB}(\mathbf{x})(\mathbf{y})(w)\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(q)] \\ &= [a_0 \left(\prod_{k=1}^{0} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} w_k]) a_k \right) b_0 \left(\prod_{j=1}^{0} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} q_k]) b_k \right)] \\ &= [a_0 b_0] \\ &= [a_0 b_o] \left(\prod_{k=1}^{0} \mathrm{S}([\lambda \mathbf{x} \varphi_k]) c_k \right) \end{aligned}$$

and

$$
\begin{aligned}
&\mathrm{SF}(\boldsymbol{x})(\nu)([wq])\\
=&[\mathrm{SF}(\boldsymbol{x})(\nu)(w)\,\mathrm{SF}(\boldsymbol{x})(\nu)(q)]\\
=&[\mathrm{S}(\boldsymbol{x})(\nu)(a_0)\left(\prod_{k=1}^{0}[\lambda\boldsymbol{x}w_k]\,\mathrm{S}(\boldsymbol{x})(\nu)(a_k)\right)\mathrm{S}(\boldsymbol{x})(\nu)(b_0)\\
&\left(\prod_{j=1}^{0}[\lambda\boldsymbol{x}q_k]\,\mathrm{S}(\boldsymbol{x})(\nu)(b_k)\right)]\\
=&[\mathrm{S}(\boldsymbol{x})(\nu)(a_0)\,\mathrm{S}(\boldsymbol{x})(\nu)(b_0)]\\
=&\,\mathrm{S}(\boldsymbol{x})(\nu)([a_0b_0])\\
=&\,\mathrm{S}(\boldsymbol{x})(\nu)([a_0b_0])\left(\prod_{k=1}^{0}[\lambda\boldsymbol{x}\varphi_k]\,\mathrm{S}(\boldsymbol{x})(\nu)(c_k)\right).
\end{aligned}
$$

If on the other hand $m > 0$ and $l = 0$, let $c_0 = [a_0$, $p_k = w_k$ for all $k \in [1, m]$, $c_k = a_k$ for all $k \in [1, m[$, $c_m = a_m b_0]$, Then

$$
\begin{aligned}
[wq] &= [a_0\left(\prod_{k=1}^{m}[\lambda\boldsymbol{x}w_k]a_k\right)b_0\left(\prod_{j=1}^{0}[\lambda\boldsymbol{x}q_j]b_j\right)]\\
&= c_0\left(\prod_{k=1}^{m-1}[\lambda\boldsymbol{x}w_k]a_k\right)[\lambda\boldsymbol{x}w_m]a_m b_0]\\
&= c_0\left(\prod_{k=1}^{m-1}[\lambda\boldsymbol{x}p_k]c_k\right)[\lambda\boldsymbol{x}p_m]c_m\\
&= c_0\left(\prod_{k=1}^{m}[\lambda\boldsymbol{x}p_k]c_k\right).
\end{aligned}
$$

Furthermore

$$p_k \in \{w_k \,|\, 0 < k \leq m\} \subseteq U.$$

If $[wq]$ is well-formed, w and q are well-formed, whereby $p_k \in W$ for all $k \in [1, m]$, like before. Finally

$$
\begin{aligned}
&\mathrm{SB}(\boldsymbol{x})(\boldsymbol{y})([wq])\\
=&[\mathrm{SB}(\boldsymbol{x})(\boldsymbol{y})(w)\,\mathrm{SB}(\boldsymbol{x})(\boldsymbol{y})(q)]\\
=&[a_0\left(\prod_{k=1}^{m}\mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{x}w_k])a_k\right)b_0\left(\prod_{j=1}^{0}\mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{x}q_j])b_j\right)]\\
=&c_0\left(\prod_{k=1}^{m-1}\mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{x}w_k])a_k\right)\mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{x}w_m])a_m b_0]\\
=&c_0\left(\prod_{k=1}^{m-1}\mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{x}p_k])c_k\right)\mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{x}p_m])c_m\\
=&c_0\left(\prod_{k=1}^{m}\mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda\boldsymbol{x}p_k])c_k\right)
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathrm{SF}(\boldsymbol{x})(\nu)([wq])\\
=&[\mathrm{SF}(\boldsymbol{x})(\nu)(w)\,\mathrm{SF}(\boldsymbol{x})(\nu)(q)]\\
=&[\mathrm{SF}(\boldsymbol{x})(\nu)(w)\,\mathrm{S}(\boldsymbol{x})(\nu)(b_0)\left(\prod_{j=1}^{0}[\lambda\boldsymbol{x}q_j]\,\mathrm{S}(\boldsymbol{x})(\nu)(b_j)\right)]\\
=&[\mathrm{SF}(\boldsymbol{x})(\nu)(w)\,\mathrm{S}(\boldsymbol{x})(\nu)(b_0)]\\
=&[\mathrm{S}(\boldsymbol{x})(\nu)(a_0)\left(\prod_{k=1}^{m}[\lambda\boldsymbol{x}w_k]\,\mathrm{S}(\boldsymbol{x})(\nu)(a_k)\right)\mathrm{S}(\boldsymbol{x})(\nu)(b_0)]\\
=&\,\mathrm{S}(\boldsymbol{x})(\nu)([a_0)\left(\prod_{k=1}^{m-1}[\lambda\boldsymbol{x}w_k]\,\mathrm{S}(\boldsymbol{x})(\nu)(a_k)\right)[\lambda\boldsymbol{x}w_m]\,\mathrm{S}(\boldsymbol{x})(\nu)(a_m b_0])\\
=&\,\mathrm{S}(\boldsymbol{x})(\nu)(c_0)\left(\prod_{k=1}^{m-1}[\lambda\boldsymbol{x}p_k]\,\mathrm{S}(\boldsymbol{x})(\nu)(c_k)\right)[\lambda\boldsymbol{x}p_m]\,\mathrm{S}(\boldsymbol{x})(\nu)(c_m)\\
=&\,\mathrm{S}(\boldsymbol{x})(\nu)(c_0)\left(\prod_{k=1}^{m}[\lambda\boldsymbol{x}p_k]\,\mathrm{S}(\boldsymbol{x})(\nu)(c_k)\right).
\end{aligned}
$$

If otherwise $m = 0$ and $l > 0$, let $c_0 = [a_0 b_0$, $p_k = q_k$ for all $k \in [1, l]$, $c_k = b_k]$ for all $k \in [1, l[$, and $c_l = b_l]$. Then

$$\begin{aligned}
[wq] &= [a_0 \left(\prod_{k=1}^{0} [\lambda x w_k] a_k\right) b_0 \left(\prod_{j=1}^{l} [\lambda x q_j] b_j\right)] \\
&= [a_0 b_0 \left(\prod_{j=1}^{l-1} [\lambda x q_j] b_j\right) [\lambda x q_l] b_l] \\
&= c_0 \left(\prod_{k=1}^{l-1} [\lambda x p_k] c_k\right) [\lambda x p_l] c_l \\
&= c_0 \left(\prod_{k=1}^{l} [\lambda x p_k] c_k\right).
\end{aligned}$$

Furthermore

$$p_k \in \{q_k \,|\, 0 < k \leq l\} \subseteq U.$$

Thus w and q are well-formed if $[wq]$ is well-formed, whereby $p_k \in W$ for all $k \in [1, m + l]$, like above. Finally

$$\begin{aligned}
&\mathrm{SB}(\boldsymbol{x})(\boldsymbol{y})([wq]) \\
=&[\mathrm{SB}(\boldsymbol{x})(\boldsymbol{y})(w)\, \mathrm{SB}(\boldsymbol{x})(\boldsymbol{y})(q)] \\
=&[a_0 \left(\prod_{k=1}^{0} \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda x w_k]) a_k\right) b_0 \left(\prod_{j=1}^{l} \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda x q_j]) b_j\right)] \\
=&[a_0 b_0 \left(\prod_{j=1}^{l-1} \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda x q_j]) b_j\right) \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda x q_l]) b_l] \\
=&c_0 \left(\prod_{k=1}^{l-1} \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda x p_k]) c_k\right) \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda x p_l]) c_l \\
=&c_0 \left(\prod_{k=1}^{l} \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})([\lambda x p_k]) c_k\right)
\end{aligned}$$

and

$$
\begin{aligned}
&\mathrm{SF}(x)(\nu)([wq]) \\
=&[\mathrm{SF}(x)(\nu)(w)\,\mathrm{SF}(x)(\nu)(q)] \\
=&[\mathrm{S}(x)(\nu)(a_0)\left(\prod_{k=1}^{0}[\lambda x w_k]\,\mathrm{S}(x)(\nu)(a_k)\right)\mathrm{S}(x)(\nu)(b_0) \\
&\left(\prod_{j=1}^{l}[\lambda x q_j]\,\mathrm{S}(x)(\nu)(b_k)\right)] \\
=&\,\mathrm{S}(x)(\nu)([a_0 b_0)\left(\prod_{j=1}^{l-1}[\lambda x q_j]\,\mathrm{S}(x)(\nu)(b_j)\right)[\lambda x q_l]\,\mathrm{S}(x)(\nu)(b_l)] \\
=&\,\mathrm{S}(x)(\nu)(c_0)\left(\prod_{k=1}^{l-1}[\lambda x p_k]\,\mathrm{S}(x)(\nu)(c_k)\right)[\lambda x p_l]\,\mathrm{S}(x)(\nu)(c_l) \\
=&\,\mathrm{S}(x)(\nu)(c_0)\left(\prod_{k=1}^{l}[\lambda x p_k]\,\mathrm{S}(x)(\nu)(c_k)\right).
\end{aligned}
$$

If finally $m > 0$ and $l > 0$, let $c_0 = [a_0$, $p_k = w_k$ for all $k \in [1, m]$, $p_k = q_{k-m}$ for all $k \in [m+1, m+l]$, $c_k = a_k$ for all $k \in [1, m[$, $c_m = a_m b_0$, $c_k = b_{k-m}]$ for all $k \in [m+1, m+l[$, and $c_{m+l} = b_l]$. Then

$$
\begin{aligned}
[wq] &= [a_0\left(\prod_{k=1}^{m}[\lambda x w_k]a_k\right)b_0\left(\prod_{j=1}^{l}[\lambda x q_j]b_j\right)] \\
&= c_0\left(\prod_{k=1}^{m-1}[\lambda x w_k]a_k\right)[\lambda x w_m]a_m b_0\left(\prod_{j=1}^{l-1}[\lambda x q_j]b_j\right)[\lambda x q_l]b_l] \\
&= c_0\left(\prod_{k=1}^{m-1}[\lambda x p_k]c_k\right)[\lambda x p_m]c_m\left(\prod_{k=m+1}^{m+l-1}[\lambda x q_{k-m}]b_{k-m}\right)[\lambda x p_{m+l}]c_{m+l} \\
&= c_0\left(\prod_{k=1}^{m}[\lambda x p_k]c_k\right)\left(\prod_{k=m+1}^{m+l-1}[\lambda x p_k]c_k\right)[\lambda x p_{m+l}]c_{m+l} \\
&= c_0\left(\prod_{k=1}^{m+l}[\lambda x p_k]c_k\right).
\end{aligned}
$$

Furthermore

$$
p_k \in \{w_k \mid 0 < k \leq m\} \cup \{q_k \mid 0 < k \leq l\} \subseteq U.
$$

If $[wq]$ is well-formed, w and q are well-formed, whereby $p_k \in W$ for all $k \in [1, m+l]$, like above. Finally

$$
\begin{aligned}
&\mathrm{SB}(x)(y)([wq]) \\
=&[\mathrm{SB}(x)(y)(w)\,\mathrm{SB}(x)(y)(q)] \\
=&[a_0 \left(\prod_{k=1}^{m} \mathrm{S}(x)(y)([\lambda x w_k])a_k \right) b_0 \left(\prod_{j=1}^{l} \mathrm{S}(x)(y)([\lambda x q_j])b_j \right)] \\
=&c_0 \left(\prod_{k=1}^{m-1} \mathrm{S}(x)(y)([\lambda x w_k])a_k \right) \mathrm{S}(x)(y)([\lambda x w_m])a_m b_0 \left(\prod_{j=1}^{l} \mathrm{S}(x)(y)([\lambda x q_j])b_j \right)] \\
=&c_0 \left(\prod_{k=1}^{m-1} \mathrm{S}(x)(y)([\lambda x p_k])c_k \right) \mathrm{S}(x)(y)([\lambda x p_m])c_m \left(\prod_{j=1}^{l} \mathrm{S}(x)(y)([\lambda x q_j])b_j \right)] \\
=&c_0 \left(\prod_{k=1}^{m} \mathrm{S}(x)(y)([\lambda x p_k])c_k \right) \left(\prod_{j=1}^{l-1} \mathrm{S}(x)(y)([\lambda x q_j])b_j \right) \mathrm{S}(x)(y)([\lambda x q_l])b_l] \\
=&c_0 \left(\prod_{k=1}^{m} \mathrm{S}(x)(y)([\lambda x p_k])c_k \right) \left(\prod_{k=m+1}^{m+l-1} \mathrm{S}(x)(y)([\lambda x q_{k-m}])b_{k-m} \right) \mathrm{S}(x)(y)([\lambda x p_{m+l}])c_{m+l} \\
=&c_0 \left(\prod_{k=1}^{m} \mathrm{S}(x)(y)([\lambda x p_k])c_k \right) \left(\prod_{k=m+1}^{m+l-1} \mathrm{S}(x)(y)([\lambda x p_k])c_k \right) \mathrm{S}(x)(y)([\lambda x p_{m+l}])c_{m+l} \\
=&c_0 \left(\prod_{k=1}^{m+l} \mathrm{S}(x)(y)([\lambda x p_k])c_k \right)
\end{aligned}
$$

and

$$
\begin{aligned}
&\mathrm{SF}(\boldsymbol{x})(v)([wq])\\
=&[\mathrm{SF}(\boldsymbol{x})(v)(w)\,\mathrm{SF}(\boldsymbol{x})(v)(q)]\\
=&[\mathrm{S}(\boldsymbol{x})(v)(a_0)\left(\prod_{k=1}^{m}[\lambda\boldsymbol{x}w_k]\,\mathrm{S}(\boldsymbol{x})(v)(a_k)\right)\mathrm{S}(\boldsymbol{x})(v)(b_0)\\
&\left(\prod_{j=1}^{l}[\lambda\boldsymbol{x}q_j]\,\mathrm{S}(\boldsymbol{x})(v)(b_k)\right)]\\
=&\,\mathrm{S}(\boldsymbol{x})(v)([a_0)\left(\prod_{k=1}^{m-1}[\lambda\boldsymbol{x}w_k]\,\mathrm{S}(\boldsymbol{x})(v)(a_k)\right)[\lambda\boldsymbol{x}w_m]\,\mathrm{S}(\boldsymbol{x})(v)(a_m b_0)\\
&\left(\prod_{j=1}^{l}[\lambda\boldsymbol{x}q_j]\,\mathrm{S}(\boldsymbol{x})(v)(b_k)\right)]\\
=&\,\mathrm{S}(\boldsymbol{x})(v)(c_0)\left(\prod_{k=1}^{m-1}[\lambda\boldsymbol{x}p_k]\,\mathrm{S}(\boldsymbol{x})(v)(c_k)\right)[\lambda\boldsymbol{x}p_m]\,\mathrm{S}(\boldsymbol{x})(v)(c_m)\\
&\left(\prod_{j=1}^{l}[\lambda\boldsymbol{x}q_j]\,\mathrm{S}(\boldsymbol{x})(v)(b_k)\right)]\\
=&\,\mathrm{S}(\boldsymbol{x})(v)(c_0)\left(\prod_{k=1}^{m}[\lambda\boldsymbol{x}p_k]\,\mathrm{S}(\boldsymbol{x})(v)(c_k)\right)\\
&\left(\prod_{j=1}^{l-1}[\lambda\boldsymbol{x}q_j]\,\mathrm{S}(\boldsymbol{x})(v)(b_j)\right)[\lambda\boldsymbol{x}q_l]\,\mathrm{S}(\boldsymbol{x})(v)(b_l)]\\
=&\,\mathrm{S}(\boldsymbol{x})(v)(c_0)\left(\prod_{k=1}^{m}[\lambda\boldsymbol{x}p_k]\,\mathrm{S}(\boldsymbol{x})(v)(c_k)\right)\\
&\left(\prod_{k=m+1}^{m+l-1}[\lambda\boldsymbol{x}q_{k-m}]\,\mathrm{S}(\boldsymbol{x})(v)(b_{k-m})\right)[\lambda\boldsymbol{x}p_{m+l}]\,\mathrm{S}(\boldsymbol{x})(v)(c_{m+l})\\
=&\,\mathrm{S}(\boldsymbol{x})(v)(c_0)\left(\prod_{k=1}^{m}[\lambda\boldsymbol{x}p_k]\,\mathrm{S}(\boldsymbol{x})(v)(c_k)\right)\\
&\left(\prod_{k=m+1}^{m+l-1}[\lambda\boldsymbol{x}p_k]\,\mathrm{S}(\boldsymbol{x})(v)(c_k)\right)[\lambda\boldsymbol{x}p_{m+l}]\,\mathrm{S}(\boldsymbol{x})(v)(c_{m+l})\\
=&\,\mathrm{S}(\boldsymbol{x})(v)(c_0)\left(\prod_{k=1}^{m+l}[\lambda\boldsymbol{x}p_k]\,\mathrm{S}(\boldsymbol{x})(v)(c_k)\right).
\end{aligned}
$$

This concludes the proof of the theorem. □

3 Formal Deduction in the Simple Theory of Types

Thus far we have regarded formulas as merely some structured set of strings, but we have not made any significant progress towards a reasonable interpretation of said formulas. In particular, almost all of the claims in the previous section concerns all formulas rather than just the well-formed ones, and we have yet to motivate the type assignment. Since we are doing type theory, one should expect that the types of well-formed formulas are highly relevant.

In this final section we will thus define a formal notion of inference in the language we constructed above. The deductive system hence defined will share many properties of familiar deductive systems, e.g. natural deduction in first order logic. We will begin this endeavour by introducing a series of abbreviations, or shorthand notations, in terms of which the axioms (Definition 3.4) and rules of inference (Definition 3.2) will be stated. This will ensure that these can be formulated in a rather natural way. We then prove some fundamental properties of the inference relation, concluding with the Deduction Theorem (3.13).

Abbreviations

Well-formed formulas written out literally (that is, as sequences of symbols) quickly become unwieldily long, while at the same time seeming slightly unintelligible. To somewhat remedy these issues, we will introduce the following names of some special formulas, for all $\alpha, \beta \in \mathcal{T}$, all $\varphi, \psi \in W$ and all $\mathbf{x} \in \mathcal{X}$.

First among these are just names of some specific variables. They are not intended to have any special meaning in the language. Rather, they are used to make some of the following formulas uniquely defined. Note that the use of the subscript α is consistent with the earlier specified use.

- $f_\alpha = a(6, \alpha)$
- $g_\alpha = a(7, \alpha)$
- $h_\alpha = a(8, \alpha)$
- $m_\alpha = a(13, \alpha)$
- $n_\alpha = a(14, \alpha)$
- $p_\alpha = a(16, \alpha)$
- $q_\alpha = a(17, \alpha)$
- $r_\alpha = a(18, \alpha)$
- $s_\alpha = a(19, \alpha)$
- $t_\alpha = a(20, \alpha)$

- $x_\alpha = a(24, \alpha)$
- $y_\alpha = a(25, \alpha)$
- $z_\alpha = a(26, \alpha)$

The following mean what you would expect them to, in order to make propositional logic feasible.

- $(\neg\varphi_o) = [N_{oo}\varphi_o]$
- $(\varphi_o \vee \psi_o) = [[A_{\langle oo\rangle o}\varphi_o]\psi_o]$
- $(\varphi_o \wedge \psi_o) = (\neg((\neg\varphi_o) \vee (\neg\psi_o)))$
- $(\varphi_o \to \psi_o) = ((\neg\varphi_o) \vee \psi_o)$
- $(\varphi_o \leftrightarrow \psi_o) = ((\varphi_o \to \psi_o) \wedge (\psi_o \to \varphi_o))$

The below are in the same spirit as above, but for predicate logic. Note that we use $\equiv$ to denote (what is to be) the formal equality, to avoid unnecessary confusion.

- $(\forall \mathbf{x}_\alpha \varphi_o) = [\Pi_{o\langle o\alpha\rangle}[\lambda \mathbf{x}_\alpha \varphi_o]]$
- $(\exists \mathbf{x}_\alpha \varphi_o) = (\neg(\forall \mathbf{x}_\alpha(\neg\varphi_o)))$
- $(\imath \mathbf{x}_\alpha \varphi_o) = [\iota_{\alpha\langle o\alpha\rangle}[\lambda \mathbf{x}_\alpha \varphi_o]]$
- $Q_{\langle o\alpha\rangle\alpha} = [\lambda x_\alpha[\lambda y_\alpha(\forall f_{o\alpha}([f_{o\alpha}x_\alpha] \to [f_{o\alpha}y_\alpha]))]]$[4]
- $(\varphi_\alpha \equiv \psi_\alpha) = [[Q_{\langle o\alpha\rangle\alpha}\varphi_\alpha]\psi_\alpha]$
- $(\varphi_\alpha \not\equiv \psi_\alpha) = (\neg(\varphi_\alpha \equiv \psi_\alpha))$

Finally, these are some special functions, included for the sake of completeness. The last two are meant as the successor operation on, and set of, natural numbers, for any type α. They are included here mostly because there are two formal axioms (see 3.4) referring to them. We refer to [2] for a derivation from these of a number of properties of the natural numbers.

- $\mathrm{Id}_{\alpha\alpha} = [\lambda x_\alpha x_\alpha]$
- $\pi_{\langle\alpha\beta\rangle\alpha} = [\lambda x_\alpha[\lambda y_\beta x_\alpha]]$
- $0_{\alpha'} = [\lambda f_{\alpha\alpha}[\mathrm{Id}_{\alpha\alpha}]]$
- $1_{\alpha'} = [\lambda f_{\alpha\alpha}[\lambda x_\alpha[f_{\alpha\alpha}x_\alpha]]]$
- $2_{\alpha'} = [\lambda f_{\alpha\alpha}[\lambda x_\alpha[f_{\alpha\alpha}[f_{\alpha\alpha}x_\alpha]]]]$

[4] In [1] is presented a system similar to the presently studied, where this function, the formal equality, is a constant, and thus a primitive notion.

- $S_{\alpha'\alpha'} = [\lambda n_{\alpha'}[\lambda f_{\alpha\alpha}[\lambda x_\alpha[f_{\alpha\alpha}[[n_{\alpha'} f_{\alpha\alpha}]x_\alpha]]]]]$
- $N_{o\alpha'} = [\lambda n_{\alpha'}(\forall f_{o\alpha'}([f_{o\alpha'} 0_{\alpha'}] \to ((\forall x_{\alpha'}([f_{o\alpha'} x_{\alpha'}] \to [f_{o\alpha'}[S_{\alpha'\alpha'} x_{\alpha'}]])) \to [f_{o\alpha'} n_{\alpha'}])))]$

These abbreviations behave just like their first order counterparts when it comes to substitutions of variables.

3.1 Lemma. *Let* $\alpha \in \mathcal{T}$*,* $\boldsymbol{x}, \boldsymbol{y} \in X_\alpha$*,* $\varphi_o, \psi_o, \theta_\alpha \in W$ *and* $F \in \{\mathrm{S}(\boldsymbol{x})(\theta), \mathrm{SF}(\boldsymbol{x})(\theta_\alpha), \mathrm{SB}(\boldsymbol{x})(\boldsymbol{y})\}$.

1. $F((\neg\varphi)) = (\neg F(\varphi))$.
2. $F((\varphi \vee \psi)) = (F(\varphi) \vee F(\psi))$.
3. $F((\varphi \wedge \psi)) = (F(\varphi) \wedge F(\psi))$.
4. $F((\varphi \to \psi)) = (F(\varphi) \to F(\psi))$.
5. *If* $\boldsymbol{z} \neq \boldsymbol{x}$*, then* $F((\forall \boldsymbol{z}\varphi)) = ((\forall \boldsymbol{z} F(\varphi)))$*. Moreover* $\mathrm{SF}(\boldsymbol{x})(\theta)((\forall \boldsymbol{x}\varphi)) = (\forall \boldsymbol{x}\varphi)$ *and* $\mathrm{SB}(\boldsymbol{x})(\boldsymbol{y})((\forall \boldsymbol{x}\varphi)) = (\forall \boldsymbol{y}\, \mathrm{S}(\boldsymbol{x})(\boldsymbol{y})(\varphi))$.

Proof. 1.

$$\begin{aligned} F((\neg\varphi)) &= F([N\varphi]) \\ &= [F(N)F(\varphi)] \\ &= [NF(\varphi)] \\ &= (\neg F(\varphi)). \end{aligned}$$

2.

$$\begin{aligned} F((\varphi \vee \psi)) &= F([[A\varphi]\psi]) \\ &= [[F(A)F(\varphi)]F(\psi)] \\ &= [[AF(\varphi)]F(\psi)] \\ &= (F(\varphi) \vee F(\psi)). \end{aligned}$$

3.

$$\begin{aligned} F((\varphi \wedge \psi)) &= F((\neg((\neg\varphi) \vee (\neg\psi)))) \\ &= (\neg((\neg F(\varphi)) \vee (\neg F(\psi)))) \\ &= (F(\varphi) \wedge F(\psi)). \end{aligned}$$

4.

$$\begin{aligned} F((\varphi \to \psi)) &= F(((\neg\varphi) \vee \psi)) \\ &= ((\neg F(\varphi)) \vee F(\psi)) \\ &= (F(\varphi) \to F(\psi)). \end{aligned}$$

5. Let $\beta \in \mathcal{T}$ and $z_\beta \in \mathcal{X} \setminus \{x_\alpha\}$. Then

$$\begin{aligned} \mathsf{F}((\forall z_\beta \varphi_o)) &= \mathsf{F}([\Pi_{o\langle o\beta\rangle}[\lambda z_\beta \varphi_o]]) \\ &= [\mathsf{F}(\Pi_{o\langle o\beta\rangle})\mathsf{F}([\lambda z_\beta \varphi_o])] \\ &= [\Pi_{o\langle o\beta\rangle}[\lambda z_\beta \mathsf{F}(\varphi_o)]] \\ &= ((\forall z_\beta \mathsf{F}(\varphi_o))). \end{aligned}$$

On the other hand

$$\begin{aligned} \mathrm{SF}(x_\alpha)(\theta_\alpha)((\forall x_\alpha \varphi_o)) &= [\mathrm{SF}(x_\alpha)(\theta_\alpha)(\Pi_{o\langle o\alpha\rangle})\,\mathrm{SF}(x_\alpha)(\theta_\alpha)([\lambda x_\alpha \varphi_o])] \\ &= [\Pi_{o\langle o\alpha\rangle}[\lambda x_\alpha \varphi_o]] \\ &= (\forall x_\alpha \varphi_o) \end{aligned}$$

and

$$\begin{aligned} \mathrm{SB}(x)(y)((\forall x \varphi)) &= [\mathrm{SB}(x_\alpha)(y_\alpha)(\Pi_{o\langle o\alpha\rangle})\,\mathrm{SB}(x_\alpha)(y_\alpha)([\lambda x_\alpha \varphi_o])] \\ &= [\Pi_{o\langle o\alpha\rangle}[\lambda y_\alpha\, \mathrm{S}(x_\alpha)(y_\alpha)(\varphi_o)]] \\ &= (\forall y_\alpha\, \mathrm{S}(x_\alpha)(y_\alpha)(\varphi_o)). \end{aligned}$$

□

Formal Proofs

3.2 Definition (Rules of inference)**.** The *rules of inference* are the following

α-conversion Let $\alpha, \beta \in \mathcal{T}$ and $x, y \in X_\alpha$. For any well-formed formulas φ_β, ψ_o such that $a\varphi b = \psi$ for some $a, b \in \mathcal{S}^*$ where x is not free in φ and y does not occur in φ, we may from ψ infer $a\,\mathrm{S}(x)(y)(\varphi)b$

β-contraction Let $\alpha, \beta \in \mathcal{T}$. For any $x \in X_\alpha$, any well-formed formulas φ_β, ψ_α and ϕ_o such that neither x nor any free variable of ψ is a bound variable of φ, and $a[[\lambda x\varphi]\psi]b = \phi$ for some $a, b \in \mathcal{S}^*$, we may from ϕ infer $a\,\mathrm{S}(x)(\psi)(\varphi)b$.

β-expansion From any well-formed formula φ_o we may infer ψ_o, where ψ is any well-formed formula from which φ could be inferred by β-contraction.

Substitution Let $\alpha \in \mathcal{T}$. For any well-formed formulas $\varphi_{o\alpha}, \psi_\alpha \in W$ and any variable x_α such that x is not free in φ, we may from $[\varphi x]$ infer $[\varphi\psi]$

Modus ponens For any well-formed formulas $\varphi_o, \psi_o \in W$ we may from $(\varphi \to \psi)$ and φ infer ψ.

Generalisation Let $\alpha \in \mathcal{T}$. For any well-formed formula $\varphi_{o\alpha}$ and any variable x_α such that x is not free in φ, we may from $[\varphi x]$ infer $[\Pi_{o\langle o\alpha\rangle}\varphi_{o\alpha}]$.

For $V \subseteq W_o$, we say that $a \in \mathcal{S}^*$ can be inferred from V by α-conversion, β-contraction, β-expansion, Substitution or Generalisation, if there is a formula $\varphi \in V$ such that a can be inferred from φ by the corresponding rule. We say that $a \in \mathcal{S}^*$ can be inferred from V by Modus ponens if there are formulas $\varphi, \psi \in V$ so that a can be inferred from φ and ψ by Modus ponens.

If $a \in \mathcal{S}^*$ can be inferred from V by any rule like above, we simply say that a can be inferred from V.

Essentially, α-conversion allows us to change the names of bound variables, β-contraction is function application and Substitution and Generalisation makes free variable behave like universal quantification in practise and symbolism, respectively.

A desirable property of the rules of inference, especially if we strive to study semantic notions, would be to preserve the property of being well-formed, to ensure that we cannot infer gibberish out of meaningful statements. Fortunately this is the case.

3.3 Theorem. *Let* $V \subseteq W_o$. *If* $a \in \mathcal{S}^*$ *can be inferred from* V *then* $a \in W_o$.

Proof. We consider cases depending on the rule by which a can be inferred from V.

α-conversion Let $\psi \in V$ be such that a can be inferred from ψ by α-conversion. Then there are $\alpha, \beta \in \mathcal{T}$, $x, y \in X_\alpha$, and $\varphi \in W_\beta$ such that $b\varphi c = \psi$ for some $b, c \in S^*$ where x is not free in φ and y does not occur in φ, and

$$a = b\,S(x)(y)(\varphi)c.$$

Now if φ is a single character string, in particular $\varphi \neq x$ since x is not free in φ. Thus

$$a = b\,S(x)(y)(\varphi)c = b\varphi c = \psi \in W_o.$$

If on the other hand φ is not a single character string, then in any case

$$S(x)(y)(\varphi) \in W_\beta$$

by Corollary 2.20. Hence

$$a = b\,S(x)(y)(\varphi)c \in W_o$$

by Lemma 2.10.

β-contraction Let $\psi \in V$ be such that a can be inferred from ψ by β-contraction. Thus there are $\alpha, \beta \in \mathcal{T}$, $x \in \mathcal{X}_\alpha$ and $\varphi_\beta, \vartheta_\alpha \in W$ such that neither x nor any free variable of ϑ is a bound variable of φ, and $b[[\lambda x \varphi]\vartheta]c = \psi$ for some $b, c \in S^*$. Furthermore $a = b\,S(x)(\vartheta)(\varphi)c$. Since clearly $[[\lambda x \varphi]\vartheta]$ is not a single character string, and $S(x)(\vartheta)(\varphi) \in W_\beta$ by Corollary 2.20,

$$a = b\,S(x)(\vartheta)(\varphi)c \in W_o$$

by Lemma 2.10.

β-expansion Let $\psi \in V$ be such that a can be inferred from ψ by β-expansion. Then $a \in W_o$ by definition.

Substitution Let $\psi \in V$ be such that a can be inferred from ψ by Substitution. Then there are $\varphi_{o\alpha}, \vartheta_\alpha \in W$ and $\mathbf{x} \in X_\alpha$ such that $\mathbf{x}$ is not free in φ, $\psi = [\varphi\mathbf{x}]$ and $a = [\varphi\vartheta] \in W_o$.

Modus ponens Let $\varphi, \psi \in V$ be such that a can be inferred from φ and ψ by Modus ponens. Then $a \in W_o$ by definition.

Generalisation Let $\psi \in V$ be such that a can be inferred from ψ by Generalisation. Hence there are $\varphi \in W_{o\alpha}$ and $\mathbf{x} \in X_\alpha$ such that $\mathbf{x}_\alpha$ is not free in φ and $\psi = [\varphi\mathbf{x}]$. Then $a = [\Pi_{o\langle o\alpha\rangle}\varphi_{o\alpha}] \in W_o$ by definition.

□

3.4 Definition (Axioms)**.** The following are the axioms of simple type theory.

1. $((p_o \vee p_o) \rightarrow p_o)$.
2. $(p_o \rightarrow (p_o \vee q_o))$.
3. $((p_o \vee q_o) \rightarrow (q_o \vee p_o))$.
4. $((p_o \rightarrow q_o) \rightarrow ((r_o \vee p_o) \rightarrow (r_o \vee q_o)))$.
5. For all $\alpha \in \mathcal{T}$: $([\Pi_{o\langle o\alpha\rangle} f_{o\alpha}] \rightarrow [f_{o\alpha} x_\alpha])$.
6. For all $\alpha \in \mathcal{T}$: $((\forall x_\alpha(p_o \vee [f_{o\alpha} x_\alpha])) \rightarrow (p_o \vee [\Pi_{o\langle o\alpha\rangle} f_{o\alpha}]))$.
7. $(\exists x_i(\exists y_i(x_i \not\equiv y_i)))$.
8. $([N_{oi'} x_{i'}] \rightarrow ([N_{oi'} y_{i'}] \rightarrow (([S_{i'i'} x_{i'}] \equiv [S_{i'i'} y_{i'}]) \rightarrow (x_{i'} \equiv y_{i'}))))$.
9. For all $\alpha \in \mathcal{T}$: $([f_{o\alpha} x_\alpha] \rightarrow ((\forall y_\alpha([f_{o\alpha} y_\alpha] \rightarrow (x_\alpha \equiv y_\alpha))) \rightarrow [f_{o\alpha}[\iota_{\alpha\langle o\alpha\rangle} f_{o\alpha}]]))$.
10. For all $\alpha, \beta \in \mathcal{T}$: $((\forall x_\beta([f_{\alpha\beta} x_\beta] \equiv [g_{\alpha\beta} x_\beta])) \rightarrow (f_{\alpha\beta} \equiv g_{\alpha\beta}))$.

10^o. $((x_o \leftrightarrow y_o) \rightarrow (x_o \equiv y_o))$.[5]

11. For all $\alpha \in \mathcal{T}$: $([f_{o\alpha} x_\alpha] \rightarrow [f_{o\alpha}[\iota_{\alpha\langle o\alpha\rangle} f_{o\alpha}]])$.

As indicated above, not all instances of the simple theory of types use all of the above axioms. For instance, in [3] Henkin proved complete the simple theory of types with axioms 1-6, 10, 10^o and 11.

Certain sub-schemata of the above axioms in themselves describes certain "logical subsystems". As will be seen below, axioms 1-4 give rise to propositional logic, while 1-6 describes predicate logic. Furthermore, axioms 1-9 encompasses elementary number theory and axioms 1-10 and 11 allows one to obtain real analysis. See Church [2] for details and a development of elementary number theory.

[5]This axiom is not part of Church's formulation in [2], although he remarks on the possibility of including it. In Henkin's description in [3] it is a part of 10.

3.5 Definition (Formal Proof)**.** Let $\Gamma \subseteq \overline{W}_o$, $\varphi \in W_o$ and $n \in \mathbb{N}$. A *formal proof* of length $n+1$ of φ on the assumptions Γ is a string $\mathfrak{P} \in (W_o)^{n+1}$ such that $\mathfrak{P}(n) = \varphi$, where for every $0 \le k \le n$, $\mathfrak{P}(k)$ is either a formal axiom, a formula of Γ, or can be inferred from $\{\mathfrak{P}(l) \,|\, 0 \le 1 < k\}$. We write

$$\Gamma \vdash \varphi$$

when there is a formal proof (of any length) of φ on the assumptions Γ. If $\varnothing \vdash \varphi$ we call φ a *formal theorem* and write

$$\vdash \varphi$$

This is Henkin's ([3]) definition of formal proof. It differs from Church's ([2]) in that Church only allows finite sets (sequences) Γ, as well as allowing variables to occur free in the formulas of Γ, with the added restriction that Substitution or Generalisation upon such variables is not allowed. We stick with Henkin since it (in my opinion) gives a cleaner notion of proof, and since by allowing non-logical constants the need for free variables in Γ is considerably decreased.

3.6 Lemma. *In a simple theory of types endowed with axioms 1-4 of Definition 3.4, the following hold:*

1. $\vdash ((p_o \to q_o) \to ((r_o \to p_o) \to (r_o \to q_o)))$.
2. $\vdash (p_o \to p_o)$.
3. $\vdash (p_o \to (q_o \vee p_o))$.
4. $\vdash (p_o \to (q_o \to p_o))$.

Proof. The proofs are given in great detail. Be warned.

1.

$((p_o \to q_o) \to ((r_o \vee p_o) \to (r_o \vee q_o)))$	by axiom 4	(1)
$[[\lambda x_o((p_o \to q_o) \to ((x_o \vee p_o) \to (x_o \vee q_o)))]r_o]$	by β-expansion	(2)
$[[\lambda x_o((p_o \to q_o) \to ((x_o \vee p_o) \to (x_o \vee q_o)))](\neg r_o)]$	by Substitution	(3)
$((p_o \to q_o) \to (((\neg r_o) \vee p_o) \to ((\neg r_o) \vee q_o)))$	by β-contraction	(4)

Hence

$$\vdash ((p_o \to q_o) \to ((r_o \to p_o) \to (r_o \to q_o)))$$

2.

$[[\lambda x_o((x_o \to q_o) \to ((r_o \to x_o) \to (r_o \to q_o)))]p_o]$	by above and β-expansion	(5)
$[[\lambda x_o((x_o \to q_o) \to ((r_o \to x_o) \to (r_o \to q_o)))](p_o \vee p_o)]$	by Substitution	(6)
$(((p_o \vee p_o) \to q_o) \to ((r_o \to (p_o \vee p_o)) \to (r_o \to q_o)))$	by β-contraction	(7)
$[[\lambda x_o(((p_o \vee p_o) \to q_o) \to ((x_o \to (p_o \vee p_o)) \to (x_o \to q_o)))]r_o]$	by β-expansion	(8)
$[[\lambda x_o(((p_o \vee p_o) \to q_o) \to ((x_o \to (p_o \vee p_o)) \to (x_o \to q_o)))]p_o]$	by Substitution	(9)
$(((p_o \vee p_o) \to q_o) \to ((p_o \to (p_o \vee p_o)) \to (p_o \to q_o)))$	by β-contraction	(10)
$[[\lambda x_o(((p_o \vee p_o) \to x_o) \to ((p_o \to (p_o \vee p_o)) \to (p_o \to x_o)))]q_o]$	by β-expansion	(11)
$[[\lambda x_o(((p_o \vee p_o) \to x_o) \to ((p_o \to (p_o \vee p_o)) \to (p_o \to x_o)))]p_o]$	by Substitution	(12)
$(((p_o \vee p_o) \to p_o) \to ((p_o \to (p_o \vee p_o)) \to (p_o \to p_o)))$	by β-contraction	(13)

$((p_o \lor p_o) \to p_o)$	by axiom 1	(14)
$((p_o \to (p_o \lor p_o)) \to (p_o \to p_o))$	by 13, 14 and Modus ponens	(15)
$(p_o \to (p_o \lor q_o))$	by axiom 2	(16)
$[[\lambda x_o(p_o \to (p_o \lor x_o))]q_o]$	by β-expansion	(17)
$[[\lambda x_o(p_o \to (p_o \lor x_o))]p_o]$	by Substitution	(18)
$(p_o \to (p_o \lor p_o))$	by β-contraction	(19)
$(p_o \to p_o)$	by 15, 19 and Modus ponens	(20)

3.

$[[\lambda y_o[[\lambda x_o((x_o \to y_o) \to ((r_o \to x_o) \to (r_o \to y_o)))]p_o]]q_o]$	by 5 and β-expansion	(21)
$[[\lambda y_o[[\lambda x_o((x_o \to y_o) \to ((r_o \to x_o) \to (r_o \to y_o)))]p_o]](q_o \lor p_o)]$	by Substitution	(22)
$[[\lambda x_o((x_o \to (q_o \lor p_o)) \to ((r_o \to x_o) \to (r_o \to (q_o \lor p_o))))]p_o]$	by β-contraction	(23)
$[[\lambda x_o((x_o \to (q_o \lor p_o)) \to ((r_o \to x_o) \to (r_o \to (q_o \lor p_o))))](p_o \lor q_o)]$	by Substitution	(24)
$(((p_o \lor q_o) \to (q_o \lor p_o)) \to ((r_o \to (p_o \lor q_o)) \to (r_o \to (q_o \lor p_o))))$	by β-contraction	(25)
$[[\lambda x_o(((p_o \lor q_o) \to (q_o \lor p_o)) \to ((x_o \to (p_o \lor q_o)) \to (x_o \to (q_o \lor p_o))))]r_o]$	by β-expansion	(26)
$[[\lambda x_o(((p_o \lor q_o) \to (q_o \lor p_o)) \to ((x_o \to (p_o \lor q_o)) \to (x_o \to (q_o \lor p_o))))]p_o]$	by Substitution	(27)
$(((p_o \lor q_o) \to (q_o \lor p_o)) \to ((p_o \to (p_o \lor q_o)) \to (p_o \to (q_o \lor p_o))))$	by β-contraction	(28)
$((p_o \lor q_o) \to (q_o \lor p_o))$	by axiom 3	(29)
$((p_o \to (p_o \lor q_o)) \to (p_o \to (q_o \lor p_o)))$	by 28, 29 and Modus ponens	(30)
$(p_o \to (p_o \lor q_o))$	by axiom 2	(31)
$(p_o \to (q_o \lor p_o))$	by 30, 31 and Modus ponens	(32)

4.

$[[\lambda x_o(p_o \to (x_o \lor p_o))]q_o]$	by β-expansion	(33)
$[[\lambda x_o(p_o \to (x_o \lor p_o))](\neg q_o)]$	by Substitution	(34)
$(p_o \to ((\neg q_o) \lor p_o))$	by β-contraction	(35)

Hence $\vdash (p_o \to (q_o \to p_o))$ as desired.

□

3.7 Proposition. *In a simple theory of types with axioms 1-6,* $\vdash ((\forall x_\alpha[f_{o\alpha}x_\alpha]) \to [f_{o\alpha}x_\alpha])$ *for every* $\alpha \in \mathcal{T}$.

Proof.

$([\Pi_{o\langle o\alpha\rangle}f_{o\alpha}] \to [f_{o\alpha}x_\alpha])$	by axiom 5	(36)
$[[\lambda g_{o\alpha}([\Pi_{o\langle o\alpha\rangle}g_{o\alpha}] \to [g_{o\alpha}x_\alpha])]f_{o\alpha}]$	by β-expansion	(37)
$[[\lambda g_{o\alpha}([\Pi_{o\langle o\alpha\rangle}g_{o\alpha}] \to [g_{o\alpha}x_\alpha])][\lambda x_\alpha[f_{o\alpha}x_\alpha]]]$	by Substitution	(38)
$([\Pi_{o\langle o\alpha\rangle}[\lambda x_\alpha[f_{o\alpha}x_\alpha]]] \to [[\lambda x_\alpha[f_{o\alpha}x_\alpha]]x_\alpha])$	by β-contraction	(39)
$([\Pi_{o\langle o\alpha\rangle}[\lambda x_\alpha[f_{o\alpha}x_\alpha]]] \to [f_{o\alpha}x_\alpha])$	by β-contraction	(40)

Hence $\vdash ((\forall x_\alpha[f_{o\alpha}x_\alpha]) \to [f_{o\alpha}x_\alpha])$. □

Derived Rules of Inference

Unsurprisingly, as can be seen above, making formal proofs is quite a tedious business. Fortunately, there are some general properties of the relation $\vdash$ which can be used to prove that a formal proof exists, and hence be used much as the rules of inference (Definition 3.2), courtesy of the following remark.

Remark 10. If ϕ may be inferred from φ by any of the rules of inference, we have that

$$\Gamma \vdash \varphi \Rightarrow \Gamma \vdash \phi$$

by definition. Furthermore, if $\varphi \in \Gamma$ or is an axiom (3.4), then

$$\Gamma \vdash \varphi.$$

First among these, the below properties can be seen as generalisations of the rules of inference.

3.8 Theorem. *Whichever set of axioms is chosen, we have the following relations, for every $\alpha \in \mathcal{T}$, $\Gamma \subseteq \overline{W}_o$, $\varphi_o, \phi_o, \psi_\alpha \in W$, and every $\mathbf{x} \in X_\alpha$.*

1. *For every $\mathbf{y} \in X_\alpha$ which does not occur in φ*

$$\Gamma \vdash \varphi \Leftrightarrow \Gamma \vdash \mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi).$$

2. *If no free variable of ψ, except possibly $\mathbf{x}$, is bound in φ, then*

$$\Gamma \vdash \varphi \Rightarrow \Gamma \vdash \mathrm{SF}(\mathbf{x})(\psi)(\varphi).$$

3.

$$\Gamma \vdash \varphi \Rightarrow \Gamma \vdash (\forall \mathbf{x} \varphi)$$

Proof. 1. Let $\mathbf{y} \in X_\alpha$ be such that $\mathbf{y} \not\prec \varphi$. By Theorem 2.25 φ can be written as

$$\varphi = r_0 \left(\prod_{k=1}^{n} [\lambda \mathbf{x} \varphi_k] r_k \right)$$

such that

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi) = r_0 \left(\prod_{k=1}^{n} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} \varphi_k]) r_k \right).$$

Assume that $\Gamma \vdash \varphi$. By iterated α-conversion

$$\Gamma \vdash r_0 \left(\prod_{k=1}^{n} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} \varphi_k]) r_k \right).$$

Hence

$$\Gamma \vdash \mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi).$$

Conversely, assume that $\Gamma \vdash \mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)$, i.e.

$$\Gamma \vdash r_0 \left(\prod_{k=1}^{n} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} \varphi_k]) r_k \right).$$

Since $\mathbf{y}$ is not free in

$$\mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} \varphi_k]) = [\lambda \mathbf{y}\, \mathrm{S}(\mathbf{x})(\mathbf{y})(\varphi_k)]$$

for any $k \in [0, n]$, iterated α-conversion gives

$$\Gamma \vdash r_0 \left(\prod_{k=1}^{n} \mathrm{S}(\mathbf{y})(\mathbf{x})(\mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} \varphi_k])) r_k \right)$$

where

$$\begin{aligned} r_0 \left(\prod_{k=1}^{n} \mathrm{S}(\mathbf{y})(\mathbf{x})(\mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} \varphi_k])) r_k \right) &= r_0 \left(\prod_{k=1}^{n} \mathrm{S}(\mathbf{x})(\mathbf{x})([\lambda \mathbf{x} \varphi_k]) r_k \right) \\ &= r_0 \left(\prod_{k=1}^{n} [\lambda \mathbf{x} \varphi_k] r_k \right) \\ &= \varphi. \end{aligned}$$

Hence $\Gamma \vdash \varphi$.

2. If $\mathbf{x}$ is not free in φ, then $\mathrm{SF}(\mathbf{x})(\psi)(\varphi) = \varphi$ and the statement holds. Otherwise let $\mathbf{y} \in X_\alpha$ be such that $\mathbf{y} \not\preceq \mathbf{x}, \varphi, \psi$. By Theorem 2.25 there is a representation

$$\varphi = r_0 \left(\prod_{k=1}^{n} [\lambda \mathbf{x} \varphi_k] r_k \right)$$

such that

$$\begin{aligned} \mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi) &= r_0 \left(\prod_{k=1}^{n} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda \mathbf{x} \varphi_k]) r_k \right) \\ \mathrm{SF}(\mathbf{x})(\psi)(\varphi) &= \mathrm{S}(\mathbf{x})(\psi)(r_0) \left(\prod_{k=1}^{n} [\lambda \mathbf{x} \varphi_k]\, \mathrm{S}(\mathbf{x})(\psi)(r_k) \right). \end{aligned}$$

By 1

$$\Gamma \vdash \mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)$$

whereby

$$\Gamma \vdash [[\lambda \mathbf{x}\, \mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)] \mathbf{x}]$$

by β-expansion, since $\mathbf{x}$ is not bound in $\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)$ and $\mathrm{S}(\mathbf{x})(\mathbf{x})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)) = \mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)$. Hence we may use Substitution for $\mathbf{x}$ and ψ, that is

$$\Gamma \vdash [[\lambda\mathbf{x}\,\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)]\psi]$$

since $\mathbf{x}$ is certainly not free in $[\lambda\mathbf{x}\,\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)]$. Furthermore no free variable of ψ, except possibly $\mathbf{x}$, is bound in φ. Thus, by Lemmas 2.12 and 2.22, no free variable of ψ is bound in $\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)$, whence

$$\Gamma \vdash \mathrm{S}(\mathbf{x})(\psi)(\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi))$$

by β-contraction. Observe that

$$\begin{aligned}\mathrm{S}(\mathbf{x})(\psi)(\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)) &= \mathrm{S}(\mathbf{x})(\psi)(r_0\left(\prod_{k=1}^{n}\mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda\mathbf{x}\varphi_k])r_k\right))\\ &= \mathrm{S}(\mathbf{x})(\psi)(r_0)\left(\prod_{k=1}^{n}\mathrm{S}(\mathbf{x})(\psi)(\mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda\mathbf{x}\varphi_k]))\,\mathrm{S}(\mathbf{x})(\psi)(r_k)\right)\\ &= \mathrm{S}(\mathbf{x})(\psi)(r_0)\left(\prod_{k=1}^{n}\mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda\mathbf{x}\varphi_k])\,\mathrm{S}(\mathbf{x})(\psi)(r_k)\right).\end{aligned}$$

As above, $\mathbf{y}$ is not free in $\mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda\mathbf{x}\varphi_k])$ for any $k \in [0, n]$, so by iterated α-conversion

$$\Gamma \vdash \mathrm{S}(\mathbf{x})(\psi)(r_0)\left(\prod_{k=1}^{n}\mathrm{S}(\mathbf{y})(\mathbf{x})(\mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda\mathbf{x}\varphi_k]))\,\mathrm{S}(\mathbf{x})(\psi)(r_k)\right)$$

where indeed

$$\begin{aligned}&\mathrm{S}(\mathbf{x})(\psi)(r_0)\left(\prod_{k=1}^{n}\mathrm{S}(\mathbf{y})(\mathbf{x})(\mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda\mathbf{x}\varphi_k]))\,\mathrm{S}(\mathbf{x})(\psi)(r_k)\right)\\ =&\,\mathrm{S}(\mathbf{x})(\psi)(r_0)\left(\prod_{k=1}^{n}\mathrm{S}(\mathbf{x})(\mathbf{x})([\lambda\mathbf{x}\varphi_k])\,\mathrm{S}(\mathbf{x})(\psi)(r_k)\right)\\ =&\,\mathrm{S}(\mathbf{x})(\psi)(r_0)\left(\prod_{k=1}^{n}[\lambda\mathbf{x}\varphi_k]\,\mathrm{S}(\mathbf{x})(\psi)(r_k)\right)\\ =&\,\mathrm{SF}(\mathbf{x})(\psi)(\varphi).\end{aligned}$$

Thus $\Gamma \vdash \mathrm{SF}(\mathbf{x})(\psi)(\varphi)$.

3. Let $\mathbf{y} \in X_\alpha$ be such that $\mathbf{y} \not\preceq \mathbf{x}, \varphi$. By Theorem 2.25 there is a representation

$$\varphi = r_0\left(\prod_{k=1}^{n}[\lambda\mathbf{x}\varphi_k]r_k\right)$$

such that

$$\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi) = r_0\left(\prod_{k=1}^{n} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda\mathbf{x}\varphi_k])r_k\right).$$

Assume that $\Gamma \vdash \varphi$. Then 1 ensures that

$$\Gamma \vdash \mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi).$$

Since $\mathbf{x}$ is not bound in $\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)$, and $\mathrm{S}(\mathbf{x})(\mathbf{x})(\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)) = \mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)$, we have that

$$\Gamma \vdash [[\lambda\mathbf{x}\,\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)]\mathbf{x}]$$

by β-expansion. Now $\mathbf{x}$ is not free in $[\lambda\mathbf{x}\,\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)]$, whence

$$\Gamma \vdash [\Pi_{o\langle o\alpha\rangle}[\lambda\mathbf{x}\,\mathrm{SB}(\mathbf{x})(\mathbf{y})(\varphi)]]$$

by Generalisation, that is,

$$\Gamma \vdash (\forall\mathbf{x}r_0\left(\prod_{k=1}^{n} \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda\mathbf{x}\varphi_k])r_k\right)).$$

Since for every $k \in [1,n]$, $\mathbf{y}$ is not free in $\mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda\mathbf{x}\varphi_k])$ and $\mathbf{x} \not\preceq \mathrm{S}(\mathbf{x})(\mathbf{y})([\lambda\mathbf{x}\varphi_k])$, iterated α-conversion gives that $\Gamma \vdash (\forall\mathbf{x}r_0\,(\prod_{k=1}^{n}[\lambda\mathbf{x}\varphi_k]r_k))$ i.e.

$$\Gamma \vdash (\forall\mathbf{x}\varphi).$$

□

The following lemma, while having a slightly unintelligible air, will mainly be used to prove such intuitive statements as: *if* $\vdash (p_o \to (q_o \to p_o))$, *then* $\vdash (\varphi \to (\psi \to \varphi))$ *for all* $\varphi, \psi \in W_o$. A more precise (though in the current context undefined) statement is that we may substitute a free variable $\mathbf{x}$ for a well-formed formula ψ in a provable well-formed formula φ to obtain a new provable formula, as long as the variable $\mathbf{x}$ does not occur within the scope of a binding of a free variable of ψ in φ.

3.9 Proposition. *Let* $\Gamma \subseteq \overline{W}_o$ *and* $\mathfrak{x} \in \mathcal{X}^*$, *and let* $F : \mathbb{N} \longrightarrow (U^U)^{U^*}$ *be defined thus:*

$$F(0)(\mathfrak{v})(u) = u$$
$$F(n+1)(\mathfrak{v})(u) = \begin{cases} \mathrm{SF}(\mathfrak{x}(n))(\mathfrak{v}(n))(F(n)(\mathfrak{v})(u)) & \text{if } n < \mathrm{lh}(\mathfrak{x}), \mathrm{lh}(\mathfrak{v}) \\ N_{oo} & \text{otherwise} \end{cases}$$

for all $n \in \mathbb{N}$, *all* $\mathfrak{v} \in U^*$ *and all* $u \in U$.

For all $n \in \mathbb{N}$, $\Psi \in W^*$ *such that* $n < \mathrm{lh}(\Psi) = \mathrm{lh}(\mathfrak{x})$, $T(\Psi(k)) = T(\mathfrak{x}(k))$ *and* $\mathfrak{x}(k) \notin VF(\Psi(l))$ *for all* $k, l \in [0, \mathrm{lh}(\mathfrak{x})[$, *and* $\varphi \in W_o$ *such that* $VB(\varphi) = \varnothing$ *we have that*

$$\Gamma \vdash F(n)(\psi)(\varphi) \Rightarrow \Gamma \vdash F(n+1)(\Psi)(\varphi).$$

Proof. We first prove the following

Sublemma 3.9.1. *Let* $\Psi \in W^*$ *and* $\varphi \in W$ *such that* $\mathrm{lh}(\Psi) = \mathrm{lh}(\mathfrak{x})$ *and* $\mathsf{T}(\Psi(k)) = \mathsf{T}(\mathfrak{x}(k))$ *for all* $k \in [0, \mathrm{lh}(\mathfrak{x})[$. *For all* $n \in [0, \mathrm{lh}(\mathfrak{x})]$,

$$\mathsf{F}(n)(\Psi)(\varphi) \in W_{\mathsf{T}(\varphi)}.$$

Proof. By induction on n.

- By definition

$$\mathsf{F}(0)(\Psi)(\varphi) = \varphi \in W_{\mathsf{T}(\varphi)}.$$

- Assume that $m \in [0, \mathrm{lh}(\mathfrak{x})]$ is such that $\mathsf{F}(m)(\Psi)(\varphi) \in W_{\mathsf{T}(\varphi)}$, and consider $m+1$. If $m+1 \notin [0, \mathrm{lh}(\mathfrak{x})]$ there is nothing to prove. Otherwise

$$\mathsf{F}(m+1)(\Psi)(\varphi) = \mathrm{SF}(\mathfrak{x}(m))(\Psi(m))(\mathsf{F}(m)(\Psi)(\varphi)) \in W_{\mathsf{T}(\varphi)}.$$

  by Lemma 2.16.

□

Sublemma 3.9.2. *Let* $u, v \in U$ *and* $\mathfrak{v} \in U^*$ *be such that* $\mathrm{lh}(\mathfrak{v}) = \mathrm{lh}(\mathfrak{x})$. *For all* $n \in [0, \mathrm{lh}(\mathfrak{x})]$,

$$\mathsf{F}(n)(\mathfrak{v})([uv]) = [\mathsf{F}(n)(\mathfrak{v})(u)\mathsf{F}(n)(\mathfrak{v})(v)].$$

Proof. By induction on n.

- By definition

$$\begin{aligned}\mathsf{F}(0)(\mathfrak{v})([uv]) &= [uv] \\ &= [\mathsf{F}(0)(\mathfrak{v})(u)\mathsf{F}(0)(\mathfrak{v})(v)].\end{aligned}$$

- Assume that $m \in [0, \mathrm{lh}(\mathfrak{x})]$ is such that $\mathsf{F}(m)(\mathfrak{v})([uv]) = [\mathsf{F}(m)(\mathfrak{v})(u)\mathsf{F}(m)(\mathfrak{v})(v)]$, and consider $m+1$. If $m+1 \notin [0, \mathrm{lh}(\mathfrak{x})]$ there is nothing to prove. Otherwise

$$\begin{aligned}\mathsf{F}(m+1)(\mathfrak{v})([uv]) &= \mathrm{SF}(\mathfrak{x}(m))(\mathfrak{v}(m))(\mathsf{F}(m)(\mathfrak{v})([uv])) \\ &= \mathrm{SF}(\mathfrak{x}(m))(\mathfrak{v}(m))([\mathsf{F}(m)(\mathfrak{v})(u)\mathsf{F}(m)(\mathfrak{v})(v)]) \\ &= [\mathrm{SF}(\mathfrak{x}(m))(\mathfrak{v}(m))(\mathsf{F}(m)(\mathfrak{v})(u))\,\mathrm{SF}(\mathfrak{x}(m))(\mathfrak{v}(m))(\mathsf{F}(m)(\mathfrak{v})(v))] \\ &= [\mathsf{F}(m+1)(\mathfrak{v})(u)\mathsf{F}(m+1)(\mathfrak{v})(v)]\end{aligned}$$

  by definition.

□

Sublemma 3.9.3. *Let* $\mathsf{s} \in \mathcal{B}$ *and* $\mathfrak{v} \in \mathsf{U}^*$ *be such that* $\mathrm{lh}(\mathfrak{v}) = \mathrm{lh}(\mathfrak{x})$ *and* $\mathfrak{x}(\mathsf{k}) \notin \mathsf{VF}(\mathfrak{v}(\mathsf{l}))$ *for all* $\mathsf{k}, \mathsf{l} \in [0, \mathrm{lh}(\mathfrak{x})[$.

If $\mathsf{s} = \mathfrak{x}(\mathsf{k})$ *for some* $\mathsf{k} \in [0, \mathrm{lh}(\mathfrak{x})[$, *then for all* $\mathsf{n} \in [0, \mathrm{lh}(\mathfrak{x})]$

$$\mathsf{F}(\mathsf{n})(\mathfrak{v})(\mathsf{s}) = \begin{cases} \mathfrak{v}(\mathsf{l}) & \textit{if } \mathsf{n} > \mathsf{l} \\ \mathsf{s} & \textit{otherwise} \end{cases}$$

where $\mathsf{l} \in [0, \mathrm{lh}(\mathfrak{x})[$ *is the least such that* $\mathfrak{x}(\mathsf{l}) = \mathsf{s}$.

Otherwise

$$\mathsf{F}(\mathsf{n})(\mathfrak{v})(\mathsf{s}) = \mathsf{s}$$

for all $\mathsf{n} \in [0, \mathrm{lh}(\mathfrak{x})]$.

Proof. By induction on n.

- Firstly, $0 \leq \mathsf{l}$ for every $\mathsf{l} \in [0, \mathrm{lh}(\mathfrak{x})[$, and

$$\mathsf{F}(0)(\mathfrak{v})(\mathsf{s}) = \mathsf{s}.$$

- Assume that $\mathsf{m} \in [0, \mathrm{lh}(\mathfrak{x})]$ is such that if $\mathsf{s} = \mathfrak{x}(\mathsf{k})$ for some $\mathsf{k} \in [0, \mathrm{lh}(\mathfrak{x})[$, then

$$\mathsf{F}(\mathsf{m})(\mathfrak{v})(\mathsf{s}) = \begin{cases} \mathfrak{v}(\mathsf{l}) & \text{if } \mathsf{m} > \mathsf{l} \\ \mathsf{s} & \text{otherwise} \end{cases}$$

where $\mathsf{l} \in [0, \mathrm{lh}(\mathfrak{x})[$ is the least such that $\mathfrak{x}(\mathsf{l}) = \mathsf{s}$, and if $\mathsf{s} \neq \mathfrak{x}(\mathsf{k})$ for all $\mathsf{k} \in [0, \mathrm{lh}(\mathfrak{x})[$ then

$$\mathsf{F}(\mathsf{m})(\mathfrak{v})(\mathsf{s}) = \mathsf{s},$$

and consider $\mathsf{m} + 1$. If $\mathsf{m} + 1 \notin [0, \mathrm{lh}(\mathfrak{x})]$ there is nothing to prove. Otherwise, assume that $\mathsf{s} = \mathfrak{x}(\mathsf{k})$ for some $\mathsf{k} \in [0, \mathrm{lh}(\mathfrak{x})[$, and let $\mathsf{l} \in [0, \mathrm{lh}(\mathfrak{x})[$ be the least such. Now, if $\mathsf{m} + 1 > \mathsf{l}$ then either $\mathsf{m} > \mathsf{l}$ or $\mathsf{m} = \mathsf{l}$. In the former case

$$\begin{aligned} \mathsf{F}(\mathsf{m}+1)(\mathfrak{v})(\mathsf{s}) &= \mathrm{SF}(\mathfrak{x}(\mathsf{m}))(\mathfrak{v}(\mathsf{m}))(\mathsf{F}(\mathsf{m})(\mathfrak{v})(\mathsf{s})) \\ &= \mathrm{SF}(\mathfrak{x}(\mathsf{m}))(\mathfrak{v}(\mathsf{m}))(\mathfrak{v}(\mathsf{l})) \\ &= \mathfrak{v}(\mathsf{l}) \end{aligned}$$

by induction hypothesis, since $\mathfrak{x}(\mathsf{m})$ is not free in $\mathfrak{v}(\mathsf{l})$; while in the latter case,

$$\begin{aligned} \mathsf{F}(\mathsf{m}+1)(\mathfrak{v})(\mathsf{s}) &= \mathrm{SF}(\mathfrak{x}(\mathsf{m}))(\mathfrak{v}(\mathsf{m}))(\mathsf{F}(\mathsf{m})(\mathfrak{v})(\mathsf{s})) \\ &= \mathrm{SF}(\mathfrak{x}(\mathsf{m}))(\mathfrak{v}(\mathsf{m}))(\mathsf{s}) \\ &= \mathfrak{v}(\mathsf{m}) \\ &= \mathfrak{v}(\mathsf{l}). \end{aligned}$$

If $\mathfrak{m}+1 \leq \mathfrak{l}$, then $\mathfrak{m} < \mathfrak{l}$ and $\mathsf{s} \neq \mathfrak{x}(\mathfrak{m})$, whence

$$\begin{aligned} \mathsf{F}(\mathfrak{m}+1)(\mathfrak{v})(\mathsf{s}) &= \mathrm{SF}(\mathfrak{x}(\mathfrak{m}))(\mathfrak{v}(\mathfrak{m}))(\mathsf{F}(\mathfrak{m})(\mathfrak{v})(\mathsf{s})) \\ &= \mathrm{SF}(\mathfrak{x}(\mathfrak{m}))(\mathfrak{v}(\mathfrak{m}))(\mathsf{s}) \\ &= \mathsf{s} \end{aligned}$$

by definition.

Finally assume that $\mathsf{s} \neq \mathfrak{x}(\mathsf{k})$ for all $\mathsf{k} \in [0, \mathrm{lh}(\mathfrak{x})[$. In particular $\mathsf{s} \neq \mathfrak{x}(\mathfrak{m})$, whereby

$$\begin{aligned} \mathsf{F}(\mathfrak{m}+1)(\mathfrak{v})(\mathsf{s}) &= \mathrm{SF}(\mathfrak{x}(\mathfrak{m}))(\mathfrak{v}(\mathfrak{m}))(\mathsf{F}(\mathfrak{m})(\mathfrak{v})(\mathsf{s})) \\ &= \mathrm{SF}(\mathfrak{x}(\mathfrak{m}))(\mathfrak{v}(\mathfrak{m}))(\mathsf{s}) \\ &= \mathsf{s} \end{aligned}$$

again by definition.

□

Sublemma 3.9.4. *Let* $\boldsymbol{y}, \boldsymbol{z} \in \mathcal{X}$, $\mathfrak{n} \in [0, \mathrm{lh}(\mathfrak{x})]$ *and* $\mathfrak{v} \in \mathsf{U}^*$ *be such that* $\mathrm{lh}(\mathfrak{v}) = \mathrm{lh}(\mathfrak{x})$ *and* $\mathfrak{x}(\mathsf{k}) \notin \mathsf{VF}(\mathfrak{v}(\mathfrak{l}))$ *for all* $\mathsf{k}, \mathfrak{l} \in [0, \mathrm{lh}(\mathfrak{x})[$. *Let* $\mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathfrak{v}) \in \mathsf{U}^{\mathrm{lh}(\mathfrak{v})}$ *be such that* $\mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathfrak{v})(\mathsf{k}) = \mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathfrak{v}(\mathsf{k}))$ *for all* $\mathsf{k} \in [0, \mathrm{lh}(\mathfrak{v})[$. *Then*

$$\mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathsf{F}(\mathfrak{n})(\mathfrak{v})(\mathfrak{u})) = \mathsf{F}(\mathfrak{n})(\mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathfrak{v}))(\mathfrak{u})$$

for all $\mathfrak{u} \in \mathsf{U}$ *such that* $\mathsf{VB}(\mathfrak{u}) = \varnothing$.

Proof. By induction on $\mathfrak{u}$.

- Let $\mathsf{s} \in \mathcal{B}$. Of course $\mathsf{VB}(\mathsf{s}) = \varnothing$. By the previous sublemma (3.9.3) we have that if $\mathsf{s} \neq \mathfrak{x}(\mathsf{k})$ for all $\mathsf{k} \in [0, \mathrm{lh}(\mathfrak{x})[$, then $\mathsf{F}(\mathfrak{n})(\mathfrak{v})(\mathsf{s}) = \mathsf{s}$; otherwise

$$\mathsf{F}(\mathfrak{n})(\mathfrak{v})(\mathsf{s}) = \begin{cases} \mathfrak{v}(\mathfrak{l}) & \text{if } \mathfrak{n} > \mathfrak{l} \\ \mathsf{s} & \text{otherwise} \end{cases}$$

for the least $\mathfrak{l} \in [0, \mathrm{lh}(\mathfrak{x})[$ such that $\mathsf{s} \neq \mathfrak{x}(\mathfrak{l})$. In the former case

$$\begin{aligned} \mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathsf{F}(\mathfrak{n})(\mathfrak{v})(\mathsf{s})) &= \mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathsf{s}) \\ &= \mathsf{s} \\ &= \mathsf{F}(\mathfrak{n})(\mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathfrak{v}))(\mathsf{s}), \end{aligned}$$

and in the latter case

$$\begin{aligned} \mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathsf{F}(\mathfrak{n})(\mathfrak{v})(\mathsf{s})) &= \begin{cases} \mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathfrak{v}(\mathfrak{l})) & \text{if } \mathfrak{n} > \mathfrak{l} \\ \mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathsf{s}) & \text{otherwise} \end{cases} \\ &= \begin{cases} \mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathfrak{v})(\mathfrak{l}) & \text{if } \mathfrak{n} > \mathfrak{l} \\ \mathsf{s} & \text{otherwise} \end{cases} \\ &= \mathsf{F}(\mathfrak{n})(\mathrm{SB}(\boldsymbol{y})(\boldsymbol{z})(\mathfrak{v}))(\mathsf{s}) \end{aligned}$$

by 3.9.3 again. In any case

$$\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathsf{F}(n)(\mathfrak{v})(s)) = \mathsf{F}(n)(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathfrak{v}))(s).$$

- Let $\mathfrak{a} \in \mathcal{X}$ and $v \in \mathsf{U}$ be such that if $\mathsf{VB}(v) = \varnothing$ then $\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathsf{F}(n)(\mathfrak{v})(v)) = \mathsf{F}(n)(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathfrak{v}))(v)$, and consider $[\lambda\mathfrak{a}v] \in \mathsf{U}$. Since $\mathfrak{a} \in \mathsf{VB}([\lambda\mathfrak{a}v])$, there is nothing to prove.

- Let $v, w \in \mathsf{U}$ be such that if $\mathsf{VB}(q) = \varnothing$ then

$$\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathsf{F}(n)(\mathfrak{v})(q)) = \mathsf{F}(n)(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathfrak{v}))(q)$$

  for $q \in \{v, w\}$ and consider $[vw] \in \mathsf{U}$. Assume that $\mathsf{VB}([vw]) = \varnothing$. Hence $\mathsf{VB}(v) = \mathsf{VB}(w) = \varnothing$, whereby

$$\begin{aligned}\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathsf{F}(n)(\mathfrak{v})([vw])) &= \mathrm{SB}(\mathfrak{y})(\mathfrak{z})([\mathsf{F}(n)(\mathfrak{v})(v)\mathsf{F}(n)(\mathfrak{v})(w)]) \\ &= [\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathsf{F}(n)(\mathfrak{v})(v))\,\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathsf{F}(n)(\mathfrak{v})(w))] \\ &= [\mathsf{F}(n)(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathfrak{v}))(v)\mathsf{F}(n)(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathfrak{v}))(w)] \\ &= \mathsf{F}(n)(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathfrak{v}))([vw])\end{aligned}$$

  by 3.9.2.

□

Now let $n \in \mathbb{N}$ and $\varphi \in W_o$ be such that $n < \mathrm{lh}(\mathfrak{x})$ and $\mathsf{VB}(\varphi) = \varnothing$. By 3.9.1, we have that

$$\mathsf{F}(n)(\Psi)(\varphi) \in W_o$$

for all $\Psi \in W^*$ such that $\mathrm{lh}(\Psi) = \mathrm{lh}(\mathfrak{x})$ and $\mathsf{T}(\Psi(k)) = \mathsf{T}(\mathfrak{x}(k))$ for all $k \in [0, \mathrm{lh}(\mathfrak{x})[$. We conclude the proof by showing that

$$\Gamma \vdash \mathsf{F}(n)(\Psi)(\varphi) \Rightarrow \Gamma \vdash \mathsf{F}(n+1)(\Psi)(\varphi)$$

for all $\Psi \in W^*$ such that $\mathrm{lh}(\Psi) = \mathrm{lh}(\mathfrak{x})$, $\mathsf{T}(\Psi(k)) = \mathsf{T}(\mathfrak{x}(k))$ and $\mathfrak{x}(k) \notin \mathsf{VF}(\Psi(l))$ for all $k, l \in [0, \mathrm{lh}(\mathfrak{x})[$, by induction on $|\mathsf{VF}(\Psi(n)) \cap \mathsf{VB}(\mathsf{F}(n)(\Psi)(\varphi)) \setminus \{\mathfrak{x}(n)\}|$.

- Assume that $\Theta \in W^*$ is such that $\mathrm{lh}(\Theta) = \mathrm{lh}(\mathfrak{x})$, $\mathsf{T}(\Theta(k)) = \mathsf{T}(\mathfrak{x}(k))$, $\mathfrak{x}(k) \notin \mathsf{VF}(\Theta(l))$ for all $k, l \in [0, \mathrm{lh}(\mathfrak{x})[$ and $|\mathsf{VF}(\Theta(n)) \cap \mathsf{VB}(\mathsf{F}(n)(\Theta)(\varphi)) \setminus \{\mathfrak{x}(n)\}| = 0$. Thus no free variable of $\Theta(n)$, except possibly $\mathfrak{x}(n)$, is bound in $\mathsf{F}(n)(\Theta)(\varphi)$. Therefore, if

$$\Gamma \vdash \mathsf{F}(n)(\Theta)(\varphi)$$

  then it follows that

$$\Gamma \vdash \mathrm{SF}(\mathfrak{x}(n))(\Theta(n))(\mathsf{F}(n)(\Theta)(\varphi))$$

  by Theorem 3.8, and

$$\mathrm{SF}(\mathfrak{x}(n))(\Theta(n))(\mathsf{F}(n)(\Theta)(\varphi)) = \mathsf{F}(n+1)(\Theta)(\varphi)$$

  by definition.

- Let $\mathfrak{m} \in \mathbb{Z}^+$ and assume that

$$\Gamma \vdash \mathsf{F}(\mathfrak{n})(\Upsilon)(\varphi) \Rightarrow \Gamma \vdash \mathsf{F}(\mathfrak{n}+1)(\Upsilon)(\varphi)$$

for all $\Upsilon \in W^*$ such that $\mathrm{lh}(\Upsilon) = \mathrm{lh}(\mathfrak{x})$, $\mathsf{T}(\Upsilon(\mathsf{k})) = \mathsf{T}(\mathfrak{x}(\mathsf{k}))$, $\mathfrak{x}(\mathsf{k}) \notin \mathsf{VF}(\Upsilon(\mathfrak{l}))$ for all $\mathsf{k}, \mathfrak{l} \in [0, \mathrm{lh}(\mathfrak{x})[$ and $|\mathsf{VF}(\Upsilon(\mathfrak{n})) \cap \mathsf{VB}(\mathsf{F}(\mathfrak{n})(\Upsilon)(\varphi)) \setminus \{\mathfrak{x}(\mathfrak{n})\}| < \mathfrak{m}$. Furthermore, let $\Phi \in W^*$ be such that $\mathrm{lh}(\Phi) = \mathrm{lh}(\mathfrak{x})$, $\mathsf{T}(\Phi(\mathsf{k})) = \mathsf{T}(\mathfrak{x}(\mathsf{k}))$, $\mathfrak{x}(\mathsf{k}) \notin \mathsf{VF}(\Phi(\mathfrak{l}))$ for all $\mathsf{k}, \mathfrak{l} \in [0, \mathrm{lh}(\mathfrak{x})[$ and $|\mathsf{VF}(\Phi(\mathfrak{n})) \cap \mathsf{VB}(\mathsf{F}(\mathfrak{n})(\Phi)(\varphi)) \setminus \{\mathfrak{x}(\mathfrak{n})\}| = \mathfrak{m}$. Hence there is some $\beta \in \mathcal{T}$ and $\mathfrak{y}_\beta \in \mathsf{VF}(\Phi(\mathfrak{n})) \cap \mathsf{VB}(\mathsf{F}(\mathfrak{n})(\Phi)(\varphi)) \setminus \{\mathfrak{x}(\mathfrak{n})\}$. Let $\mathfrak{z} \in \mathsf{X}_\beta$ be such that $\mathfrak{z} \not\preceq \mathfrak{x}(\mathfrak{l}), \mathfrak{y}, \varphi, \Phi(\mathfrak{l}), \mathsf{F}(\mathfrak{l})(\Phi)(\varphi), \mathsf{F}(\mathfrak{n}+1)(\Phi)(\varphi)$ for all $\mathfrak{l} \in [0, \mathrm{lh}(\mathfrak{x})[$. Suppose that

$$\Gamma \vdash \mathsf{F}(\mathfrak{n})(\Phi)(\varphi).$$

Similarly to before, let $\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi) \in W^{\mathrm{lh}(\Phi)}$ denote the string such that

$$\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi)(\mathsf{k}) = \mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi(\mathsf{k}))$$

for all $\mathsf{k} \in [0, \mathrm{lh}(\Phi)[$. Then $\mathsf{T}(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi)(\mathsf{k})) = \mathsf{T}(\Phi(\mathsf{k})) = \mathsf{T}(\mathfrak{x}(\mathsf{k}))$ for all $\mathsf{k} \in [0, \mathrm{lh}(\mathfrak{x})[$ by 2.16. Since

$$\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathsf{F}(\mathfrak{n})(\Phi)(\varphi)) = \mathsf{F}(\mathfrak{n})(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi))(\varphi)$$

by 3.9.4, we have from Theorem 3.8 that

$$\Gamma \vdash \mathsf{F}(\mathfrak{n})(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi))(\varphi).$$

Furthermore, by Lemma 2.22

$$\begin{aligned}
&\mathsf{VF}(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi)(\mathfrak{n})) \cap \mathsf{VB}(\mathsf{F}(\mathfrak{n})(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi))(\varphi)) \setminus \{\mathfrak{x}(\mathfrak{n})\} \\
\subseteq &\mathsf{VF}(\Phi(\mathfrak{n})) \cap \mathsf{VB}(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathsf{F}(\mathfrak{n})(\Phi)(\varphi))) \setminus \{\mathfrak{x}(\mathfrak{n})\} \\
\subseteq &(\mathsf{VF}(\Phi(\mathfrak{n})) \cap (\{\mathfrak{z}\} \cup (\mathsf{VB}(\mathsf{F}(\mathfrak{n})(\Phi)(\varphi)) \setminus \{\mathfrak{y}\}))) \setminus \{\mathfrak{x}(\mathfrak{n})\} \\
= &(\mathsf{VF}(\Phi(\mathfrak{n})) \cap \mathsf{VB}(\mathsf{F}(\mathfrak{n})(\Phi)(\varphi))) \setminus \{\mathfrak{x}(\mathfrak{n})\} \setminus \{\mathfrak{y}\},
\end{aligned}$$

since $\mathfrak{z} \not\preceq \Phi(\mathfrak{n})$, whereby

$$|\mathsf{VF}(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi)(\mathfrak{n})) \cap \mathsf{VB}(\mathsf{F}(\mathfrak{n})(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi))(\varphi)) \setminus \{\mathfrak{x}(\mathfrak{n})\}| \leq \mathfrak{m} - 1.$$

Additionally $\mathfrak{x}(\mathsf{k}) \notin \mathsf{VF}(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi)(\mathfrak{l}))$ for all $\mathsf{k}, \mathfrak{l} \in [0, \mathrm{lh}(\mathfrak{x})[$ by Lemma 2.22, whence

$$\Gamma \vdash \mathsf{F}(\mathfrak{n}+1)(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi))(\varphi)$$

by induction hypothesis. Moreover

$$\mathsf{F}(\mathfrak{n}+1)(\mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\Phi))(\varphi) = \mathrm{SB}(\mathfrak{y})(\mathfrak{z})(\mathsf{F}(\mathfrak{n}+1)(\Phi)(\varphi)).$$

Applying Theorem 3.8 again gives

$$\Gamma \vdash \mathsf{F}(\mathfrak{n}+1)(\Phi)(\varphi).$$

Thus

$$\Gamma \vdash \mathsf{F}(n)(\Psi)(\varphi) \Rightarrow \Gamma \vdash \mathsf{F}(n+1)(\Psi)(\varphi)$$

for all $\Psi \in W^*$ such that $\mathrm{lh}(\Psi) = \mathrm{lh}(\mathfrak{x})$, $\mathsf{T}(\Psi(k)) = \mathsf{T}(\mathfrak{x}(k))$ and $\mathfrak{x}(k) \notin \mathsf{VF}(\Psi(l))$ for all $k, l \in [0, \mathrm{lh}(\mathfrak{x})[$, as claimed. □

Furthermore, the relation $\vdash$ has the following recognisable properties.

3.10 Lemma. *In a theory of types with axioms 1-4 (see Definition 3.4) the following hold for all $\Gamma, \Delta \subseteq \overline{W}_o$ and $\varphi, \psi, \theta \in W_o$.*

1. *If $\Delta \subseteq \Gamma$ and $\Delta \vdash \varphi_o$ then $\Gamma \vdash \varphi_o$.*
2. *If $\Gamma \vdash (\varphi_o \to \psi_o)$ and $\Delta \vdash (\psi_o \to \theta_o)$, then $\Gamma \cup \Delta \vdash (\varphi_o \to \theta_o)$. (Transitivity)*
3. *If $\Gamma \vdash (\varphi_o \to \psi_o)$ and $\Delta \vdash (\theta_o \to \psi_o)$, then $\Gamma \cup \Delta \vdash ((\varphi_o \vee \theta_o) \to \psi_o)$. (Or-elimination)*

Proof. 1. Let $\mathfrak{P}$ be the proof of φ from Δ and $n \in \mathbb{Z}^+$ be its length. Then $\mathfrak{P}(n-1) = \varphi$ and for every $k \in [0, n[$ we have that $\mathfrak{P}(k) \in \Delta \subseteq \Gamma$, $\mathfrak{P}(k)$ is a formal axiom or $\mathfrak{P}(k)$ can be inferred from $\{\mathfrak{P}(l) \,|\, 0 \le l < k\}$ by some rule of inference; that is $\mathfrak{P}(k)$ is also a proof of φ from Γ. Hence $\Gamma \vdash \varphi$.

2. By above, $\Gamma \cup \Delta \vdash (\varphi \to \psi)$ and $\Gamma \cup \Delta \vdash (\psi \to \theta)$. Furthermore,

$$\vdash ((p_o \to q_o) \to ((r_o \to p_o) \to (r_o \to q_o)))$$

by Lemma 3.6, so

$$\Gamma \cup \Delta \vdash ((p_o \to q_o) \to ((r_o \to p_o) \to (r_o \to q_o))).$$

By Theorem 3.8,

$$\Gamma \cup \Delta \vdash ((\mathbf{p} \to \mathbf{q}) \to ((\mathbf{r} \to \mathbf{p}) \to (\mathbf{r} \to \mathbf{q})))$$

for distinct $\mathbf{p}, \mathbf{q}, \mathbf{r} \in X_o$ such that $\mathbf{p}, \mathbf{q}, \mathbf{r} \not\preceq p_o, q_o, r_o, \varphi, \psi, \theta$. Since

$$\begin{aligned} &((\psi \to \theta) \to ((\varphi \to \psi) \to (\varphi \to \theta))) \\ =&\mathrm{SF}(\mathbf{r})(\varphi)(\mathrm{SF}(\mathbf{q})(\theta)(\mathrm{SF}(\mathbf{p})(\psi)(((\mathbf{p} \to \mathbf{q}) \to ((\mathbf{r} \to \mathbf{p}) \to (\mathbf{r} \to \mathbf{q})))))) \end{aligned}$$

(or, in the language of Proposition 3.9,

$$\begin{aligned} &((\psi \to \theta) \to ((\varphi \to \psi) \to (\varphi \to \theta))) \\ =&\mathsf{F}(3)(\Phi)(((\mathbf{p} \to \mathbf{q}) \to ((\mathbf{r} \to \mathbf{p}) \to (\mathbf{r} \to \mathbf{q})))) \end{aligned}$$

for $\mathfrak{x} = \mathbf{pqr}$, where $\Phi(0) = \psi$, $\Phi(1) = \theta$ and $\Phi(2) = \varphi$) we have that

$$\Gamma \cup \Delta \vdash ((\psi \to \theta) \to ((\varphi \to \psi) \to (\varphi \to \theta)))$$

by Proposition 3.9, whence $\Gamma \cup \Delta \vdash ((\varphi \to \psi) \to (\varphi \to \theta))$ and

$$\Gamma \cup \Delta \vdash (\varphi \to \theta)$$

by Modus ponens, as desired.

3. Like before, $\Gamma \cup \Delta \vdash (\varphi \to \psi)$ and $\Gamma \cup \Delta \vdash (\theta \to \psi)$. Furthermore

$$\Gamma \cup \Delta \vdash ((p_o \to q_o) \to ((r_o \vee p_o) \to (r_o \vee q_o)))$$

by axiom 4. Let $\mathbf{p}, \mathbf{q}, \mathbf{r} \in X_o$ be distinct variables such that $\mathbf{p}, \mathbf{q}, \mathbf{r} \not\preceq p_o, q_o, r_o, \varphi, \psi, \theta$, whereby

$$\Gamma \cup \Delta \vdash ((\mathbf{p} \to \mathbf{q}) \to ((\mathbf{r} \vee \mathbf{p}) \to (\mathbf{r} \vee \mathbf{q})))$$

by Theorem 3.8, and thus

$$\Gamma \cup \Delta \vdash ((\theta \to \psi) \to ((\varphi \vee \theta) \to (\varphi \vee \psi))),$$
$$\Gamma \cup \Delta \vdash ((\varphi \to \psi) \to ((\psi \vee \varphi) \to (\psi \vee \psi)))$$

by Proposition 3.9. Hence

$$\Gamma \cup \Delta \vdash ((\varphi \vee \theta) \to (\varphi \vee \psi)),$$
$$\Gamma \cup \Delta \vdash ((\psi \vee \varphi) \to (\psi \vee \psi))$$

by Modus ponens. Likewise

$$\Gamma \cup \Delta \vdash ((p_o \vee q_o) \to (q_o \vee p_o))$$

by axiom 3, whence

$$\Gamma \cup \Delta \vdash ((\mathbf{p} \vee \mathbf{q}) \to (\mathbf{q} \vee \mathbf{p}))$$

by Theorem 3.8, and

$$\Gamma \cup \Delta \vdash ((\varphi \vee \psi) \to (\psi \vee \varphi))$$

by Proposition 3.9. Thus

$$\Gamma \cup \Delta \vdash ((\varphi \vee \theta) \to (\psi \vee \varphi))$$

and

$$\Gamma \cup \Delta \vdash ((\varphi \vee \theta) \to (\psi \vee \psi))$$

by Transitivity. Finally

$$\vdash ((p_o \vee p_o) \to p_o),$$

whereby

$$\vdash ((\mathbf{p} \vee \mathbf{p}) \to \mathbf{p})$$

by Theorem 3.8, and thus

$$\vdash ((\psi \vee \psi) \to \psi)$$

by Proposition 3.9, so that

$$\Gamma \cup \Delta \vdash ((\varphi \vee \theta) \to \psi)$$

by Transitivity, as desired.

□

The following lemma will illustrate the decrease in tediousness, augmented by the convention of not making any explicit appeals to part 2 of Theorem 3.8. We will here mainly be interested in the 10:th theorem, since it will be vital in the proof of the Deduction Theorem 3.13.

3.11 Lemma. *The following are formal theorems of a simple theory of types with axioms 1-4.*

1. $\vdash (p_o \rightarrow (\neg(\neg p_o)))$.
2. $\vdash ((\neg(\neg p_o)) \rightarrow p_o)$.
3. $\vdash ((p_o \rightarrow q_o) \rightarrow ((\neg q_o) \rightarrow (\neg p_o)))$.
4. $\vdash (((p_o \vee q_o) \vee r_o) \rightarrow (p_o \vee (q_o \vee r_o)))$.
5. $\vdash ((p_o \vee (q_o \vee r_o)) \rightarrow ((p_o \vee q_o) \vee r_o))$.
6. $\vdash ((p_o \rightarrow (q_o \rightarrow r_o)) \rightarrow (q_o \rightarrow (p_o \rightarrow r_o)))$.
7. $\vdash ((p_o \rightarrow q_o) \rightarrow ((q_o \rightarrow r_o) \rightarrow (p_o \rightarrow r_o)))$.
8. $\vdash (((p_o \wedge q_o) \rightarrow r_o) \rightarrow (p_o \rightarrow (q_o \rightarrow r_o)))$.
9. $\vdash ((p_o \rightarrow (q_o \rightarrow r_o)) \rightarrow ((p_o \wedge q_o) \rightarrow r_o))$.
10. $\vdash ((p_o \rightarrow (q_o \rightarrow r_o)) \rightarrow ((p_o \rightarrow q_o) \rightarrow (p_o \rightarrow r_o)))$.

Proof. We wish to remind our readers that the symbol = is the equality in the meta language. Below, unnumbered lines beginning with = means that the formula on that line is the same as the formula following the $\vdash$-sign on the previous line, courtesy of the abbreviations in 3.

1.

$\vdash((\neg p_o) \rightarrow (\neg p_o))$	3.6.2	(1)
$= ((\neg(\neg p_o)) \vee (\neg p_o))$		
$\vdash(((\neg(\neg p_o)) \vee (\neg p_o)) \rightarrow ((\neg p_o) \vee (\neg(\neg p_o))))$	axiom 3	(2)
$\vdash((\neg p_o) \vee (\neg(\neg p_o)))$	Modus ponens, 1, 2	(3)
$= (p_o \rightarrow (\neg(\neg p_o)))$		

2.

$\vdash(p_o \rightarrow p_o)$	3.6.2	(1)
$= ((\neg p_o) \vee p_o)$		
$\vdash(((\neg p_o) \vee p_o) \rightarrow (p_o \vee (\neg p_o)))$	axiom 3	(2)
$\vdash((\neg p_o) \rightarrow (\neg(\neg(\neg p_o))))$	3.11.1	(3)
$\vdash(((\neg p_o) \rightarrow (\neg(\neg(\neg p_o)))) \rightarrow ((p_o \vee (\neg p_o)) \rightarrow (p_o \vee (\neg(\neg(\neg p_o))))))$	axiom 4	(4)
$\vdash((p_o \vee (\neg p_o)) \rightarrow (p_o \vee (\neg(\neg(\neg p_o)))))$	Modus ponens, 3, 4	(5)
$\vdash(((\neg p_o) \vee p_o) \rightarrow (p_o \vee (\neg(\neg(\neg p_o)))))$	Transitivity, 2, 5	(6)
$\vdash((p_o \vee (\neg(\neg(\neg p_o)))) \rightarrow ((\neg(\neg(\neg p_o))) \vee p_o))$	axiom 3	(7)
$\vdash(((\neg p_o) \vee p_o) \rightarrow ((\neg(\neg(\neg p_o))) \vee p_o))$	Transitivity, 6, 7	(8)
$\vdash((\neg(\neg(\neg p_o))) \vee p_o)$	Modus ponens, 1, 8	(9)
$= ((\neg(\neg p_o)) \rightarrow p_o)$		

3.

$\vdash(q_o \to (\neg(\neg q_o)))$	3.11.1	(1)
$\vdash((q_o \to (\neg(\neg q_o))) \to (((\neg p_o) \lor q_o) \to ((\neg p_o) \lor (\neg(\neg q_o)))))$	axiom 4	(2)
$\vdash(((\neg p_o) \lor q_o) \to ((\neg p_o) \lor (\neg(\neg q_o))))$	Modus ponens, 1, 2	(3)
$\vdash(((\neg p_o) \lor (\neg(\neg q_o))) \to ((\neg(\neg q_o)) \lor (\neg p_o)))$	axiom 3	(4)
$\vdash(((\neg p_o) \lor q_o) \to ((\neg(\neg q_o)) \lor (\neg p_o)))$	Transitivity, 3, 4	(5)
$= ((p_o \to q_o) \to ((\neg q_o) \to (\neg p_o)))$		

4.

$\vdash(p_o \to (p_o \lor (q_o \lor r_o)))$	axiom 2	(1)
$\vdash(q_o \to (q_o \lor r_o))$	axiom 2	(2)
$\vdash((q_o \lor r_o) \to (p_o \lor (q_o \lor r_o)))$	3.6.3	(3)
$\vdash(q_o \to (p_o \lor (q_o \lor r_o)))$	Transitivity, 2, 3	(4)
$\vdash(r_o \to (q_o \lor r_o))$	3.6.3	(5)
$\vdash(r_o \to (p_o \lor (q_o \lor r_o)))$	Transitivity, 5, 3	(6)
$\vdash((p_o \lor q_o) \to (p_o \lor (q_o \lor r_o)))$	Or-elimination, 1, 4	(7)
$\vdash(((p_o \lor q_o) \lor r_o) \to (p_o \lor (q_o \lor r_o)))$	Or-elimination, 7, 6	(8)

5.

$\vdash(p_o \to (p_o \lor q_o))$	axiom 2	(1)
$\vdash((p_o \lor q_o) \to ((p_o \lor q_o) \lor r_o))$	axiom 2	(2)
$\vdash(p_o \to ((p_o \lor q_o) \lor r_o))$	Transitivity, 1, 2	(3)
$\vdash(q_o \to (p_o \lor q_o))$	3.6.3	(4)
$\vdash(q_o \to ((p_o \lor q_o) \lor r_o))$	Transitivity, 4, 2	(5)
$\vdash(r_o \to ((p_o \lor q_o) \lor r_o))$	3.6.3	(6)
$\vdash((q_o \lor r_o) \to ((p_o \lor q_o) \lor r_o))$	Or-elimination, 5, 6	(7)
$\vdash((p_o \lor (q_o \lor r_o)) \to ((p_o \lor q_o) \lor r_o))$	Or-elimination, 3, 7	(8)

6.

$\vdash(((\neg p_o) \lor (\neg q_o)) \to ((\neg q_o) \lor (\neg p_o)))$	axiom 3	(1)
$\vdash(((\neg q_o) \lor (\neg p_o)) \to (((\neg q_o) \lor (\neg p_o)) \lor r_o))$	axiom 2	(2)
$\vdash(((\neg p_o) \lor (\neg q_o)) \to (((\neg q_o) \lor (\neg p_o)) \lor r_o))$	Transitivity, 1, 2	(3)
$\vdash(r_o \to (((\neg q_o) \lor (\neg p_o)) \lor r_o))$	3.6.3	(4)
$\vdash((((\neg p_o) \lor (\neg q_o)) \lor r_o) \to (((\neg q_o) \lor (\neg p_o)) \lor r_o))$	Or-elimination, 3, 4	(5)
$\vdash(((\neg p_o) \lor ((\neg q_o) \lor r_o)) \to (((\neg p_o) \lor (\neg q_o)) \lor r_o))$	3.11.5	(6)
$= (((\neg p_o) \lor (q_o \to r_o)) \to (((\neg p_o) \lor (\neg q_o)) \lor r_o))$		
$= ((p_o \to (q_o \to r_o)) \to (((\neg p_o) \lor (\neg q_o)) \lor r_o))$		
$\vdash((p_o \to (q_o \to r_o)) \to (((\neg q_o) \lor (\neg p_o)) \lor r_o))$	Transitivity, 6, 5	(7)
$\vdash((((\neg q_o) \lor (\neg p_o)) \lor r_o) \to ((\neg q_o) \lor ((\neg p_o) \lor r_o)))$	3.11.4	(8)
$\vdash((p_o \to (q_o \to r_o)) \to ((\neg q_o) \lor ((\neg p_o) \lor r_o)))$	Transitivity, 7, 8	(9)
$= ((p_o \to (q_o \to r_o)) \to ((\neg q_o) \lor (p_o \to r_o)))$		
$= ((p_o \to (q_o \to r_o)) \to (q_o \to (p_o \to r_o)))$		

7.

$\vdash((q_o \to r_o) \to ((p_o \to q_o) \to (p_o \to r_o)))$	3.6.1	(1)
$\vdash(((q_o \to r_o) \to ((p_o \to q_o) \to (p_o \to r_o))) \to ((p_o \to q_o) \to ((q_o \to r_o) \to (p_o \to r_o))))$	3.11.6	(2)
$\vdash((p_o \to q_o) \to ((q_o \to r_o) \to (p_o \to r_o)))$	Transitivity, 1, 2	(3)

8.

$\vdash((\neg(\neg((\neg p_o)\vee(\neg q_o)))) \rightarrow ((\neg p_o)\vee(\neg q_o)))$	3.11.2	(1)
$\vdash(((\neg p_o)\vee(\neg q_o)) \rightarrow (((\neg p_o)\vee(\neg q_o))\vee r_o))$	axiom 2	(2)
$\vdash((\neg(\neg((\neg p_o)\vee(\neg q_o)))) \rightarrow (((\neg p_o)\vee(\neg q_o))\vee r_o))$	Transitivity, 1, 2	(3)
$\vdash(r_o \rightarrow (((\neg p_o)\vee(\neg q_o))\vee r_o))$	3.6.3	(4)
$\vdash(((\neg(\neg((\neg p_o)\vee(\neg q_o))))\vee r_o) \rightarrow (((\neg p_o)\vee(\neg q_o))\vee r_o))$	Or-elimination, 3, 4	(5)
$\vdash((((\neg p_o)\vee(\neg q_o))\vee r_o) \rightarrow ((\neg p_o)\vee((\neg q_o)\vee r_o)))$	3.11.4	(6)
$\vdash(((\neg(\neg((\neg p_o)\vee(\neg q_o))))\vee r_o) \rightarrow ((\neg p_o)\vee((\neg q_o)\vee r_o)))$	Transitivity, 5, 6	(7)
$= (((\neg((\neg p_o)\vee(\neg q_o))) \rightarrow r_o) \rightarrow ((\neg p_o)\vee(q_o \rightarrow r_o)))$		
$= (((p_o \wedge q_o) \rightarrow r_o) \rightarrow (p_o \rightarrow (q_o \rightarrow r_o)))$		

9.

$\vdash(((\neg p_o)\vee((\neg q_o)\vee r_o)) \rightarrow (((\neg p_o)\vee(\neg q_o))\vee r_o))$	3.11.5	(1)
$\vdash(((\neg p_o)\vee(\neg q_o)) \rightarrow (\neg(\neg((\neg p_o)\vee(\neg q_o)))))$	3.11.1	(2)
$\vdash((\neg(\neg((\neg p_o)\vee(\neg q_o)))) \rightarrow ((\neg(\neg((\neg p_o)\vee(\neg q_o))))\vee r_o))$	axiom 2	(3)
$\vdash(((\neg p_o)\vee(\neg q_o)) \rightarrow ((\neg(\neg((\neg p_o)\vee(\neg q_o))))\vee r_o))$	Transitivity, 2, 3	(4)
$\vdash(r_o \rightarrow ((\neg(\neg((\neg p_o)\vee(\neg q_o))))\vee r_o))$	3.6.3	(5)
$\vdash((((\neg p_o)\vee(\neg q_o))\vee r_o) \rightarrow ((\neg(\neg((\neg p_o)\vee(\neg q_o))))\vee r_o))$	Or-elimination, 2, 5	(6)
$\vdash(((\neg p_o)\vee((\neg q_o)\vee r_o)) \rightarrow ((\neg(\neg((\neg p_o)\vee(\neg q_o))))\vee r_o))$	Transitivity, 1, 6	(7)
$= (((\neg p_o)\vee(q_o \rightarrow r_o)) \rightarrow ((\neg((\neg p_o)\vee(\neg q_o))) \rightarrow r_o))$		
$= ((p_o \rightarrow (q_o \rightarrow r_o)) \rightarrow ((p_o \wedge q_o) \rightarrow r_o))$		

10. Due to the limited space in this proof, below MP will denote Modus ponens, T will denote Transitivity and $\vee$E will denote Or-elimination.

$\vdash(q_o \rightarrow (\neg(\neg q_o)))$	3.11.1	(1)
$\vdash((q_o \rightarrow (\neg(\neg q_o))) \rightarrow (((\neg(\neg q_o)) \rightarrow (\neg p_o)) \rightarrow (q_o \rightarrow (\neg p_o))))$	3.11.7	(2)
$\vdash(((\neg(\neg q_o)) \rightarrow (\neg p_o)) \rightarrow (q_o \rightarrow (\neg p_o)))$	MP, 1, 2	(3)
$\vdash((q_o \rightarrow (\neg p_o)) \rightarrow ((p_o \rightarrow q_o) \rightarrow (p_o \rightarrow (\neg p_o))))$	3.6.1	(4)
$\vdash(((\neg(\neg q_o)) \rightarrow (\neg p_o)) \rightarrow ((p_o \rightarrow q_o) \rightarrow (p_o \rightarrow (\neg p_o))))$	T, 3, 4	(5)
$\vdash((p_o \rightarrow (\neg q_o)) \rightarrow ((\neg(\neg q_o)) \rightarrow (\neg p_o)))$	3.11.3	(6)
$\vdash((p_o \rightarrow (\neg q_o)) \rightarrow ((p_o \rightarrow q_o) \rightarrow (p_o \rightarrow (\neg p_o))))$	T, 6, 5	(7)
$= ((\neg((\neg p_o)\vee(\neg q_o)))\vee((\neg(p_o \rightarrow q_o))\vee((\neg p_o)\vee(\neg p_o))))$		
$= ((p_o \wedge q_o)\vee((\neg(p_o \rightarrow q_o))\vee((\neg p_o)\vee(\neg p_o))))$		
$\vdash(((\neg p_o)\vee(\neg p_o)) \rightarrow (\neg p_o))$	axiom 1	(8)
$\vdash((\neg p_o) \rightarrow ((\neg(p_o \rightarrow q_o))\vee(\neg p_o)))$	3.6.3	(9)
$\vdash(((\neg p_o)\vee(\neg p_o)) \rightarrow ((\neg(p_o \rightarrow q_o))\vee(\neg p_o)))$	T, 8, 9	(10)
$\vdash((\neg(p_o \rightarrow q_o)) \rightarrow ((\neg(p_o \rightarrow q_o))\vee(\neg p_o)))$	axiom 3	(11)
$\vdash(((\neg(p_o \rightarrow q_o))\vee((\neg p_o)\vee(\neg p_o))) \rightarrow ((\neg(p_o \rightarrow q_o))\vee(\neg p_o)))$	$\vee$E, 11, 10	(12)
$\vdash(((\neg(p_o \rightarrow q_o))\vee(\neg p_o)) \rightarrow ((p_o \wedge q_o)\vee((\neg(p_o \rightarrow q_o))\vee(\neg p_o))))$	3.6.3	(13)
$\vdash(((\neg(p_o \rightarrow q_o))\vee((\neg p_o)\vee(\neg p_o))) \rightarrow ((p_o \wedge q_o)\vee((\neg(p_o \rightarrow q_o))\vee(\neg p_o))))$	T, 12, 13	(14)
$\vdash((p_o \wedge q_o) \rightarrow ((p_o \wedge q_o)\vee((\neg(p_o \rightarrow q_o))\vee(\neg p_o))))$	axiom 2	(15)
$\vdash(((p_o \wedge q_o)\vee((\neg(p_o \rightarrow q_o))\vee((\neg p_o)\vee(\neg p_o)))) \rightarrow ((p_o \wedge q_o)\vee((\neg(p_o \rightarrow q_o))\vee(\neg p_o))))$	$\vee$E, 15, 14	(16)
$\vdash((p_o \wedge q_o)\vee((\neg(p_o \rightarrow q_o))\vee(\neg p_o)))$	MP, 6, 16	(17)
$\vdash(((p_o \wedge q_o)\vee((\neg(p_o \rightarrow q_o))\vee(\neg p_o))) \rightarrow (((\neg(p_o \rightarrow q_o))\vee(\neg p_o))\vee(p_o \wedge q_o)))$	3.11.5	(18)
$\vdash(((\neg(p_o \rightarrow q_o))\vee(\neg p_o))\vee(p_o \wedge q_o))$	MP, 17, 18	(19)
$\vdash(((\neg(p_o \rightarrow q_o))\vee(\neg p_o)) \rightarrow (\neg(\neg((\neg(p_o \rightarrow q_o))\vee(\neg p_o)))))$	3.11.1	(20)
$\vdash((\neg(\neg((\neg(p_o \rightarrow q_o))\vee(\neg p_o)))) \rightarrow ((\neg(\neg((\neg(p_o \rightarrow q_o))\vee(\neg p_o))))\vee(p_o \wedge q_o)))$	axiom 3	(21)
$\vdash(((\neg(p_o \neg q_o))\vee(\neg p_o)) \rightarrow ((\neg(\neg((\neg(p_o \rightarrow q_o))\vee(\neg p_o))))\vee(p_o \wedge q_o)))$	T, 20, 21	(22)
$\vdash((p_o \wedge q_o) \rightarrow ((\neg(\neg((\neg(p_o \rightarrow q_o))\vee(\neg p_o))))\vee(p_o \wedge q_o)))$	3.6.3	(23)
$\vdash((((\neg(p_o \rightarrow q_o))\vee(\neg p_o))\vee(p_o \wedge q_o)) \rightarrow ((\neg(\neg((\neg(p_o \rightarrow q_o))\vee(\neg p_o))))\vee(p_o \wedge q_o)))$	3.11.1	(24)
$\vdash((\neg(\neg((\neg(p_o \rightarrow q_o))\vee(\neg p_o))))\vee(p_o \wedge q_o))$	MP, 19, 24	(25)
$= ((\neg((p_o \rightarrow q_o)\wedge p_o))\vee(p_o \wedge q_o))$		
$= (((p_o \rightarrow q_o)\wedge p_o) \rightarrow (p_o \wedge q_o))$		

$\vdash((((p_o \to q_o) \wedge p_o) \to (p_o \wedge q_o)) \to (((p_o \wedge q_o) \to r_o) \to (((p_o \to q_o) \wedge p_o) \to r_o)))$	3.11.7	(26)
$\vdash(((p_o \wedge q_o) \to r_o) \to (((p_o \to q_o) \wedge p_o) \to r_o))$	MP, 25, 26	(27)
$\vdash((p_o \to (q_o \to r_o)) \to ((p_o \wedge q_o) \to r_o))$	3.11.9	(28)
$\vdash((p_o \to (q_o \to r_o)) \to (((p_o \to q_o) \wedge p_o) \to r_o))$	T, 28, 27	(29)
$\vdash((((p_o \to q_o) \wedge p_o) \to r_o) \to ((p_o \to q_o) \to (p_o \to r_o)))$	3.11.8	(30)
$\vdash((p_o \to (q_o \to r_o)) \to ((p_o \to q_o) \to (p_o \to r_o)))$	T, 29, 30	(31)

□

We finally reach the Deduction Theorem, in some sense the goal of this section. Note that, while the proof is not too short, the following converse is immediate, by Lemma 3.10.

3.12 Proposition. *In a simple theory of types with axioms 1-4, we have for every* $\Gamma \subseteq \overline{W}_o$, $\varphi \in \overline{W}_o$, $\psi \in W_o$ *that* $\Gamma \vdash (\varphi \to \psi) \Rightarrow \Gamma \cup \{\varphi\} \vdash \psi$.

Proof. If $\Gamma \vdash (\varphi \to \psi)$, then

$$\Gamma \cup \{\varphi\} \vdash (\varphi \to \psi).$$

Since $\Gamma \cup \{\varphi\} \vdash \varphi$, we see that

$$\Gamma \cup \{\varphi\} \vdash \psi$$

by Modus ponens. □

3.13 Theorem (The Deduction Theorem)**.** *In a simple theory of types with axioms 1-6, we have for every* $\Gamma \subseteq \overline{W}_o$, $\varphi \in \overline{W}_o$, $\psi \in W_o$ *that*

$$\Gamma \cup \{\varphi\} \vdash \psi \Rightarrow \Gamma \vdash (\varphi \to \psi).$$

Proof. Let $\mathfrak{P}$ be a proof of length $n \in \mathbb{Z}^+$ of ψ from $\Gamma \cup \{\varphi\}$. We will prove that $\Gamma \vdash (\varphi \to \mathfrak{P}(k))$ for all $k \in [0, n[$, by complete induction.

Let $m \leq n$ and assume that

$$\Gamma \vdash (\varphi \to \mathfrak{P}(k))$$

for all $0 \leq k < m$. Then $\mathfrak{P}(m) \in \Gamma \cup \{\varphi\}$, $\mathfrak{P}(m)$ is a formal axiom or $\mathfrak{P}(m)$ is obtained from $\{\mathfrak{P}(k) \mid 0 \leq k < m\}$ by some rule of inference.

If $\mathfrak{P}(m) = \varphi$ we have, since $\Gamma \vdash (p_o \to p_o)$ by Lemma 3.6, that

$$\Gamma \vdash (\varphi \to \mathfrak{P}(m))$$

by Theorem 3.8.

If instead $\mathfrak{P}(m)$ is in Γ or a formal axiom we have $\Gamma \vdash \mathfrak{P}(m)$; moreover, since $\Gamma \vdash (p_o \to (q_o \to p_o))$ by Lemma 3.6,

$$\Gamma \vdash (\mathfrak{P}(m) \to (p_o \to \mathfrak{P}(m)))$$

and

$$\Gamma \vdash (\mathfrak{P}(\mathfrak{m}) \to (\varphi \to \mathfrak{P}(\mathfrak{m})))$$

(since $\mathsf{VF}(\varphi) = \varnothing$) by Theorem 3.8, whereby

$$\Gamma \vdash (\varphi \to \mathfrak{P}(\mathfrak{m}))$$

by Modus ponens.

If $\mathfrak{P}(\mathfrak{m})$ was obtained from $\{\mathfrak{P}(\mathsf{k}) \,|\, 0 \leq \mathsf{k} < \mathfrak{m}\}$ by α-conversion, then there are $0 \leq \mathfrak{l} < \mathfrak{m}$, $\alpha \in \mathcal{T}$, $\mathbf{x}, \mathbf{y} \in \mathsf{X}_\alpha$, $\theta \in \mathcal{W}$ and $\mathfrak{a}, \mathfrak{b} \in \mathcal{S}^*$ such that $\mathbf{x}$ is not free in θ, $\mathbf{y}$ does not occur in θ,

$$\mathfrak{P}(\mathfrak{l}) = \mathfrak{a}\theta\mathfrak{b}$$

and

$$\mathfrak{P}(\mathfrak{m}) = \mathfrak{a}\,\mathrm{S}(\mathbf{x})(\mathbf{y})(\theta)\mathfrak{b}.$$

Hence by the induction hypothesis $\Gamma \vdash (\varphi \to \mathfrak{a}\theta\mathfrak{b})$, which since $(\varphi \to \mathfrak{a}\theta\mathfrak{b}) = [[\mathsf{A}[\mathsf{N}\varphi]]\mathfrak{a}\theta\mathfrak{b}]$ and $(\varphi \to \mathfrak{P}(\mathfrak{m})) = [[\mathsf{A}[\mathsf{N}\varphi]]\mathfrak{a}\,\mathrm{S}(\mathbf{x})(\mathbf{y})(\theta)\mathfrak{b}]$ implies that

$$\Gamma \vdash (\varphi \to \mathfrak{P}(\mathfrak{m})),$$

by α-conversion.

If $\mathfrak{P}(\mathfrak{m})$ was obtained from $\{\mathfrak{P}(\mathsf{k}) \,|\, 0 \leq \mathsf{k} < \mathfrak{m}\}$ by β-contraction, then there are $0 \leq \mathfrak{l} < \mathfrak{m}$, $\alpha, \beta \in \mathcal{T}$, $\mathbf{x} \in \mathsf{X}_\alpha$, $\theta_\alpha, \vartheta_\beta \in \mathcal{W}$ and $\mathfrak{a}, \mathfrak{b} \in \mathcal{S}^*$ such that

$$\mathfrak{P}(\mathfrak{l}) = \mathfrak{a}[[\lambda\mathbf{x}\vartheta]\theta]\mathfrak{b},$$
$$\mathfrak{P}(\mathfrak{m}) = \mathfrak{a}\,\mathrm{S}(\mathbf{x})(\theta)(\vartheta)\mathfrak{b}$$

and neither $\mathbf{x}$ nor any free variable of θ is bound in ϑ. Hence by the induction hypothesis $\Gamma \vdash (\varphi \to \mathfrak{a}[[\lambda\mathbf{x}\vartheta]\theta]\mathfrak{b})$, whereby

$$\Gamma \vdash (\varphi \to \mathfrak{P}(\mathfrak{m}))$$

by β-contraction, similarly to above.

If $\mathfrak{P}(\mathfrak{m})$ was obtained from $\{\mathfrak{P}(\mathsf{k}) \,|\, 0 \leq \mathsf{k} < \mathfrak{m}\}$ by β-expansion, then there are $0 \leq \mathfrak{l} < \mathfrak{m}$, $\alpha, \beta \in \mathcal{T}$, $\mathbf{x} \in \mathsf{X}_\alpha$, $\theta_\alpha, \vartheta_\beta \in \mathcal{W}$ and $\mathfrak{a}, \mathfrak{b} \in \mathcal{S}^*$ such that

$$\mathfrak{P}(\mathfrak{l}) = \mathfrak{a}\,\mathrm{S}(\mathbf{x})(\theta)(\vartheta)\mathfrak{b},$$
$$\mathfrak{P}(\mathfrak{m}) = \mathfrak{a}[[\lambda\mathbf{x}\vartheta]\theta]\mathfrak{b}$$

and neither $\mathbf{x}$ nor any free variable of θ is bound in ϑ. Hence by the induction hypothesis $\Gamma \vdash (\varphi \to \mathfrak{a}\,\mathrm{S}(\mathbf{x})(\theta)(\vartheta)\mathfrak{b})$, whereby

$$\Gamma \vdash (\varphi \to \mathfrak{P}(\mathfrak{m}))$$

by β-expansion.

If $\mathfrak{P}(\mathfrak{m})$ was obtained from $\{\mathfrak{P}(\mathfrak{k})\,|\,0 \le \mathfrak{k} < \mathfrak{m}\}$ by Substitution there are $0 \le \mathfrak{l} < \mathfrak{m}$, $\alpha \in \mathcal{T}$, $\theta_\alpha, \vartheta_{o\alpha} \in W$ and $\boldsymbol{x} \in X_\alpha$ such that $\mathfrak{P}(\mathfrak{l}) = [\vartheta\boldsymbol{x}]$, $\mathfrak{P}(\mathfrak{m}) = [\vartheta\theta]$ and $\boldsymbol{x}$ is not free in ϑ. By the induction hypothesis $\Gamma \vdash (\varphi \to [\vartheta\boldsymbol{x}])$. Let $\mathbf{y}_\alpha, \mathbf{p}_o, \mathbf{f}_{o\alpha} \in \mathcal{X}$ be such that $\mathbf{y}, \mathbf{p}, \mathbf{f} \not\preceq \boldsymbol{x}, \varphi, \vartheta, \theta$. Then

$$\Gamma \vdash (\varphi \to [\vartheta\mathbf{y}])$$

by Theorem 3.8, where $(\varphi \to [\vartheta\mathbf{y}]) = \mathrm{SF}(\mathbf{f})(\vartheta)(\mathrm{SF}(\mathbf{p})(\varphi)((\mathbf{p} \to [\mathbf{f}\mathbf{y}])))$. Since

$$(\varphi \to [\vartheta\theta]) = \mathrm{SF}(\mathbf{y})(\theta)(\mathrm{SF}(\mathbf{f})(\vartheta)(\mathrm{SF}(\mathbf{p})(\varphi)((\mathbf{p} \to [\mathbf{f}\mathbf{y}]))))$$

we conclude by Proposition 3.9 that $\Gamma \vdash (\varphi \to [\vartheta\theta])$, i.e.

$$\Gamma \vdash (\varphi \to \mathfrak{P}(\mathfrak{m}))$$

as desired.

If $\mathfrak{P}(\mathfrak{m})$ was obtained from $\{\mathfrak{P}(\mathfrak{k})\,|\,0 \le \mathfrak{k} < \mathfrak{m}\}$ by Modus ponens there are $0 \le \mathfrak{l}, \mathfrak{j} < \mathfrak{m}$ such that

$$\mathfrak{P}(\mathfrak{l}) = (\mathfrak{P}(\mathfrak{j}) \to \mathfrak{P}(\mathfrak{m})).$$

By the induction hypothesis $\Gamma \vdash (\varphi \to \mathfrak{P}(\mathfrak{l}))$ and $\Gamma \vdash (\varphi \to \mathfrak{P}(\mathfrak{j}))$. Furthermore

$$\Gamma \vdash ((p_o \to (q_o \to r_o)) \to ((p_o \to q_o) \to (p_o \to r_o)))$$

by Lemmas 3.11 and 3.10, whereby

$$\Gamma \vdash ((\varphi \to (\mathfrak{P}(\mathfrak{j}) \to \mathfrak{P}(\mathfrak{m}))) \to ((\varphi \to \mathfrak{P}(\mathfrak{j})) \to (\varphi \to \mathfrak{P}(\mathfrak{m}))))$$

by Proposition 3.9. Thus

$$\Gamma \vdash ((\varphi \to \mathfrak{P}(\mathfrak{j})) \to (\varphi \to \mathfrak{P}(\mathfrak{m})))$$

and

$$\Gamma \vdash (\varphi \to \mathfrak{P}(\mathfrak{m}))$$

by Modus ponens.

If $\mathfrak{P}(\mathfrak{m})$ was obtained from $\{\mathfrak{P}(\mathfrak{k})\,|\,0 \le \mathfrak{k} < \mathfrak{m}\}$ by Generalisation there are $0 \le \mathfrak{l} < \mathfrak{m}$, $\alpha \in \mathcal{T}$, $\theta_{o\alpha} \in W$ and $\boldsymbol{x} \in X_\alpha$ such that $\boldsymbol{x}$ is not free in θ,

$$\mathfrak{P}(\mathfrak{l}) = [\theta\boldsymbol{x}]$$

and

$$\mathfrak{P}(\mathfrak{m}) = [\Pi_{o\langle o\alpha\rangle}\theta].$$

By the induction hypothesis $\Gamma \vdash (\varphi \to \mathfrak{P}(l))$. Furthermore

$$\Gamma \vdash ((\forall x_\alpha(p_o \vee [f_{o\alpha} x_\alpha])) \to (p_o \vee [\Pi_{o\langle o\alpha\rangle} f_{o\alpha}]))$$

by axiom 6. Let $\mathfrak{y} \in X_\alpha$ be a variable distinct from x_α and $\mathfrak{x}$ which does not occur in θ or φ. Then

$$\Gamma \vdash (\varphi \to [\theta \mathfrak{y}]),$$

$$\Gamma \vdash (\forall \mathfrak{y}(\varphi \to [\theta \mathfrak{y}]))$$

and

$$\Gamma \vdash ((\forall \mathfrak{y}(p_o \vee [f_{o\alpha} \mathfrak{y}])) \to (p_o \vee [\Pi_{o\langle o\alpha\rangle} f_{o\alpha}]))$$

by Theorem 3.8. Since $\mathfrak{y} \notin VF(\theta)$ and $VF((\neg\varphi)) = VF(\varphi) = \varnothing$,

$$\Gamma \vdash ((\forall \mathfrak{y}(p_o \vee [\theta \mathfrak{y}])) \to (p_o \vee [\Pi_{o\langle o\alpha\rangle} \theta]))$$

and

$$\Gamma \vdash ((\forall \mathfrak{y}(\varphi \to [\theta \mathfrak{y}])) \to (\varphi \to [\Pi_{o\langle o\alpha\rangle} \theta]))$$

by the same theorem. Thus

$$\Gamma \vdash (\varphi \to \mathfrak{P}(m))$$

by Modus ponens.

In particular, since $\mathfrak{P}(n-1) = \psi$, $\Gamma \vdash (\varphi \to \psi)$. □

Church proved the above theorem with his definition of a formal proof, which allows one to derive e.g. that $\vdash p_o \to p_o$ by showing that $p_o \vdash p_o$. This is of course impossible in the current context since $p_o \notin \overline{W}$, i.e. $p_o \vdash p_o$ does not make sense. However, the corresponding statement holds for any closed well-formed formula φ_o.

3.14 Corollary. *Let $\varphi \in \overline{W}_o$ and $\psi \in W_o$. Then*

$$\varphi \vdash \psi \Leftrightarrow \vdash (\varphi \to \psi).$$

Moreover, this version of the theorem is actually enough to be applicable in the proof of the compactness theorem, as can be seen in [3].

It seems reasonable to assume that Church's more general theorem could in some sense be salvaged by temporarily expanding the language with a new constant symbol, appealing to the above theorem and replacing all occurrences of that symbol in the obtained proof by e.g. p_o like above. However, an investigation of this idea is beyond the scope of this paper.

References

[1] Peter B. Andrews. *An Introduction to Mathematical Logic and Type Theory: To Truth Through Proof.* Number 27 in Applied Logic Series. Kluwer Academic Publishers, Dordrecht, second edition, 2002.

[2] Alonzo Church. A formulation of the simple theory of types. *Journal of Symbolic Logic*, 5(2):56–68, June 1940.

[3] Leon Henkin. Completeness in the theory of types. *Journal of Symbolic Logic*, 15(2):81–91, June 1950.

[4] Yannis Moschovakis. *Notes on Set Theory.* Undergraduate Texts in Mathematics. Springer, second edition, 2006.