# Why Ramify?

## Harold T. Hodes

**Abstract**    This paper considers two reasons that might support Russell's (and Whitehead's) choice of a ramified-type theory over a simple-type theory. The first reason is the existence of purported paradoxes that can be formulated in any simple-type language, including an argument that Russell considered in 1903. These arguments depend on certain converse-compositional principles. When we take account of Russell's doctrine that a propositional function is not a constituent of its values, these principles turn out to be too implausible to make these arguments troubling. The second reason is conditional on a substitutional interpretation of quantification over types other than that of individuals. This reason stands up to investigation: a simple-type language will not sustain such an interpretation, but a ramified-type language will. And there is evidence that Russell was tacitly inclined towards such an interpretation. A strong construal of that interpretation opens a way to make sense of Russell's simultaneous repudiation of propositions and his willingness to quantify over them. But that way runs into trouble with Russell's commitment to the finitude of human understanding.

The Whitehead–Russell project in logic and the foundations of mathematics is something of a historical anomaly. Usually, scientists and mathematicians try out simple ideas before complex ideas, moving on to consider complex ideas only after simpler ideas turn out to be inadequate. But when Whitehead and Russell developed their (philosophical) theory of types, and their logical system that articulated that theory, roughly from 1907 to 1910, they went straight to ramified types. The earliest published consideration of simple types was in work by Leon Chwistek and Frank Ramsey in the 1920s.[1]

In this paper, I will consider simple-type assignment systems and ask whether, given Russell's background views, he would have had any good reasons to be unhappy with simple-type theory as captured by such systems. I will consider two possible such reasons.

© 2015 by University of Notre Dame      10.1215/00294527-2864352

Discussions of reasons for ramification have focused on the avoidance of semantic paradoxes.[2] For example, Quine wrote: "[Russell] thought that these distinctions [of order] were helpful against a class of paradoxes...: the paradoxes known nowadays as semantic" (see Quine [39, p. 254]).[3] This understanding of Russell's motives goes back to Frank Ramsey, who first distinguished between the mathematical and semantic paradoxes in [40, p. 20]. He characterized semantic paradoxes, those constituting his group B, as those "containing some reference to thought, language or symbolism, which are not formal but empirical terms"—presumably terms like "true," "denotes," and so forth.

Russell clearly thought that ramified types helped avoid the liar and Richard's paradoxes, which are uncontroversially semantic, both in the current sense (of involving "semantic ascent") and in Ramsey's sense: the liar involved reference to a speaker and the making of a statement; Richard's involved reference to syllables and naming. I will not dispute what seems to be the current consensus: that semantic paradoxes originate from semantic ascent, rather than from a lack of ramification of whatever types organize the logical syntax of our language. Since those paradoxes do not afflict simple-type logic as such (i.e., without machinery to express any semantic concepts), their avoidance is not a good reason for ramification.

But Russell also found paradoxical an argument naturally formulated in any simple-type language, one that is not, at least on its face, semantic. And one might think that, given certain Russellian commitments about the structure and constitution of propositions, it contaminates simple-type logic. Below, I will discuss this purported paradox and two simplifications of it. In the end, I will argue that they are not paradoxical: they rely on converse-compositional principles that, on examination, are implausible in their own terms. So these arguments do not give us a good reason for ramification.

I will go on to discuss a reason to ramify that has nothing to do with paradoxes: the substitutional interpretation of quantification over types other than that of individuals. If one favors such an interpretation, one has a good reason for ramification. I will consider a way of using a strong form of that interpretation to make sense of Russell's puzzling post-1908 doctrine that propositions are "incomplete symbols." Unfortunately, it seems that Russell could not have endorsed this connection, since it would require the comprehensibility of infinitary propositions.

My purpose here is not to defend or criticize any philosophical view but rather to better understand Russell's philosophy and his century-old project in the foundations of logic and mathematics.

## 1  Simple Types and Terms

What follows will make use of standard mathematical machinery. For $n \in \omega$, let $(n) = \{1, \ldots, n\}$.

I will inductively define the simple types (sometimes to be called "s-types" or, if confusion is unlikely, just "types"). Each type will be assigned a place number.

Think of s-types as linguistic expressions and thus as able to occur in terms.[4]

**Definition**    **i** is an s-type (for individuals); $\langle \rangle$ is an s-type (for propositions). We have place(**i**) = place($\langle \rangle$) = 0. For $n \in \omega - \{0\}$, if $t_1, \ldots, t_n$ are s-types, then $\langle t_1, \ldots, t_n \rangle$ is an s-type (for $n$-place propositional functions taking arguments of s-type $t_i$ at the $i$th place), and place($\langle t_1, \ldots, t_n \rangle$) = $n$. Let $\mathcal{T}^s$ be the set of s-types.

On one standard approach to variables in typed languages, going back to Alonzo Church, each variable comes associated with a unique type.[5] Another approach, going back to Haskell Curry, uses "in situ" typing for bound occurrences of variables, and "external" typing by type contexts for free occurrences of variables.[6] I will take the latter approach (which takes a step toward incorporating Russell and Whitehead's preference for "ambiguity" of type.[7]) I will be careless about the use-mention distinction and will use Greek letters as variables added to English to range over terms in the object-language, sometimes mixing upper with lower case to increase readability.

**Definitions**     A type assignment for an expression $\gamma$ is an expression of the form $\gamma : t$ for $t \in \mathcal{T}$. Think of the type assignment $\gamma : t$ as the ordered pair $\langle \gamma, t \rangle$; so the domain and range of a set of them can be understood in the usual way.

Fix an infinite set Var of variables. Slightly modifying the notion of a type context in Hindley [19, p. 14], let $C$ be a simple-type (until further notice I will omit "simple") context if and only if $C$ is a single-valued set of type assignments for variables; that is, $\text{dom}(C) \subseteq \text{Var}$ and for any $v : t, v : t' \in C$ then $t = t'$, such that $\text{Var} - \text{dom}(C)$ is infinite.[8] Let

$$C, v_1 : t_1, \ldots, v_n : t_n = C \cup \{v_1 : t_1, \ldots, v_n : t_n\},$$

provided that it is a context and $v_1 : t_1, \ldots, v_n : t_n \notin C$. In specifications of type contexts, I will sometimes omit the curly brackets of set formation.

To model the roles played by interpreted (i.e., meaning-bearing) expressions, consider a fresh set of expressions to serve as nonlogical constants.

**Definitions**     $\mathcal{S}$ is a simply typed vocabulary set or signature (where confusion is unlikely I will omit "simply typed") if and only if it is a single-valued set of type assignments for nonlogical constants. If $\gamma : \mathbf{i} \in \mathcal{S}$, then $\gamma$ is an individual constant in $\mathcal{S}$; if $\gamma : t \in \mathcal{S}$ for $t \neq \mathbf{i}$, then $\text{place}(t) > 0$ and $\gamma$ is a $\text{place}(t)$-place predicate constant in $\mathcal{S}$.

In effect, Whitehead and Russell [55] and Russell [41] take as terms to constitute their broadest syntactic category; formulas are special terms. I will do the same (though this is inessential; see Hodes [20]).

**Definitions**     Given a vocabulary set $\mathcal{S}$, what follows is the simultaneous inductive definition of the s-type assignment system $\Rightarrow^s_{\mathcal{S}}$ and the free-variable assignment $FV$ for that system. (Where confusion is unlikely, I will omit mention of $\mathcal{S}$ and the superscript for "simple.")

(1) If $v : t \in C$, then $C \Rightarrow v : t$ and $FV(v) = \{v\}$.

(2) If $\tau : t \in \mathcal{S}$, then $C \Rightarrow \tau : t$ and $FV(\tau) = \{\}$.

(3) If $C \Rightarrow \tau : \langle t_1, \ldots, t_n \rangle$ with $n > 0$ and $C \Rightarrow \tau_i : t_i$ for each $i \in (n)$, then $C \Rightarrow \tau(\tau_1, \ldots, \tau_n) : \langle\rangle$ and

$$FV\big(\tau(\tau_1, \ldots, \tau_n)\big) = FV(\tau) \cup \bigcup_{i \in (n)} FV(\tau_i).$$

(4.1) We have $C \Rightarrow \bot : \langle\rangle$ and $FV(\bot) = \{\}$.

(4.2) If $C \Rightarrow \tau_i : \langle\rangle$ for $i = 0, 1$, then $C \Rightarrow (\tau_0 \supset \tau_1) : \langle\rangle$, and

$$FV\big((\tau_0 \supset \tau_1)\big) = FV(\tau_0) \cup FV(\tau_1).$$

Define &, $\vee$, and $\leftrightarrow$ as usual, and let $\neg\tau$ abbreviate $(\tau \supset \bot)$.[9]

(5) If $C, \nu : t \Rightarrow \tau : \langle\rangle$, then $C \Rightarrow \exists \nu : t.\tau : \langle\rangle$, and

$$FV(\exists \nu : t.\tau) = FV(\tau) - \{\nu\}.$$

Define $\forall$ as usual.[10]

(6) If $n > 0$, $C, \nu_1 : t_1, \ldots, \nu_n : t_n \Rightarrow \tau : \langle\rangle$ for distinct $\nu_1, \ldots, \nu_n \in FV(\tau)$,
then $C \Rightarrow (\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\tau) : \langle t_1, \ldots, t_n \rangle$ and

$$FV\big((\lambda \nu_1 : t_1, \ldots \nu_n : t_n.\tau)\big) = FV(\tau) - \{\nu_{i \in (n)}\}.$$

When $\{\nu_1 : t_1, \ldots, \nu_n : t_n.\tau\}$ occurs to the left of $\Rightarrow$, I will omit the curly brackets; in particular, $\Rightarrow \tau : t$ iff $\{\} \Rightarrow \tau : t$.

The relation $\Rightarrow$ determines its corresponding simply typed language thus:

$$L^s(\mathcal{S}) = \big\{\langle C, \tau \rangle : \text{for some } t, C \Rightarrow_{\mathcal{S}} \tau : t\big\}.$$

**Definitions**     $\tau$ is a term of $L^s$ relative to $C$ iff $\langle C, \tau \rangle \in L^s$; $\tau$ is a formula of $L^s$ relative to $C$ iff $C \Rightarrow_{\mathcal{S}} \tau : \langle\rangle$; $\tau$ is a term [formula] of $L^s$ iff it is a term [formula] in $L^s$ relative to some type context $C$. A term is closed if and only if it is a term in $L^s$ relative to $\{\}$. A sentence is a closed formula.

Following the convention used in Barendregt [2], we will treat terms as identical if and only if they are "alphabet variants" (i.e., "identical modulo relettering of bound variables"), following [2] and [19],[11] I will follow contemporary practice and use $\equiv$ to express identity of terms.

Here are some easy observations.

As usual for formation rules, all of these conditionals reverse.

If $\tau$ is not a variable and is a term relative to both $C_0$ and $C_1$, then it is a formula relative to $C_0$ iff it is one relative to $C_1$.

If $C \Rightarrow \tau : t$ and $C \Rightarrow \tau : t'$, then $t = t'$.

If $C \Rightarrow \tau(\tau_1, \ldots, \tau_n) : t$ or $C \Rightarrow (\tau_0 \supset \tau_1) : t$ or $C \Rightarrow \exists \nu : t'.\tau : t$, then $t = \langle\rangle$. Similarly for any other logical constants we may take as primitive. Also, if $C \Rightarrow (\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\sigma) : t$, then $t = \langle t_1, \ldots, t_n \rangle$ and $n > 0$.

**Definition**     For each $t \in \mathcal{T}$:

$$(\tau_0 =^t \tau_1) \equiv_{\text{def}} \forall \nu : \langle t \rangle \big(\nu(\tau_0) \leftrightarrow \nu(\tau_1)\big),$$

for $\nu$ a variable not in $FV(\tau_0) \cup FV(\tau_1)$.

**Definition**     For distinct variables $\nu_1, \ldots, \nu_n$,

$$\rho[\nu_1, \ldots, \nu_n := \tau_1, \ldots, \tau_n]$$

is the result of simultaneously substituting each $\tau_i$ for the occurrences of $\nu_i$ free in $\rho$ for $i \in (n)$, with the understanding that bound variables in $\rho$ are "relettered" so as to prevent free occurrences of variables in $\tau_i$ from being "captured" by the substitution. The definition is by the usual sort of induction on the construction of $\rho$.

**Definition**     A $\lambda I$-term is one of the form $(\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\tau)$.[12] Following [2] and [19], $I$ here reminds us that $\{\nu_{i \in (n)}\} \subseteq FV(\tau)$ and $n > 0$.[13] Note that these $\lambda I$-terms differ from the so-called terms in the standard texts on $\lambda$-calculi in being "multivariate" (allowing binding of more than one variable; for more on multivariate terms, see Pottinger [36]).

The following appropriates standard concepts and results from $\lambda$-calculi.

**Definitions**    A term of the form $\tau(\tau_1, \ldots, \tau_n)$ is a predication. It is an atomic predication if and only if $\tau$ is either a variable or a predicate-constant; otherwise, it is a $\beta$-redex. Let a $\beta$-redex $(\lambda v_1 : t_1, \ldots, v_n : t_n.\sigma)(\tau_1, \ldots, \tau_n)$ $\beta$-convert to $\sigma[v_1, \ldots, v_n := \tau_1, \ldots, \tau_n]$. We will call the latter the $\beta$-convert of the given $\beta$-redex. Note that if a $\beta$-redex is a term relative to $C$, both it and its $\beta$-convert are formulas relative to $C$, and they have the same free variables.

$\tau$ 1-step $\beta$-reduces to $\tau'$ (in symbols $\tau \to_\beta \tau'$) iff $\tau'$ results from replacing exactly one occurrence of a $\beta$-redex in $\tau$ by its $\beta$-convert.[14]

An $\eta$-redex is a term of the form

$$\left(\lambda v_1 : t_1, \ldots, v_n : t_n.\tau(v_1, \ldots, v_n)\right)$$

with distinct $v_1, \ldots, v_n \in \text{Var} - FV(\tau)$ and $n > 0$. Such an $\eta$-redex $\eta$-converts to $\tau$.

A term $\tau$ 1-step $\eta$-reduces to $\tau'$ (in symbols, $\tau \to_\eta \tau'$) iff $\tau'$ results from replacing a single occurrence of an $\eta$-redex in $\tau$ by its $\eta$-convert.[15]

A term $\tau$ 1-step $\beta\eta$-reduces to $\tau'$ (in symbols, $\tau \to_{\beta\eta} \tau'$) iff either $\tau \to_\beta \tau'$ or $\tau \to_\eta \tau'$.

$\tau$ $\beta\eta$-reduces to $\tau'$ (hereafter $\tau \twoheadrightarrow_{\beta\eta} \tau'$) iff there is a finite $\to_{\beta\eta}$-chain from $\tau$ to $\tau'$.

$\tau$ is $\beta\eta$-normal iff $\tau$ contains no $\beta$- or $\eta$-redex, that is, $\tau$ does not 1-step $\beta\eta$-reduce to anything. Let $\tau'$ be a $\beta\eta$-normalization of $\tau$ iff $\tau \twoheadrightarrow_{\beta\eta} \tau'$ and $\tau'$ is $\beta\eta$-normal.

**Observation**    We have that $\twoheadrightarrow_{\beta\eta}$, and thus $\beta\eta$-reduction, preserve termhood, type, and free variables relative to a type context.[16]

**Definition**    $\tau$ is strongly $\beta\eta$-normalizable (hereafter SN) iff every $\to_{\beta\eta}$-chain is finite.

**Strong normalization theorem**    *Every term is SN. Tait's well-known proof suffices.*[17]

**Church–Rosser theorem**    *Every term has a unique $\beta\eta$-normalization.*[18]

Though we will not here consider any proof-theoretic system, the reader should be aware of this: since termhood, and thus formulahood, is relative to type context that types free variables, being a deduction in any reasonable proof-theoretic system of a Curry-style language is also relative to such a type context.[19] Note also that if $\tau$ is a formula and $\tau \twoheadrightarrow_{\beta\eta} \tau'$, we want whatever notion of deduction we adopt to allow us to infer $\tau'$ from $\tau$, and to infer $\tau$ from $\tau'$.

According to the type theory that motivates systems of the form $\Rightarrow^s$, propositions are not individuals. So such systems cannot distinguish a 1-place predication whose argument is a proposition from the result of applying a 1-place operator to a proposition; in effect, a predicate of type $\langle\langle\rangle\rangle$ is a 1-place operator. A truth predicate $T$ would be a merely redundant operator: for $T : \langle\langle\rangle\rangle \in \mathcal{S}$, the $T$-principle would be this axiom: $\forall v : \langle\rangle(T(v) \leftrightarrow v)$.

Because it admits quantification into formula position, some sentences of $L^s$ are not close regimentations of any English sentences. For example, $\exists v : \langle\rangle.v$ is a sentence. English does not allow quantification into formula positions. If one had to try to express the proposition that it purports to signify in English, "some proposition is true" would be as close as one could get, though the latter would be more closely regimented by $\exists v : \langle\rangle.T(v)$, which is trivially equivalent to $\exists v : \langle\rangle.v$.[20] Although we start out understanding formal languages in terms of translation between them and a

natural language, in some cases we can eventually become able to understand formal languages in their own terms. Does this hold for $\exists v : \langle\rangle.v$? The answer is hardly obvious. I will return to this in Section 5.

## 2  Three Paradoxes, Perhaps

On the next to the last page of [44], Russell presents an argument that he takes to be paradoxical, one that appears to have played a significant role in motivating ramification.[21] Here is Russell's presentation.

> If $m$ be a class of propositions, the proposition "every $m$ is true" may or may not be itself an $m$. But there is a one–one relation of this proposition to $m$: if $n$ be different from $m$, "every $n$ is true" is not the same proposition as "every $m$ is true." Consider now the whole class of propositions of the form "every $m$ is true," and having the property of not being members of their respective $m$'s. Let this class be $w$, and let $p$ be the proposition "every $w$ is true." If $p$ is a $w$, it must possess the defining property of $w$; but this property demands that $p$ should not be a $w$. On the other hand, if $p$ is not a $w$, then $p$ does possess the defining property of $w$, and therefore is a $w$. Thus a contradiction appears unavoidable ([44, (p. 527)]).

Contrary to its initial appearance, this argument makes no essential use of classes or of truth.[22] Regarding the former: his treatment of $m$ and $n$ as if they were common nouns deserves comment. Surely by "every $m$" he meant "every member of $m$," and so forth. But his predicational uses of $m$ and $n$ suggest that clarity would be served by replacing reference to classes with reference to the propositional functions that define these classes. Below, I will construe $m$ and $n$ to signify propositional functions of type $\langle\langle\rangle\rangle$ rather than classes of propositions. Regarding the latter, in $L^s$ truth can only be represented by a redundancy operator, so we can eliminate reference to truth by translating "every $m$ is true" as $\forall v : \langle\rangle(m(v) \supset v)$.[23]

As Russell presented it, this argument relies on the following statement: "if $n$ be different from $m$, 'every $n$ is true' is not the same proposition as 'every $m$ is true.'" Avoiding the unnecessary contraposition in Russell's formulation, the mentioned statement amounts to this: Russell's dictum (hereafter RD),

$$\left(\forall v : \langle\rangle\big(n(v) \supset v\big) =^{\langle\rangle} \forall v : \langle\rangle\big(m(v) \supset v\big)\right) \supset (n =^{\langle\langle\rangle\rangle} m).$$

RD is stronger than necessary. It implies the following, which will suffice for what follows (hereafter RD)[24]:

$$\left(\forall v : \langle\rangle\big(n(v) \supset v\big) =^{\langle\rangle} \forall v : \langle\rangle\big(m(v) \supset v\big)\right) \supset \forall v : \langle\rangle\big(n(v) \leftrightarrow m(v)\big).$$

What follows reformulates this argument in the notation of Section 1.[25] Let $G_R$ be the following:

$$\left(\lambda\xi : \langle\rangle.\exists\Gamma : \langle\langle\rangle\rangle\big((\xi =^{\langle\rangle} \forall v : \langle\rangle\big(\Gamma(v) \supset v\big)) \;\&\; \neg\Gamma(\xi)\big)\right);$$

so $\Rightarrow G_R : \langle\langle\rangle\rangle$.[26] Note that the class of propositions $\{\xi : G_R(\xi)\}$ corresponds to $w$ as used in the passage quoted above; also, Russell's $p$ would be expressed by $G_R(\forall v : \langle\rangle(G_R(v) \supset v))$. Russell's argument can be reconstructed as follows.

Assume $G_R(\forall v : \langle\rangle(G_R(v) \supset v))$. By $\beta$-conversion,

$$\exists\Gamma : \langle\langle\rangle\rangle\big((\forall v : \langle\rangle\big(G_R(v) \supset v\big) =^{\langle\rangle} \forall v : \langle\rangle\big(\Gamma(v) \supset v\big)$$
$$\&\; \neg\Gamma\big(\forall v : \langle\rangle\big(G_R(v) \supset v\big)\big)\big).$$

Fix a witnessing $\Gamma : \langle\langle\rangle\rangle$ so that

$$\big(\forall v : \langle\rangle\big(G_R(v) \supset v\big) =^{\langle\rangle} \forall v : \langle\rangle\big(\Gamma(v) \supset v\big)\big)$$
$$\& \neg\Gamma\big(\forall v : \langle\rangle\big(G_R(v) \supset v\big)\big). \tag{$*$}$$

Using weak RD, the left-conjunct in ($*$) yields $\forall v : \langle\rangle(G_R(v) \leftrightarrow \Gamma(v))$. In particular,

$$G_R\big(\forall v : \langle\rangle\big(G_R(v) \supset v\big)\big) \leftrightarrow \Gamma\big(\forall v : \langle\rangle\big(G_R(v) \supset v\big)\big).$$

By the right conjunct in ($*$), $\neg G_R(\forall v : \langle\rangle(G_R(v) \supset v))$, contrary to our initial assumption. Thus $\neg G_R(\forall v : \langle\rangle(G_R(v) \supset v))$. Thus

$$\big(\forall v : \langle\rangle\big(G_R(v) \supset v\big) =^{\langle\rangle} \forall v : \langle\rangle\big(G_R(v) \supset v\big)\big) \& \neg G_R\big(\forall v : \langle\rangle\big(G_R(v) \supset v\big)\big).$$

By existential introduction,

$$\exists\Gamma : \langle\langle\rangle\rangle\big(\big(\forall v : \langle\rangle\big(G_R(v) \supset v\big) =^{\langle\rangle} \forall v : \langle\rangle\big(\Gamma(v) \supset v\big)\big)$$
$$\& \neg\Gamma\big(\forall v : \langle\rangle\big(G_R(v) \supset v\big)\big)\big).$$

By reverse $\beta$-conversion, $G_R(\forall v : \langle\rangle(G_R(v) \supset v))$, a contradiction. Call this argument Russell's purported propositional paradox. Note that this presentation shows that the mention of truth found in Russell's presentation was unnecessary.[27]

In fact, we can go further: the use of formulas of the form $\forall v : \langle\rangle(\tau(v) \supset v)$ can be eliminated from the above argument to give us what I will call the purported paradox of propositional quantification.

Let $Q$ be $\exists$ or $\forall$ or any other quantifier constant one might add to $\Rightarrow$'s lexicon. Letting $\Pi$ and $\Xi$ represent propositional functions of type $\langle\langle\rangle\rangle$, consider the following (the propositional quantification schema, hereafter PQ):

$$\big((Qv : \langle\rangle.\Pi(v) =^{\langle\rangle} Qv : \langle\rangle.\Xi(v)) \supset (\Pi =^{\langle\langle\rangle\rangle} \Xi)\big).$$

PQ implies the following (weak PQ):

$$\big((Qv : \langle\rangle.\Pi(v) =^{\langle\rangle} Qv : \langle\rangle.\Xi(v))$$
$$\supset \forall v : \langle\rangle\big(\Pi(v) \leftrightarrow \Xi(v)\big)\big).$$

Let $G_Q$ be

$$\big(\lambda\xi : \langle\rangle.\exists\Gamma : \langle\langle\rangle\rangle\big((\xi =^{\langle\rangle} Qv : \langle\rangle.\Gamma(v)) \& \neg\Gamma(\xi)\big)\big);$$

so $\Rightarrow G_Q : \langle\langle\rangle\rangle$.

Assume $G_Q(Qv : \langle\rangle.G_Q(v))$. Thus

$$\exists\Gamma : \langle\langle\rangle\rangle\big((Qv : \langle\rangle.G_Q(v) =^{\langle\rangle} Qv : \langle\rangle.\Gamma(v)) \& \neg\Gamma(Qv : \langle\rangle.G_Q(v))\big).$$

Fix a witnessing $\Gamma : \langle\langle\rangle\rangle$ so that

$$\big(Qv : \langle\rangle.G_Q(v) =^{\langle\rangle} Qv : \langle\rangle.\Gamma(v)\big) \& \neg\Gamma\big(Qv : \langle\rangle.G_Q(v)\big). \tag{$**$}$$

Using weak PQ, the left conjunct in ($**$) yields $\forall v : \langle\rangle(G_Q(v) \leftrightarrow \Gamma(v))$. In particular,

$$G_Q\big(Qv : \langle\rangle.G_Q(v)\big) \leftrightarrow \Gamma\big(Qv : \langle\rangle.G_Q(v)\big).$$

By the right conjunct in ($**$), $\neg G_Q(Qv : \langle\rangle.G_Q(v))$, contrary to our initial assumption. Thus $\neg G_Q(Qv : \langle\rangle.G_Q(v))$. Thus

$$\big(Qv : \langle\rangle.G_Q(v) =^{\langle\rangle} Qv : \langle\rangle.G_Q(v)\big) \& \neg G_Q\big(Qv : \langle\rangle.G_Q(v)\big).$$

By existential introduction,

$$\exists\Gamma : \langle\langle\rangle\rangle\big((Qv : \langle\rangle.G_Q(v) =^{\langle\rangle} Qv : \langle\rangle.\Gamma(v)\big) \,\&\, \neg\Gamma\big(Qv : \langle\rangle.G_Q(v)\big)\big).$$

Thus $G_Q(Qv : \langle\rangle.G_Q(v))$, a contradiction.

As Frege first observed, quantification is a higher-level analogue of predication. So we should expect there to be a predicational analogue of the above argument, a purported paradox of propositional predication.

For $\Pi$ and $\Xi$ as above and any formula $\varphi$ with $v \in \mathrm{Var} - FV(\varphi)$, consider the following (the propositional predication schema, hereafter PP):

$$\big(\Pi(\varphi) =^{\langle\rangle} \Xi(\varphi)\big) \supset (\Pi =^{\langle\langle\rangle\rangle} \Xi).$$

PP implies the following (weak PP):

$$\big(\Pi(\varphi) =^{\langle\rangle} \Xi(\varphi)\big) \supset \forall\mu : \langle\rangle\big(\Pi(\mu) \leftrightarrow \Xi(\mu)\big).$$

Let $G_\varphi$ be

$$\big(\lambda v : \langle\rangle\big((v =^{\langle\rangle} \Gamma(\varphi)) \,\&\, \neg\Gamma(v)\big)\big).$$

So $\Rightarrow G_\varphi : \langle\langle\rangle\rangle$.

Assume $G_\varphi(G_\varphi(\varphi))$. Thus

$$\exists\Gamma : \langle\langle\rangle\rangle\big((G_\varphi(\varphi) =^{\langle\rangle} \Gamma(\varphi)) \,\&\, \neg\Gamma(G_\varphi(\varphi))\big).$$

Fix a witnessing $\Gamma : \langle\langle\rangle\rangle$ so that

$$\big(G_\varphi(\varphi) =^{\langle\rangle} \Gamma(\varphi)\big) \,\&\, \neg\Gamma\big(G_\varphi(\varphi)\big). \qquad\qquad (* * *)$$

Using weak PP, the left conjunct yields, $\forall\mu : \langle\rangle(G_\varphi(\mu) \leftrightarrow \Gamma(\mu))$. In particular, $G_\varphi(G_\varphi(\varphi)) \leftrightarrow \Gamma(G_\varphi(\varphi))$. By the right conjunct, $\neg G_\varphi(G_\varphi(\varphi))$, contrary to our initial assumption. Thus $\neg G_\varphi(G_\varphi(\varphi))$. Thus

$$\big(G_\varphi(\varphi) =^{\langle\rangle} G_\varphi(\varphi)\big) \,\&\, \neg G_\varphi\big(G_\varphi(\varphi)\big).$$

By existential introduction,

$$\exists\Gamma : \langle\langle\rangle\rangle\big((G_\varphi(\varphi) =^{\langle\rangle} \Gamma(\varphi)) \,\&\, \neg\Gamma\big(G_\varphi(\varphi)\big)\big),$$

that is, $G_\varphi(G_\varphi(\varphi))$. This is a contradiction.

The purported paradoxes of propositional predication and quantification make no mention of truth or of any other "terms" that would be, in Ramsey's words, "empirical" rather than "formal."[28] The same goes for our formulation of Russell's purported paradox. If these arguments are paradoxes, there is no reason for classifying them as semantic paradoxes. If avoiding them counts as a good reason for ramification, that reason should not be considered a semantic reason.

Russell's purported propositional paradox is only as paradoxical as weak RD is plausible. Similarly, the purported paradoxes of propositional quantification and predication are only as paradoxical as weak PQ and weak PP are, respectively, plausible.

Before addressing these plausibility questions, some consideration of semantic matters is appropriate.

### 3  Metaphysics and Signification

Most contemporary work in semantic theory employs terminology that distinguishes between the semantic role of occurrences of a formula in "formula position" from that of its occurrences in "argument position," for example, by saying that a formula in formula position expresses a proposition, while one in argument position designates a proposition. This corresponds to the standard way of presenting the syntax of a formal language in terms of the disjoint categories of formulas and terms. In his occasional semantic remarks on his formal language, Russell did not mark this distinction.[29] So in what follows, my terminology also will not mark this distinction. The word "signify" has a history as a fudge word for the just-mentioned distinction; in what follows, I will appropriate it for that fudge.[30] This will reflect the syntactic definitions in Section 1 (and in Section 6), in which I use "term" for the broadest syntactic category and use "formula" to carve out a subcategory of the terms.

One source of obscurity in Russell's occasional semantic remarks after 1908 was the shift in his thinking between 1908 and 1910 about the metaphysical status of propositions and propositional functions that accompanied his adoption of the multiple relation theory of judgment (hereafter the MRT).[31] The MRT remained so underdeveloped that it is hyperbolic to call it a theory.[32] For our purposes, we need only consider the following idea. Facts involving propositional attitudes do not involve a thinking subject being related to a proposition. Rather, for example, if Ann judges that Socrates is wise, there is a fact relating Ann, Socrates, and wisdom under a 3-place relation, one which holds of any individuals $x$ and $y$ and any quality $q$ in that order if and only if $x$ judges that $y$ is $q$.[33] Thus the understanding of propositional attitudes does not give us a reason to posit propositions.

Both before and after this conversion to the MRT, Russell's writings on logic make reference to individuals, propositional functions, and propositions. After his conversion, his writings in a metaphysical vein make reference to objects or particulars, universals (divided into qualities and relations), and complexes (i.e., facts)—hereafter the "real entities."[34] At this stage, he regarded particulars, universals, and complexes as robust and demoted propositions to an inferior ontological status. On the "official view" in Whitehead and Russell [55], propositions are "incomplete symbols." It is far from clear what this means, but it clearly implies that they are not real entities. In a brief "metaphysical paragraph" in the introduction to [55], we are told this: "The universe consists of objects having various qualities and standing in various relations" (p. 43), while "…what we call a 'proposition' (in the sense in which this is distinguished from the phrase expressing it) is not a single entity at all" (p. 44). More explicitly, in 1918:

> Obviously propositions are nothing . . . . When I say "Obviously propositions are nothing" it is not perhaps quite obvious. Time was when I thought there were propositions, but it does not seem to me very plausible to say that in addition to facts there are also these curious shadowy things going about such as 'That today is Wednesday' when it is in fact Tuesday. I cannot believe they go about the real world. (Russell [47, p. 223])

As Church observed (see [7, p. 748, n. 4]),[35] presumably propositional functions would have to be at least as shadowy, presumably unreal, as their values.[36]

Nonetheless, in their logical writings Russell and Whitehead freely quantify over propositions and propositional functions. Putting a sentence into the that-clause of a propositional attitude construction may "complete" the proposition signified by that sentence, but it does nothing to eliminate any quantification over types other than **i**.

Today, Quine's dictum "To be is to be the value of a variable" (see [38, p. 15]) is likely to strike us as boringly obvious. Its philosophical punch is clearer against the background of Whitehead and Russell's unembarrassed quantification over entities that they thought, in the final analysis, were unreal. In Section 6 I will suggest an interpretation of ramified-type languages that would allow Whitehead and Russell to evade this Quinean reproach, although not without friction with another Russellian principle.

Are all individuals particulars? Are all particulars individuals? Are universals individuals? I know of no writings in which Russell gives a clear answer to these basic questions. I suspect that the answer to the first question is "yes" and to the second question, "no, because complexes are not individuals." Bernard Linsky defends a "no" to the third question on the basis of passages in which Russell applies type concepts to universals (see [28]). On the other side, consider this. Before accepting the theory of types (i.e., while still clinging to the doctrine of "the unrestricted variable"), Russell apparently thought that "...there is really nothing that is not an individual" (see p. 206 of "On 'Insolubilia' and their solution by symbolic logic," in [50]). It would seem that at this time Russell considered universals to be individuals. I know of no evidence that he changed his mind on that after accepting the theory of types. Linsky did not consider the view on which universals and particulars are individuals, but nonetheless they fall into a metaphysical-type, as opposed to a logical-type, hierarchy. The metaphysical types could be represented by the $\langle\rangle$-free members of $\mathcal{T}$, except with **i** interpreted as representing the metaphysical type of all and only particulars. This double use of types would reconcile the passages Linsky cites and the thesis that universals are individuals.

Be the metaphysical details as they may, surely real entities have to play some role in Russell's post-MRT understanding of the meanings of terms in a ramified-type language; but what is that role?

If $\iota : \mathbf{i}$ and $\pi$ is a 1-place predicate of individuals, $\pi(\iota)$ signifies the proposition $p$ (in a way that fully analyzes $p$), and $\iota$ signifies a constituent of $p$.[37] But here is one clear and quite important principle within the Russell–Whitehead project: "It should be remembered that a function is not a constituent in one of its values: thus for example the function '$\widehat{x}$ is human' is not a constituent of the proposition 'Socrates is human'" (see [55, pp. 54–55]). So taking $\pi$ to signify a propositional function, it does not signify a constituent of $p$. The root of this idea is already present in [44]. There Russell says that propositional functions depend on propositions, which depend on their constituents, and this dependence is supposed to be noncircular.[38] In fact, it would seem that objects of acquaintance would be real entities. According to the MRT, propositional functions are not real entities, so on that theory Russell could not allow an occurrence of a term of type other than **i** in argument position within a sentence to signify a propositional constituent, or (more in keeping with the MRT) a component of a judgment-fact of the sort that could correspond to, for example, a judgment-attribution with that sentence in the "that" clause; the (so to speak) "constituents of propositions" would have to be objects of acquaintance and so real entities.

Does "wise" (or "is wise") correspond in some way to some constituent of the proposition signified by "Socrates is wise"? According to Peter Hylton, "The proposition that Socrates is wise contains some constituent other than Socrates, but it is unclear what this is, and how it is related to the propositional function '$\widehat{x}$ is wise'"

(see [22, p. 301, n. 23]). Nonetheless, assuming that the proposition that Socrates is wise is elementary, there is a case to be made for the thesis that the quality of being wise (i.e., wisdom) is the unique constituent other than Socrates of that proposition. One piece of evidence: in presenting the MRT in [49, p. 126], Russell says that loving is a constituent of Othello's judgment that Desdemona loves Cassio. According to this theory, an expression of a proposition is an "incomplete symbol," and "when I judge 'Socrates is human,' the meaning is completed by the act of judging" ([55, p. 44]). This suggests that in the derivative sense in which we can speak of propositions at all, being wise is a constituent of the proposition that Socrates is wise. In what follows, I will assume this.[39]

More generally, I will assume that for any sentence that provides a full analysis of the proposition that it signifies, each occurrence of a nonlogical constant (in the notation from Section 1, of a member of $\text{dom}(\mathcal{S})$) in that sentence stands for a genuine constituent of that proposition, and every genuine constituent is so represented.[40] A Russellian semantic theory would say that the occurrence of a predicate-constant $\pi$ in $\pi(\iota)$ does two kinds of work: it does the "logical" job of signifying a propositional function, and the "basic" job of standing for a quality that serves as a constituent of $p$, with the logical job uniquely determined by the basic job. More generally, an occurrence of a term in predicate position in a formula signifies a propositional function, which it does not contribute as a constituent of the proposition signified by that formula, but it might also stand for a universal which it does so contribute. Thus the distinction between universals and propositional functions requires that a Russellian semantics factor into a "basic," referential, component and a "logical" component concerning what I am calling signification.[41]

To spell this out, I will introduce a notion of proxyhood relating real to logical entities. I will continue playing the role of a Russellian trying to get by with simple types. Of course, ultimately, the following discussion would have to be rewritten using ramified rather than simple types.[42]

At type $\mathbf{i}$, signification coincides with standing-for: if $\iota : \mathbf{i} \in \mathcal{S}$ and $x$ is an individual, we can allow that $\iota$ stands for $x$ and in virtue of that signifies $x$. Let each individual be its own proxy. If $R$ is an $n$-place universal that is instanced by or relates only individuals and $\pi : \langle \mathbf{i}, \ldots, \mathbf{i} \rangle \in \mathcal{S}$, we can allow that $\pi$ stands for $R$, and, in virtue of that fact, it signifies the propositional function $f$ such that for any individuals $x_1, \ldots, x_n$, $f(x_1, \ldots, x_n)$ is the proposition that $R$ relates $x_1, \ldots, x_n$ in that order. In this case, call $R$ the proxy for $f$. If every universal is instanced by or relates only individuals, this completely describes the link between standing-for and signification.

If we want to allow for meaningful predicate constants $\pi : \langle t_1, \ldots, t_n \rangle \in \mathcal{S}$ such that for some $i \in (n)$, $t_i \neq \mathbf{i}$, we must extend our use of types by assuming that each universal is of a type and then extend our notion of proxyhood. It far from obvious how to make sense of this idea in terms in keeping with the MRT. Here is an attempt.

First, consider types in which $\langle \rangle$ does not occur. Let a universal $R$ have type $\langle t_1, \ldots, t_n \rangle \in \mathcal{T}$, with the latter $\langle \rangle$-free, if and only if $R$ is instanced by or relates $x_1, \ldots, x_n$ in that order only if for each $i \in (n)$, $x_i$ is a real entity of type $t_i$. Then for $\pi : \langle t_1, \ldots, t_n \rangle \in \mathcal{S}$, we can allow that $\pi$ stands for $R$, and in virtue of that fact it signifies the following propositional function $f$: for any logical entities $y_1, \ldots, y_n$ such that $y_i$ has type $t_i$ and $x_i$ is the proxy for $y_i$, this for each $i \in (n)$, $f(y_1, \ldots, y_n)$

is the proposition that $R$ relates $x_1, \ldots, x_n$ in that order. In this case, call $R$ the proxy for $f$.

So far, our attempt involved no sacrifice of Russellian principles. But if we want our basic semantics to allow for meaningful predicate-constants $\pi : \langle t_1, \ldots, t_n \rangle \in \mathcal{S}$ for any $n \in \omega$ and $\langle t_1, \ldots, t_n \rangle \in \mathcal{T}$, we would have to posit real entities of type $\langle \rangle$. At first glance, it may seem that the entities that Russell called "complexes" would be appropriate real entities to take type $\langle \rangle$.[43] Russell's complexes are what he later called facts. In the period under consideration, Russell took facts to be real, so to preserve the Russellian spirit, we might try taking them to be proxies for propositions. But this will not work. Russell thought that for each proposition $p$ there was a fact $c$ such that either $p$ positively corresponded to $c$ (in which case it was true) or $c$ negatively corresponded to $c$ (in which case $c$ was false).[44] Were we to take $c$ to be the proxy for $p$, it seems that we would have to take it to be a proxy for the negation of $p$ as well. For $\pi : \langle \rangle \in \mathcal{S}$, if we then allowed that $\pi$ stands for a quality $q$ of proxies for propositions, the proxy for $p$ would have $q$ if and only if the proxy for the negation of $p$ also has $q$. If what propositional functions are true of $p$ depends on what qualities $p$'s proxy has, we would be forced to accept that for a sentence $\varphi$ signifying $p$, $(\varphi =^{\langle \rangle} \neg\varphi)$ is true. A *strict* Russellian should not grant proxies to propositions.

It seems that the only reasonable way to secure a real proxy for each proposition would be to assume that for each proposition there is a corresponding situation (or state of affairs), and maintain that the latter "obtains" if and only if the former is true. Such situations would merely be Russellian propositions, according to his 1903 conception of propositions, all over again: that conception of propositionhood in effect collapses propositions to such situations. There is no evidence that between 1903 and his conversion to the MRT Russell changed his conception of propositions; it seems that the reality of such potential complexes was exactly what he repudiated when he consigned propositions to the shadows. Since situation semantics is generous in its positing of situations that may not be actual, I will call the proposal that each proposition has a corresponding situation as its proxy "the situationist path." With this non-Russellian metaphysical backdrop, a basic semantics generates a corresponding logical semantics for a language of the form $L^s(\mathcal{S})$ without constraining assumptions about $\mathcal{S}$.

Assume that basic semantics is given: for each $\iota : \mathbf{i} \in \mathcal{S}$, $\iota$ stands for a unique individual; for $n \in \omega - \{0\}$ and each $\pi : \langle t_1, \ldots, t_n \rangle \in \mathcal{S}$, $\pi$ stands for a unique universal $R^\pi$ of type $\langle t_1, \ldots, t_n \rangle$.[45] For a type context $C$, let $\mathbf{a}$ be a $C$-assignment iff $\mathbf{a}$ is a function on $\mathrm{dom}(C)$ such that for each $\nu : t \in C$, $\mathbf{a}(\nu)$ is an entity of type $t$. If $\mathbf{a}$ is a $C$ assignment, $C, \nu_1 : t_1, \ldots, \nu_n : t_n$ is a type context, $\nu_1, \ldots, \nu_n$ are distinct, and $x_i \in \mathcal{F}(t_i)$ for each $i \in (n)$, for every $\nu \in \mathrm{Var}$ let

$$\mathbf{a}[\nu_1, \ldots, \nu_n := x_1, \ldots, x_n](\nu) = \begin{cases} x_i & \text{if } \nu \equiv \nu_i \text{ for } i \in (n), \\ \mathbf{a}(\nu) & \text{otherwise.} \end{cases}$$

Thus $\mathbf{a}[\nu_1, \ldots, \nu_n := x_1, \ldots, x_n]$ is a $(C, \nu_1 : t_1, \ldots, \nu_n : t_n)$-assignment. The logical semantics for $\Rightarrow_\mathcal{S}$ will assign each $\langle C, \tau \rangle \in L^s$ such that $\tau$ is $\beta\eta$-normal, and each $C$-assignment $\mathbf{a}$, to the entity $[\tau]^{\mathbf{a}}$ that it signifies, as follows.[46]

(1) For each $\nu : t \in C$, $[\nu]^{\mathbf{a}}$ is $\mathbf{a}(\nu)$.
(2.1) For each $\iota : \mathbf{i} \in \mathcal{S}$, $[\iota]^{\mathbf{a}}$ is the individual for which $\iota$ stands.

(2.2) For each $\pi \,:\, \langle t_1, \ldots, t_n \rangle \,\in\, \mathcal{S}$, $[\pi]^{\mathbf{a}}$ is the propositional function with proxy $R^{\pi}$.

(3) If $\Rightarrow \tau \,:\, \langle t_1, \ldots, t_n \rangle$ with $n > 0$, and for each $i \in (n) \Rightarrow \tau_i : t_i$, $[\tau(\tau_1, \ldots, \tau_n)]^{\mathbf{a}}$ is $[\tau]^{\mathbf{a}}([\tau_1]^{\mathbf{a}}, \ldots, [\tau_n]^{\mathbf{a}})$.

(4.1) If for $i = 0, 1$, $C \Rightarrow \tau_i : \langle\rangle$, $[(\tau_0 \supset \tau_1)]^{\mathbf{a}}$ is the proposition that either $[\tau_0]^{\mathbf{a}}$ is false or $[\tau_1]^{\mathbf{a}}$ is true.

(4.2) $[\bot]^{\mathbf{a}}$ is the absurd proposition. (Fix one if you think that there are several.)

(5) If $C, \nu : t \Rightarrow \tau : \langle\rangle$, $[\exists \nu : t.\tau]^{\mathbf{a}}$ is the proposition that for some entity $x$ of type $t$, $[\tau]^{\mathbf{a}[\nu := x]}$ is true.

(6) If $n > 0$ and $C, \nu_1 : t_1, \ldots, \nu_n : t_n \Rightarrow \tau : \langle\rangle$ for distinct $\nu_1, \ldots, \nu_n \in FV(\tau)$, then $[(\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\tau)]^{\mathbf{a}}$ is the function $f$ so that for any entities $x_1, \ldots, x_n$ with $x_i$ of type $t_i$ for each $i \in (n)$, $f(x_1, \ldots, x_n)$ is $[\tau]^{\mathbf{a}[\nu_1, \ldots, \nu_n := x_1, \ldots, x_n]}$.[47]

Finally, we extend this definition to cover every $\langle C, \tau \rangle \in L^s$ and $C$-assignment $\mathbf{a}$: if $\tau$ is not $\beta\eta$-normal and $\tau'$ is its $\beta\eta$-normalization, let $[\tau]^{\mathbf{a}}$ be $[\tau']^{\mathbf{a}}$.

## 4 The Syntactic Picture and the Plausibility Problem

Along with [55], Russell's published writings from 1903 to 1910 show that he conceived of a proposition as having a structure that in some sense is reflected by the syntactic structure of an appropriate sentence (in an optimal formal language) that expresses it. At the moment we will leave open what form an optimal formal language would take. The works [55] and [41] extend this picture to propositional functions: Russell and Whitehead thought of a propositional function as obtainable by "punching holes" in a proposition, with free variables (which our notation then indicates with a $\lambda$-prefix) indicating these holes. Their discussion of propositional functions might seem odd by contemporary standards, since they tend to understand a formula containing at least one free variable to signify a propositional function (with its free variables representing that function's argument places)—this even though they also allow for circumflex binding of free occurrences of variables to indicate those argument places. They did not reach the contemporary understanding of free variables in a formula as "parameters," with the formula signifying a proposition only relative to a variable assignment, and with the distinction between that and a term (for us a $\lambda I$-term) that properly signifies a propositional function (though that too will be relative to a variable assignment in the case when a $\lambda I$-term contains free variables). With that in mind, what the following remark describes would correspond, in our notation, to peeling off a $\lambda$-prefix and assigning values to the previously bound variables to "fill the holes" signaled by that $\lambda$-prefix:

> By a "propositional function" we mean something which contains a variable $x$, and expresses a *proposition* as soon as a value is assigned to $x$. That is to say, it differs from a proposition solely by the fact that it is ambiguous: it contains a variable of which the value is unassigned. ([55], p. 38)

Thus the structure of a propositional function is reflected by that of the scope of a closed $\lambda I$-term that signifies that function. (This is the Russellian basis of our restriction of $\lambda$-terms to $\lambda I$-terms.[48]) Speaking more strictly, the value $q$ of a 1-place propositional function $f$ for an appropriate argument contains (in some sense) that argument, and any value of $f$ for another appropriate argument can be formed from $q$ by replacing the former argument by the latter. Similarly for multiple-place propositional functions, I will call these ideas the naive syntactic picture of propositions and propositional functions.[49]

If a sentence $\varphi$ $\beta$-converts to $\varphi'$, should one think of $\varphi$ and $\varphi'$ as signifying the same proposition? And if a closed term $\tau$ whose type has the form $\langle t_1, \ldots, t_n \rangle$ for $n > 0$ $\eta$-converts to $\tau'$, should one think of $\tau$ and $\tau'$ as signifying the same propositional function? If these are terms in an optimal language, the syntactic picture, at least if applied naively, might suggest that "no" is the correct answer to both questions: after all, $\varphi$ and $\varphi'$ are distinct sentences, and $\tau$ and $\tau'$ are distinct closed terms.

But further thought about the syntactic picture should persuade us that these answers were naive. For suppose we form a propositional function by punching a hole in a proposition; applying that function to what we just punched out would, presumably, consist in returning what we punched out to the hole that it left, yielding the original proposition and a "yes" for the first question. A similar point applies to the second question.[50]

Thus the naive syntactic picture is too strong. Here is an improvement: a proposition has a structure reflected by a $\beta\eta$-normal sentence in an optimal formal language that regiments it. Similarly, a propositional function has a structure that reflects the matrix of a closed $\beta\eta$-normal $\lambda I$-term in such a language that signifies that function.[51] (The matrix of a $\lambda I$-term is the scope of its $\lambda$-prefix.)

We are finally ready to address the plausibility questions. To give the purported paradoxes full benefit of the doubt, let us assume that $L^s$ is optimal with regard to the syntactic picture.

Whatever plausibility weak RD, weak PQ, and weak PP have, they accrue from the plausibility of RD, PQ, and PP, respectively. Before addressing the latter schemata, let us consider a paradigm of correct converse-compositional reasoning. Claim: for any $n_0$-place $\pi_0 : \langle \mathbf{i}, \ldots, \mathbf{i} \rangle \in \mathcal{S}$, any $n_1$-place $\pi_1 : \langle \mathbf{i}, \ldots, \mathbf{i} \rangle \in \mathcal{S}$, $\iota_{0,i} : \mathbf{i} \in \mathcal{S}$ for $i \in (n_0)$, and $\iota_{1,i} : \mathbf{i} \in \mathcal{S}$ for $i \in (n_1)$,

$$\left( \pi_0(\iota_{0,1}, \ldots, \iota_{0,n_0}) =^{\langle\rangle} \pi_1(\iota_{1,1}, \ldots, \iota_{1,n_0}) \right) \supset (\pi_0 =^{\langle \mathbf{i}, \ldots, \mathbf{i} \rangle} \pi_1)$$

is true. Assume the if-clause. So $[\pi_0(\iota_{0,1}, \ldots, \iota_{0,n_0})] = [\pi_1(\iota_{1,1}, \ldots, \iota_{1,n_0})]$; let $p$ be that proposition. Whether or not $L^s$ is optimal, at least the syntactic picture applies to its atomic sentences and dictates that $p$ has predicative structure: for some $n > 0$, $p$ predicates some $n$-place universal $P$ of $n$-many individuals $x_1, \ldots, x_n$ in that order; furthermore, $n_0 = n = n_1$, and $\pi_0$ and $\pi_1$ stand for $P$. So $[\pi_0] = [\pi_1]$; so $(\pi_0 =^{\langle \mathbf{i}, \ldots, \mathbf{i} \rangle} \pi_1)$ is well formed and true.

But RD, PQ, and PP are rather different counter-compositional principles. Does the syntactic picture support them?[52] This question is not yet well defined: presentation of these schemata in Section 2 was unspecific about the values of $n$ and $m$ in RD, and of $\Pi$ and $\Xi$ in PQ and PP.

For a start, assume that $\Pi : \langle\langle\rangle\rangle$, $\Xi : \langle\langle\rangle\rangle \in \mathcal{S}$. To give the purported paradoxes their best run for the money, we will extend our paradigm of converse-compositional reasoning. This will require us to have taken the situationist path in Section 3, allowing us to say that $\Pi$ stands for a quality $P$ of potential complexes and signifies propositional functions $f$ such that for any proposition $p$, $f(p)$ is true if and only if $p$'s proxy has $P$; similarly for $\Xi$ and a quality $X$.

Assume that $C \Rightarrow \varphi : \langle\rangle$ and $\mathbf{a}$ is a $C$-assignment. Consider the instance of PP for our given $\Pi$ and $\Xi$. Let $p$ be $[\Pi(\varphi)]^{\mathbf{a}}$, and let $q$ be $[\Xi(\varphi)]^{\mathbf{a}}$. The syntactic picture dictates that $p$ and $q$ are 1-place predications and that $p$ is the proposition that predicates $P$ of the proxy for $[\varphi]^{\mathbf{a}}$ (a potential complex). And similarly, $q$ predicates $X$ of that proxy. Assume that $(\Pi(\varphi) =^{\langle\rangle} \Xi(\varphi))$ is true. Thus $p = q$. Thus $p$

and $q$ have the same predicating constituent, that is, $P = X$. Thus $[\Pi] = [\Xi]$. So $(\Pi =^{\langle\langle\rangle\rangle} \Xi)$ is true, verifying the relevant instance of PP.

Now let $Q$ be $\exists$ or $\forall$ or any other quantifier constant. So

$$\Rightarrow Q\nu : \langle\rangle.\Pi(\nu) : \langle\rangle \quad \text{and} \quad \Rightarrow Q\nu : \langle\rangle.\Xi(\nu) : \langle\rangle.$$

Let $p$ be $[Q\nu : \langle\rangle.\Pi(\nu)]$, and let $q$ be $[Q\nu : \langle\rangle.\Xi(\nu)]$. The syntactic picture dictates that $p$ and $q$ are 1-place quantifications, which we can think of as predications with $[(\lambda\nu : \langle\rangle.\Pi(\nu))]$ and $[(\lambda\nu : \langle\rangle.\Xi(\nu))]$ in what corresponds to argument position. Assume that $(Q\nu : \Pi(\varphi) =^{\langle\rangle} Q\nu : \Xi(\varphi))$ is true. Thus $p = q$. Thus $p$ and $q$ have the same predicated constituent, that is, $[(\lambda\nu : \langle\rangle.\Pi(\nu))] = [(\lambda\nu : \langle\rangle.\Xi(\nu))]$. So for any $\{\nu : \langle\rangle\}$-assignment $\mathbf{a}$, $[\Pi(\nu)]^{\mathbf{a}} = [\Xi(\nu)]^{\mathbf{a}}$. For any sentence $\varphi$, let $\mathbf{a}(\nu) = [\varphi]$; so $[\Pi(\varphi)] = [\Xi(\varphi)]$, which is a proposition whose predicating constituent is both $P$ and $X$. As above, it follows that $P = X$. As above, this verifies the relevant instance of PQ.

Similar considerations show that assuming that $n : \langle\langle\rangle\rangle, m : \langle\langle\rangle\rangle \in \mathcal{S}$, the corresponding instance of RD is true.

Summing up: by taking the situationist path, we have argued from the syntactic picture to the instances of PP, PQ, and RD for which $\Pi$ and $\Xi$ are predicate constants. But now the crucial point: our paradigms of converse-compositional reasoning do not support the instances of PP, PQ, or RD used in the purported paradoxes. This is because $G_\varphi$, $G_Q$, and $G_R$ are $\lambda I$-terms, not predicate constants.[53] Even after taking the situationist path, we cannot say that these terms stand for constituents of the propositions signified by $G_\varphi(\varphi)$ (relative to an appropriate assignment), $Q\nu : \langle\rangle.G_Q(\nu)$, or $\forall\nu : \langle\rangle(n(\nu) \supset \nu)$, respectively.

We might try adding further metaphysical assumptions to the situationist path: that the propositional functions signified by these three $\lambda I$-terms have proxies. By themselves, even these extravagant assumptions will not suffice. In each purported paradox we "fix" a $\Gamma$ of type $\langle\langle\rangle\rangle$, for a use of existential elimination. ($\Gamma$ here was an eigenvariable in each argument.) We cannot say that $\Gamma$ stands for a constituent of the proposition signified by $\Gamma(\varphi)$, $Q\nu : \langle\rangle.\Gamma(\nu)$, or $\forall\nu : \langle\rangle(\Gamma(\nu) \supset \nu)$, except relative to specific $\{\Gamma : \langle\langle\rangle\rangle\}$-assignment. To transpose the above arguments concerning a pair of predicate constants into arguments applicable to a $\lambda I$-term and the eigenvariable $\Gamma$, we would need to assume that every possible value of $\Gamma$ has a proxy. If that assumption is not to be ad hoc, it would seem that we should assume that every propositional function of type $\langle\langle\rangle\rangle$ has a proxy.

This move would be quite ad hoc. In fact, we only need a $\pi : \langle\langle\rangle, \langle\rangle\rangle \in \mathcal{S}$ to show that the latter suggestion is unreasonable. Given a sentence $\varphi$, let $\Pi$ be $(\lambda\nu : \langle\rangle.\pi(\varphi, \nu))$, and let $\Xi$ be $(\lambda\nu : \langle\rangle.\pi(\nu, \varphi))$. Let $P$ and $X$ be proxies for $[\Pi]$ and $[\Xi]$. So $P$ is the predicated constituent of $[\Pi(\varphi)]$, and $X$ is the predicated constituent of $[\Xi(\varphi)]$. But $(\Pi(\varphi) =^{\langle\rangle} \Xi(\varphi))$ must be true, in which case $[\Pi(\varphi)] = [\Xi(\varphi)]$. So the predicated constituents of that proposition are identical, that is, $P = X$. So $\forall\mu : \langle\rangle(\Pi(\mu) \leftrightarrow \Xi(\mu))$ is true. But the latter is equivalent to $\forall\mu : \langle\rangle(\pi(\varphi, \mu) \leftrightarrow \pi(\mu, \varphi))$. Surely we do not want to assume that the latter is true for any $\pi : \langle\langle\rangle, \langle\rangle\rangle \in \mathcal{S}$ and sentence $\varphi$.

In conclusion: even if we take the situationist path and are generous in our assumptions about the existence of proxies, the principles on which the purported paradoxes depend are too implausible to make the latter paradoxical. Having set

aside semantic paradoxes, I conclude that the purported paradoxes provide the Russellian no good reason for ramification.

For whom would the paradoxes provide such a reason? Suppose one took variables of type $\langle t_1, \ldots, t_n \rangle$ to range over universals, rather than propositional functions, with the presupposition that every $\lambda I$-term stands for a universal. Under this construal of $L^s$, RD, PQ, and PP secure some plausibility, and the purported paradoxes acquire some bite. But even so, ramification would not be the only response. Perhaps instead we should ban the binding of variables of propositional type, thus ruling RD, PQ, and PP ill formed. The intelligiblity of such binding is not obvious—recall the last paragraph of Section 1.

## 5   The Vicious-Circle Principle

Even if the arguments of Section 2 were paradoxical, the brute justification by avoidance—"Ramification lets us avoid these paradoxes; so a ramified-type logic is better than a simple-type logic"—should feel shallow without a supplemental "philosophical" diagnosis of the paradoxes that favors ramification. As a start in that direction, Russell offered the vicious circle principle (VCP) (see [41], [55]), to which I will now turn.

In different places these texts offer different formulations of the VCP. As Gödel points out, more than one principle seems to be in play.[54] I will consider a "specificational" reading of the VCP.

Let a predicative specification of an entity be a specification expressible by a term meeting this condition: every occurrence of a bound variable in that term (or its regimentation into a formal language that handles quantification and abstraction in the usual ways) has a range which does not contain the entity in question.[55] For example, Ramsey's example, "the tallest man in the room" will not express a predicative specification: relative to any context in which it successfully specifies something (i.e., a context in which there is a unique tallest man in the room), it contains an occurrence of a bound variable whose range includes the man whom the description specifies.

Here is the specificational reading of the VCP (hereafter just the VCP): in any context of language use (including thought) in which a term signifies something, every occurrence of a bound variable in an analysis of that term ranges over entities each of which can in principle be predicatively specified. Returning to Ramsey's example, relative to any context of use, an analysis of "the tallest man in the room" contains a bound variable ranging over at least the men in the room; this reading of the VCP would require that each man in the room (and any other individuals in that range) be predicatively specifiable.[56] Presumably that is the case: in some contexts we can in principle specify men by using demonstratives, in others by using names, and so forth.

Simple-type languages would seem to flout the VCP. Assume that $\varphi$ is a sentence containing a bound occurrence of a variable of type $\langle \rangle$ and that $\varphi$ signifies something—a proposition. Then every proposition would have to have a predicatve specification. In particular, the proposition signified by $\varphi$ would have to have a predicative specification. Presumably that would have to be another sentence that signifies that proposition but contains no bound occurrence of a variable of type $\langle \rangle$. This would have to hold for every such $\varphi$, and that is implausible. This problem extends to $\lambda I$-terms. Each closed $\lambda I$-term purports to specify a propositional function with a

given place number, but if it contains a bound occurrence of a variable ranging over propositional functions with that number of places, it does not predicatively specify that propositional function. And it is implausible that every such propositional function can be predicatively specified in some other way. Without further clarification of what specification amounts to, there seems little hope of turning this plausibility argument into a demonstrative argument. But even so, its should have some force: if we like the VCP, we should shun simple-type languages, and with them the theory of simple types.

But the VCP is hardly self-evident. Why accept it? Again, it is not enough to say that by accepting it we can avoid paradoxes—as Russell recognized.[57] One might as well accept the maxim "Don't engage in deductive reasoning"; by avoiding deductive reasoning, we can avoid paradoxes. For failure to respect the VCP as serving as a diagnosis of paradoxes, we need to have a prima facie reason to think that VCP deserves respect, one independent of the paradox-avoiding consequences of respecting it.

Some authors suggest that an "in a sense constructivist" cast of mind would secure a paradox-independent motivation for accepting the VCP.[58] In what would such a cast of mind consist?

In [51], Mark Sainsbury considers "the weird and tortuous interpretation" according to which Russell and Whitehead intended that propositional functions be linguistic expressions—presumably interpreted expression types. In its support, Sainsbury quotes Russell from late in his career. First in 1940: "In the language of the second order, variables denote symbols, not what is symbolized."[59] And then in 1959: "Whitehead and I thought of a propositional function as an expression."[60] Sainsbury thinks that the only alternative to this unpleasant interpretation would be the interpretation according to which propositional functions are properties (see [51, p. 280]). In Section 3 I argued that this reading would confuse logical with real entities.

We need not go to a weird and tortuous interpretation of Russell to understand the force of these late remarks. Sainsbury in fact suggests one: Russell was thinking of quantification over types other than **i** in a tacitly substitutional way, and in this respect his cast of mind was constructivist.[61] This way of thinking of such quantification predated Russell. In the *Grundgesetze*, Frege's informal explanation of his notation for first-level universal quantification is objectual (using the phrase that Montgomery Furth translates as "for every argument" in Section 8; see Frege [15, p. 41]), while Frege's informal explanation of his notation for second-level universal quantification is substitutional (using the phrase "whatever function-name one may substitute" in Section 20; see [15, p. 71]).[62] This feature of Frege's presentation of his ideas might not be of much significance for Frege's project, but it still may have influenced Russell.

But what is it to understand a quantifier expression, or an occurrence of a quantifier prefix, substitutionally?

A substitutional interpretation of a quantifier expression is relative to a vocabulary set, the latter determining the range of relevant substitutends. Since we are here interested in s-type languages, consider an s-typed vocabulary set $\mathcal{S}$, and consider $\langle \{v : t\}, \tau \rangle \in L^s(\mathcal{S})$. One way to understand talk of a substitutional interpretation of the initial prefix of $\exists v : t.\tau$ would be in terms of truth conditions. Such an understanding would at least require that one accept this biconditional: the sentence $\exists v : t.\tau$ is true if and only if there is a $\sigma$ so that $\Rightarrow_{\mathcal{S}} \sigma : t$ and $\tau[v := \sigma]$ is true. But just accepting such a biconditional is too weak to count as an interpretation

of the quantifier prefix of a sentence of the form $\exists v : t$. It is better to put the truth-conditional construal thus: the truth of $\exists v : t.\tau$ would consist in there being a $\sigma$ so that $\Rightarrow_{\mathcal{S}} \sigma : t$ and $\tau[v := \sigma]$ is true.

An even stronger construal of talk of a substitutional interpretation is available, one that concerns the identity of signified propositions rather than merely the truth conditions for sentences. It will require allowing for infinitary propositions (which of course a Russellian may ultimately dismiss as "incomplete symbols"). On this construal, $\exists v : t.\tau$ signifies the proposition that is "the" disjunction of all propositions signified by formulas of the form $\tau[v := \sigma]$ as $\sigma$ ranges over the closed terms assigned type $t$ under $\Rightarrow_{\mathcal{S}}$.[63] On this construal, an occurrence in a sentence of an existential quantification over a type other than **i** in $L^s$ is a departure from optimality, since it does not reflect quantification within the proposition signified by that sentence; rather it reflects an infinitary disjunction within that proposition. In propositions, and thus in an optimal language, the only genuine quantification is over individuals.

I will pursue this idea in Section 6. But applied to a simple-type language $L^s$, even the weaker construal substitutional quantification (in terms of truth conditions) will not do: it will not permit an inductive specification of truth conditions for sentences containing quantification restricted to s-types other than **i**. For example, to specify inductively the truth conditions for a sentence of the form $\exists v : \langle\rangle.\varphi$ (with $v \in FV(\varphi)$), we would need to have available the truth conditions for every formula of the form $\varphi[v := \sigma]$ where $\Rightarrow \sigma : \langle\rangle$, which in turn requires us to have available the truth conditions for each such $\sigma$. But one of these $\sigma$s is $\exists v : \langle\rangle.\varphi$.

We have reached a good, though conditional, reason to ramify: if we are to interpret quantification over types other than **i** substitutionally, simple types will not do. It remains only to check that ramification will.

Before addressing that task, let me mention two consequences of interpreting quantification over types other than **i** substitutionally. On the plus side, it promises to insure that sentences that involve quantification into formula position, for example, $\exists v : \langle\rangle.v$, are intelligible.[64] But there is a big minus. Under a substitutional interpretation, the meaning of quantification over affected types is tied to a particular choice of vocabulary; adding nonlogical constants to, or dropping them from, a language's lexicon would (depending on the type) change the available substitutends and with that the meaning of quantifier prefixes restricted to such types. This underlines the difference between (a) interpreting such a prefix substitutionally and (b) interpreting it objectually but restricted to a domain of linguistic expressions. Under an interpretation of the latter sort, a change of nonlogical constants in our vocabulary would have no effect on the meaning of quantifier expressions; for example, the naming of a child would not change the meaning of "everyone" and "someone"; under the substitutional interpretation, it could.[65]

If our goal is to model natural language, for many areas of discourse a substitutional interpretation of quantification is unappealing.[66] But that was not the goal pursued by Whitehead and Russell. They were only interested in consideration of a formal language that could in principle limn reality, doing this by virtue of a specific (up to choice of linguistic expressions), though unspecified, lexicon, one that permitted signification of all propositions. Their project is not unscathed by this point. Russell wrote "It is one of the marks of a proposition of logic that, given a suitable language, such a proposition can be asserted in such a language by a per-

son who knows the syntax without knowing a single word of the vocabulary" (see Russell [43, pp. 200-1]). If a quantifier prefix in the sentence used to make such an assertion is to be interpreted substitutionally, and a relevant substitutend contained an un-understood word, the speaker would not understand a relevant substitutend and so would not understand that quantifier prefix and so would not understand that sentence!

Assume for the moment that, as Fuhrmann suggested, throughout the period leading up to his work on *Principia* Russell thought that Frege's appendix theorem exposed the fallacy in the purported propositional paradox. Then an adequate rational reconstruction of Russell's views could not attribute to him a substitutional interpretation of quantification over type $\langle\langle\rangle\rangle$, this because Frege's appendix theorem (and even just its $f(\xi) = \forall v : \langle\rangle(\xi(v) \supset v)$ instance, which Russell cites in his letter to Frege of May 24, 1903) cannot be proved on such an interpretation.[67] If such a reconstruction were to offer a substitutional interpretation of such quantification as Russell's motivation for ramification (which I have avoided doing), it would have to ascribe to Russell a change in the interpretation of such quantification, from objectual to substitutional, between his writing of that letter and his commitment to ramification. (Perhaps the move to the MRT caused such a change.)

## 6 Ramification

In this section, I will present two ramified-type assignment systems.[68]

First, we have inductive definitions of the ramified (hereafter r-) types. Each type will have a height and an order.

**Definitions** **i** is an r-type, and $ht(\mathbf{i}) = \text{ord}(\mathbf{i}) = 0$. For any $m \in \omega - \{0\}$, $\langle m \rangle$ is an r-type and $ht(\langle m \rangle) = 0$, $\text{ord}(\langle m \rangle) = m$. For any $m, n \in \omega - \{0\}$, if $t_1, \ldots, t_n$ are r-types with $\text{ord}(t_i) < m$ for each $i \in (n)$, then $\langle t_1, \ldots, t_n, m \rangle$ is an r-type, and

$$ht\big(\langle t_1, \ldots, t_n, m \rangle\big) = \max_{i \in (n)}\big\{ht(t_i) + 1\big\},$$

$$\text{ord}\big(\langle t_1, \ldots, t_n, m \rangle\big) = m.$$

Since $\max\{\} = 0$, $ht(\langle m \rangle) = 0$. Let $\mathcal{T}^r = $ the set of r-types.

Informally, **i** is the type of individuals; $\langle m \rangle$ is the type of propositions of order $m$; for $n > 0$, $\langle t_1, \ldots, t_n, m \rangle$ is the type for $n$-place propositional functions of order $m$ taking arguments of type $t_i$ at the $i$th place for each $i \in (n)$. Let $t \in \mathcal{T}^r$ be propositional if and only if for some $m \in \omega - \{0\}$, $t = \langle m \rangle$.

For $t \in \mathcal{T}^r$, let the simplification of $t$, $\text{smp}(t)$, result from deleting all the rightmost (order) components of $t$; that is, $\text{smp}(\mathbf{i}) = \mathbf{i}$; $\text{smp}(\langle t_1, \ldots, t_n, m \rangle) = \langle \text{smp}(t_1), \ldots, \text{smp}(t_n) \rangle$.

The definitions of being a type assignment, being a type context, and being a typed vocabulary set extend straightforwardly from s-types to r-types. For an r-type context $C$, let $\text{smp}(C) = \{v : \text{smp}(t) : v : t \in C\}$.

$\mathcal{S}$ is an r-type vocabulary set if and only if it is a single-valued set of r-type assignments for nonlogical constants. For such an $\mathcal{S}$, let $\text{smp}(\mathcal{S}) = \{\gamma : \text{smp}(t) : \gamma : t \in \mathcal{S}\}$.

Until further notice, $\mathcal{S}$ is a r-type vocabulary set. We will define the class of "pre-quasi-terms" based on $\mathcal{S}$ and then define a "quasi-type-assignment system" $\Rightarrow^q_{\mathcal{S}}$ and an associated $FV$ assignment; $\Rightarrow^q_{\mathcal{S}}$ determines the class of quasi-terms based on

$\mathcal{S}$ relative to an r-type context. Then we will restrict that class by defining, in two stages, an r-type assignment system based on $\mathcal{S}$.

**Definitions**      I will omit "based on $\mathcal{S}$" and the obvious simultaneous definition of $FV$. It will be convenient to simultaneously define $|\ |$ to track the depth of non-**i** quantification.

- (1) If $\nu \in \text{Var}$, then $\nu$ is a pre-quasi-term and $|\nu| = 0$.
- (2) If $\tau \in \text{dom}(\mathcal{S})$, then $\tau$ is a pre-quasi-term and $|\tau| = 0$.
- (3) If $\tau, \tau_1, \ldots, \tau_n$ are pre-quasi-terms with $n > 0$, then $\tau(\tau_1, \ldots, \tau_n)$ is a pre-quasi-term and $|\tau(\tau_1, \ldots, \tau_n)| = \max_{i \in (n)}\{|\tau|, |\tau_i|\}$.
- (4.1) $\perp$ is a pre-quasi-term and $|\perp| = 0$.
- (4.2) If $\tau_0$ and $\tau_1$ are pre-quasi-terms, then $(\tau_0 \supset \tau_1)$ is a pre-quasi-term and $|(\tau_0 \supset \tau_1)| = \max\{|\tau_0|, |\tau_1|\}$.
- (5) If $\tau$ is a pre-quasi-term, $\nu \in \text{Var}$, and $t \in \mathcal{T}^r$, then $\exists \nu : t.\tau$ is a pre-quasi-term and

$$|\exists \nu : t.\tau| = \begin{cases} |\tau| & \text{if } t = \mathbf{i}, \\ |\tau| + 1 & \text{otherwise.} \end{cases}$$

  Additional primitive logical constants may be added in the obvious ways.
- (6) If $n > 0$, $\tau$ is a pre-quasi-term, $\nu_1, \ldots, \nu_n \in FV(\tau)$ are distinct, and $t_1, \ldots, t_n \in \mathcal{T}^r$, then $(\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\tau)$ is a pre-quasi-term and $|(\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\tau)| = |\tau|$.

Consider a pre-quasi-term $\tau$ based on $\mathcal{S}$. Let $\text{smp}(\tau)$ result from replacing each occurrence of any $t \in \mathcal{T}^r$ in $\tau$ by $\text{smp}(t)$.

For a r-context $C$ and $t \in \mathcal{T}^r$, let

$$C \Rightarrow_{\mathcal{S}}^q \tau : t \quad \text{iff } \text{smp}(C) \Rightarrow_{\text{smp}(\mathcal{S})}^s \text{smp}(\tau) : \text{smp}(t).$$

Let $\tau$ be a quasi-term based on $\mathcal{S}$ and relative to $C$ if and only if for some $t \in \mathcal{T}^r$, $C \Rightarrow_{\mathcal{S}}^q \tau : t$.

For quasi-terms, being a $\beta$-redex and $\beta$-conversion of a $\beta$-redex are defined as usual; 1-step $\beta$-reduction is defined as usual. Also, being an $\eta$-redex and $\eta$-conversion of an $\eta$-redex are defined as usual; 1-step $\eta$-reduction is defined as usual. 1-step $\beta\eta$-reduction, $\beta\eta$-reduction, being $\beta\eta$-normal, and being a $\beta\eta$-normalization of a pre-quasi-term are all defined as usual.

**Theorem 6.1**      *Each quasi-term has a unique (up to choice of bound variables) $\beta\eta$-normalization.*

This follows immediately from the strong normalization and Church–Rosser theorems formulated in Section 1 by replacing quasi-terms by their simplifications.

For any quasi-term $\tau$ let $N(\tau)$ be the $\beta\eta$-normalization of $\tau$.

We are almost ready to define the ramified-type assignment system $\Rightarrow_{\mathcal{S}}^r$. We will omit mention of $\mathcal{S}$ where confusion is unlikely. This definition comes in two stages. First we define $\Rightarrow_{\mathcal{S}}^{r,\text{nf}}$.[69]

**Definition**      Let $C$ be an r-type context. There will be no need to redefine $FV$.

- ($1^r$) If $\nu : t \in C$, then $C \Rightarrow^{r,\text{nf}} \nu : t$.
- ($2^r$) If $\tau : t \in \mathcal{S}$, then $C \Rightarrow^{r,\text{nf}} \tau : t$.
- ($3^r$) If $\tau : \langle t_1, \ldots, t_n, m \rangle \in C \cup \mathcal{S}$, and for each $i \in (n)$ either $\tau_i : t_i \in C$ or $\Rightarrow^{r,\text{nf}} \tau_i : t_i$ (so $FV(\tau_i) = \{\}$), then $C \Rightarrow^{r,\text{nf}} \tau(\tau_1, \ldots, \tau_n) : \langle m \rangle$.[70]

(4.1$^r$)  $C \Rightarrow^{r,\text{nf}} \bot : \langle 1 \rangle$ and $FV(\bot) = \{\}$.

(4.2$^r$)  If $C \Rightarrow^{r,\text{nf}} \tau_i : \langle m_i \rangle$ for $i = 0, 1$, then

$$C \Rightarrow^{r,\text{nf}} (\tau_0 \supset \tau_1) : \left\langle \max_{i \in 2} m_i \right\rangle.$$

(5$^r$)  If $C, \nu : t \Rightarrow^{r,\text{nf}} \tau : \langle m \rangle$, $C \Rightarrow^{r,\text{nf}} \exists \nu : t.\tau : \langle \max(m, \text{ord}(t) + 1) \rangle$.

(6$^r$)  If $n > 0$, $C, \nu_1 : t_1, \ldots, \nu_n : t_n \Rightarrow^{r,\text{nf}} \tau : \langle m \rangle$ for distinct $\nu_1, \ldots, \nu_n \in FV(\tau)$, and $(\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\tau)$ is not an $\eta$-redex, then

$$C \Rightarrow^{r,\text{nf}} (\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\tau) : \left\langle t_1, \ldots, t_n, \max_{i \in (n)} \{m, \text{ord}(t_i) + 1\} \right\rangle.$$

Note that if $C \Rightarrow^{r,\text{nf}} \tau : t$, then $\tau$ is $\beta\eta$-normal.

Define $\Rightarrow^r_{\mathcal{S}}$ thus:

(7$^r$)  If $\tau$ is a quasi-term and $C \Rightarrow^{\text{nf}} N(\tau) : t$, then $C \Rightarrow^r \tau : t$;

So $\Rightarrow^r_{\mathcal{S}}$ determines its corresponding ramified-typed language thus:

$$L^r(\mathcal{S}) = \left\{ \langle C, \tau \rangle : \text{ for some } t \in T^r, \ C \Rightarrow^r_{\mathcal{S}} \tau : t \right\}.$$

We can define being a term of $L^r$, and being a formula of $L^r$, both relative to an r-type context $C$, in the obvious ways.

Note that the type, and thus order, of a term of $L^r$ relative to $C$ is determined by the type of its $\beta\eta$-normalization. We needed to have the class of quasi-terms and the operation of $\beta\eta$-normalization on quasi-terms available before defining $\Rightarrow^r_{\mathcal{S}}$; we provided that by defining the class of pre-quasi-terms.

Observations from Section 1 regarding $\Rightarrow^s$ carry over to $\Rightarrow^r$.

Definition (3$^r$) captures one feature of Laan and Nederpelt's reading of [55] (see [25], especially their definition of $FV$ on p. 246; they refer to [55, *9.14–15]). But one might wonder why we have not replaced (3$^r$) by this more general clause: if $\tau : \langle t_1, \ldots, t_n, m \rangle \in C \cup \mathcal{S}$, and for each $i \in (n)$ $C \Rightarrow^{r,\text{nf}} \tau_i : t_i$, then $C \Rightarrow^{r,\text{nf}} \tau(\tau_1, \ldots, \tau_n) : \langle m \rangle$. For an example of the difficulties to which this change would lead, consider $\iota : \mathbf{i}, \pi : \langle \mathbf{i}, 1 \rangle \in \mathcal{S}$ and $C = \{\nu : \langle\langle \mathbf{i}, 1 \rangle, 2 \rangle, \mu : \langle\langle 2 \rangle, 3 \rangle\}$. Under the proposed more general definition we would have $C \Rightarrow^{r,\text{nf}} \mu(\nu(\pi)) : \langle 3 \rangle$. For $\rho \equiv (\lambda \mu' : \langle \mathbf{i}, 1 \rangle.\mu'(\iota))$, $\Rightarrow^{r,\text{nf}} \rho : \langle\langle \mathbf{i}, 1 \rangle, 2 \rangle$; so we want to have

$$\mu : \langle\langle\langle \mathbf{i}, 1 \rangle, 2 \rangle, 3 \rangle \Rightarrow^{r,\text{nf}} \mu(\nu(\pi))[\nu := \rho] : \langle 3 \rangle.$$

But $\mu(\rho(\pi))$ $\beta$-reduces to $\mu(\pi(\iota))$, which is not a term of $L^r$ relative to $\{\mu : \langle\langle 2 \rangle, 3 \rangle\}$.[71]

The following definitions and lemmas are lifted from [21, Section 6]; # is a natural summation for ordinals.[72]

**Definitions**  If $\tau \in \text{Var}$, $\|\tau\| = 0$; if $\tau : \mathbf{i} \in \mathcal{S}$, $\|\tau\| = 0$; if $\tau$ is a predicate-constant in $\mathcal{S}$, $\|\tau\| = 1$; $\|\bot\| = 0$:

$$\|\tau_0(\tau_1, \ldots, \tau_n)\| = \|\tau_0\| \# \|\tau_1\| \# \cdots \# \|\tau_n\| \# 1,$$

$$\|(\varphi_0 \supset \varphi_1)\| = \|\varphi_0\| \# \|\varphi_1\| \# 1,$$

$$\|\exists \nu : t.\varphi\| = \|\varphi\| \# \omega^{\text{ord}(t)},$$

$$\|(\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\varphi)\| = \|\varphi\| \# \omega^m,$$

for $m = \max_{i \in (n)} \text{ord}(t_i)$. The following should clarify the idea behind this indexing function.

Let a quantifier prefix of the form $\exists v : t$ have order $\text{ord}(t) + 1$.[73] Let a $\lambda$-prefix $\lambda v_1 : t_1, \ldots, v_n : t_n$ have order $\max_{i \in (n)} \text{ord}(t_i) + 1$. Consider a term $\sigma$ of $L$ and $m > 0$. If $m > 1$, let $\$_m(\sigma)$ equal the number of occurrences in $\sigma$ of quantifier prefixes and $\lambda$-prefixes of order $m$. (Choice of a type context relative to which $\sigma$ is a term of $L$ is irrelevant to $\$_m(\sigma)$.) It will be convenient to define $\$_1(\sigma)$ as follows: let $\$_1^-(\sigma)$ equal the number of occurrences in $\sigma$ of quantifier prefixes and $\lambda$-prefixes of order 1, let $\$_\supset(\sigma)$ equal the number of occurrences of $\supset$ in $\sigma$, and let $\$_{\text{pc}}(\sigma)$ equal the number of occurrences of predicate constants in $\sigma$; then $\$_1(\sigma) = \$_1^-(\sigma) + \$_\supset(\sigma) + \$_{\text{pc}}(\sigma)$.

**Observation**     If $m > 0$ is the maximum of the orders of a quantifier prefix or $\lambda$-prefix in a term $\tau$ of $L$ (with the understanding that if there are no quantifier prefixes or $\lambda$-prefixes in $\tau$, then $m = 1$), then

$$\|\tau\| = \omega^{m-1} \cdot \$_m(\tau) + \cdots + \omega^0 \cdot \$_1(\tau).$$

Thus if $C \Rightarrow \tau : \langle t_1, \ldots, t_n, m \rangle$, then $\|\tau\| < \omega^m$.

**Lemma**     If $C \Rightarrow \tau : t$, then every quantifier prefix or $\lambda$-prefix occurring in $\tau$ has order $\leq \text{ord}(t)$.

**Lemma**     If $\sigma$ is a term of $L$ relative to $C \cup \{v_i : t_i\}_{i \in (n)}$, for each $i \in (n)$, $\text{ord}(t_i) < m$, and $C \Rightarrow \tau_i : t_i$, then $\$_m(\sigma[\vec{v} := \vec{\tau}]) = \$_m(\sigma)$.

**Lemma**     If $\tau \to_{\beta\eta} \tau'$, then $\|\tau\| > \|\tau'\|$.

**Observations**     (1) If $\tau_0(\tau_1, \ldots, \tau_n)$ is a term of $L$, then for each $i \in n + 1$, $\|\tau_i\| < \|\tau_0(\tau_1, \ldots, \tau_n)\|$. (2) If $(\varphi_0 \supset \varphi_1)$ is a term of $L$, then $\|\varphi_i\| < \|\varphi_0 \supset \varphi_1\|$ for $i \in 2$. (3) If $\exists v : t.\varphi$ is a term of $L$ relative to $C$ and $C \Rightarrow \tau : t$, then $\|\varphi[v := \tau]\| < \|\exists v : t.\varphi\|$.

The previous observations insure that truth for a sentence $\varphi$ can be characterized by induction on $\|\varphi\|$ treating sentences of the form $\exists v : t.\psi$ substitutionally if $\text{ord}(t) \neq 0$. This secures our truth-conditional construal of a substitutional interpretation of such quantification.[74]

We now turn to infinitary-type assignment systems and their languages. These systems will use quantification only over individuals and use infinitary disjunction (and if one wishes, conjunction) in place of other quantification. To follow the format used for the finitary case, we start with simple-type assignment systems.

Let $\kappa$ be an uncountable cardinal. Assume that $\text{card}(\text{Var}) = \kappa$. We will require of an s-type context $C$ that $\text{card}(\text{Var} - \text{dom}(C)) = \kappa$. Let $\mathcal{S}$ be a simply typed vocabulary set with $\text{card}(\mathcal{S}) < \kappa$. The simple-type assignment system $\Rightarrow_{\mathcal{S}}^{s\kappa}$ adds "levels" to types.

**Definition**     We will omit mention of $\mathcal{S}$ and all but one clause of the obvious simultaneous definition of $FV$.

(1$^\kappa$)  If $v : t \in C$, then $C \Rightarrow^{s\kappa} v : 0, t$.
(2$^\kappa$)  If $\tau : t \in \mathcal{S}$, then $C \Rightarrow \tau : 0, t$.
(3$^\kappa$)  If $C \Rightarrow^{s\kappa} \tau : l, \langle t_1, \ldots, t_n \rangle$ with $n > 0$ and $C \Rightarrow^{s\kappa} \tau_i : l_i, t_i$ for each $i \in (n)$, then $C \Rightarrow \tau(\tau_1, \ldots, \tau_n) : \max_{i \in (n)}\{l, l_i\}, \langle \rangle$.
(4.1$^\kappa$)  $C \Rightarrow^{s\kappa} \bot : 0, \langle \rangle$.
(4.2$^\kappa$)  If $C \Rightarrow^{s\kappa} \tau_i : l_i, \langle \rangle$ for $i = 0, 1$, then

$$C \Rightarrow^{s\kappa} (\tau_0 \supset \tau_1) : \max\{l_0, l_1\}, \langle \rangle.$$

$(5.1^\kappa)$ If $C, \nu : t \Rightarrow^{s\kappa} \tau : l, \langle\rangle$ (with $t \in \mathcal{T}^s$), then $C \Rightarrow^{s\kappa} \exists \nu : \mathbf{i}.\tau : l, \langle\rangle$.

$(5.2^\kappa)$ If $l \in \omega$, $\kappa'$ is an infinite cardinal and $\kappa' < \kappa$, for each $i < \kappa'$, $C \Rightarrow^{s\kappa} \tau_i : l_i, \langle\rangle$, $l = \max_{i<\kappa'} l_i$, and $\mathrm{Card}(\mathrm{Var} - \bigcup_{i<\kappa'} FV(\tau_i)) = \kappa$, then $C \Rightarrow^{s\kappa} (\bigvee_{i<\kappa'} \tau_i) : l + 1, \langle\rangle$, and $FV(\bigvee_{i<\kappa'} \tau_i) = \bigcup_{i<\kappa'} FV(\tau_i)$.

Define other connectives, including $(\bigwedge_{i<\kappa'} \tau_i)$ and $\forall$ restricted to $\mathbf{i}$, as usual, or add them as primitives.

$(6^\kappa)$ If $n > 0$, $C, \nu_1 : t_1, \ldots, \nu_n : t_n \Rightarrow \tau : l, \langle\rangle$ (with $t_i \in \mathcal{T}^s$ for $i \in (n)$) for distinct $\nu_1, \ldots, \nu_n \in FV(\tau)$, then

$$C \Rightarrow (\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\tau) : l, \langle t_1, \ldots, t_n \rangle.$$

Other notions regarding s-type assignment systems carry over to $\Rightarrow^{s\kappa}$; in particular we have this infinitary language:

$$L^{s\kappa}(\mathcal{S}) = \{\langle C, \tau \rangle : \text{ for some } l \in \omega \text{ and s-type } t, \ C \Rightarrow^{s\kappa}_{\mathcal{S}} \tau : l, t\}.$$

Informally, level measures the depth of infinitary disjunction. Note that if for $i \in 2$, $C_i \Rightarrow^{s\kappa} \tau : l_i, t_i$, then $l_0 = l_1$. Thus for any term of $L^{s\kappa}$ we can let $|\tau|^\infty =$ the $l$ such that for some $C$ and $t$, $C \Rightarrow^{s\kappa} \tau : l, t$.

Let an $\infty$-term be one of the form $(\bigvee_{i<\kappa'} \tau_i)$. $\langle \sigma; \nu_1, \ldots, \nu_n; \sigma_1, \ldots, \sigma_n \rangle$ is an outer structure for a term $\tau$ in $L^{s\kappa}$ if and only if $\sigma$ contains no $\infty$-subterm, $\nu_1, \ldots, \nu_n \in \mathrm{Var} - FV(\tau)$ are distinct, $\sigma_1, \ldots, \sigma_n$ are $\infty$-terms, and $\tau$ is $\sigma[\nu_1, \ldots, \nu_n := \sigma_1, \ldots, \sigma_n]$. Clearly every term of $L^{s\kappa}$ has an outer structure that is unique up to choice of $\nu_1, \ldots, \nu_n$ and the ordering of its maximal $\infty$-subterms.

$\beta$-conversion and $\eta$-conversion are defined as usual for terms of $L^{s\kappa}$. But $\to_{\beta\eta}$, which should be read as "one level" rather than "one step" $\beta\eta$-reduction, must be defined with care. If $\tau$ is a $\beta$-redex and $\beta$-converts to $\tau'$, let $\tau \to_{\beta\eta} \tau'$. If $\tau$ is an $\eta$-redex and $\eta$-converts to $\tau'$, let $\tau \to_{\beta\eta} \tau'$. If $\tau$ is $(\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\sigma)$ and $\sigma \to_{\beta\eta} \sigma'$, then $\tau \to_{\beta\eta} (\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\sigma')$. If $\tau$ is $\sigma(\tau_1, \ldots, \tau_n)$ and $\sigma \to_{\beta\eta} \sigma'$, then $\tau \to_{\beta\eta} \sigma'(\tau_1, \ldots, \tau_n)$; if $\tau$ is $\sigma(\tau_1, \ldots, \tau_n)$, $j \in (n)$, $\tau_j \to_{\beta\eta} \tau'_j$, and for each $i \in (n) - \{j\}$ $\tau_i$ is $\tau'_i$, then let $\tau \to_{\beta\eta} \sigma(\tau'_1, \ldots, \tau'_n)$; the clauses for $\supset$ and $\exists \nu : \mathbf{i}$ are straightforward. If $\tau$ is $(\bigvee_{i<\kappa'} \tau_i)$, for each $i < \kappa$ either $\tau_i \to_{\beta\eta} \tau'_i$ or $\tau'_i$ is $\tau_i$, and for some such $i$ $\tau_i \to_{\beta\eta} \tau'_i$, then $\tau \to_{\beta\eta} (\bigvee_{i<\kappa'} \tau'_i)$.

Define $\twoheadrightarrow_{\beta\eta}$ for terms of $L^{s\kappa}$ as usual. Define $\beta\eta$-normality and being a $\beta\eta$-normalization of a term as usual.

**Theorem 6.2**     *Every term of $L^{s\kappa}$ has a unique $\beta\eta$-normalization.*

**Proof**     This is by induction on level. Fix an outer structure

$$\langle \sigma; \nu_1, \ldots, \nu_n; \sigma_1, \ldots, \sigma_n \rangle$$

for $\tau$. If $n = 0$, $|\tau|^\infty = 0$; use Theorem 6.1. Assume that $n > 0$. Consider $j \in (n)$; let $\sigma_j$ be $(\bigvee_{i<\kappa_j} \tau_{j,i})$; for each $i < \kappa_j$ by induction fix $\tau'_{j,i}$ to be the $\beta\eta$-normalization of $\tau_{j,i}$; let $\sigma'_j$ be $(\bigvee_{i<\kappa_j} \tau'_{j,i})$. Then $\sigma'_j$ is the $\beta\eta$-normalization of $\sigma_j$. Let $\sigma'$ be the $\beta\eta$-normalization of $\sigma$. So $\sigma'[\nu_1, \ldots, \nu_n; \sigma'_1, \ldots, \sigma'_n]$ is the $\beta\eta$-normalization of $\tau$. Uniqueness follows by the usual argument.     $\square$

Next we have the infinitary ramified-type assignment, which will follow the pattern we set under the finitary ramified case. Assume that $\mathcal{S}$ is a ramified vocabulary set with $\mathrm{card}(\mathcal{S}) < \kappa$.

**Definitions**     First, we define being a $\kappa$-pre-quasiterm based on $\mathcal{S}$. Again, we omit mention of $\mathcal{S}$ and the clauses defining $FV$ for these terms.

(1*) If $\nu \in \text{Var}$, then $\nu$ is a $\kappa$-pre-quasi-term.

(2*) If $\tau \in \text{dom}(\mathcal{S})$, then $\tau$ is a $\kappa$-pre-quasi-term.

(3*) If $\tau, \tau_1, \ldots, \tau_n$ are $\kappa$-pre-quasi-terms with $n > 0$, then $\tau(\tau_1, \ldots, \tau_n)$ is a $\kappa$-pre-quasi-term.

(4.1*) $\perp$ is a $\kappa$-pre-quasi-term.

(4.2*) If $\tau_0$ and $\tau_1$ are $\kappa$-pre-quasi-terms then $(\tau_0 \supset \tau_1)$ is a $\kappa$-pre-quasi-term.

(5.1*) If $\tau$ is a $\kappa$-pre-quasi-term, $\nu \in \text{Var}$, and $t \in \mathcal{T}^r$, then $\exists \nu : \mathbf{i}.\tau$ is a $\kappa$-pre-quasi-term.

(5.2*) If $m \in \omega$, $\kappa'$ is an infinite cardinal and $\kappa' < \kappa$, for each $i < \kappa'$, $\tau_i$ is a $\kappa$-pre-quasi-term, and $\text{Card}(\text{Var} - \bigcup_{i<\kappa'} FV(\tau_i)) = \kappa$, then $(\bigvee_{i<\kappa'} \tau_i)$ is a $\kappa$-pre-quasi-term.

Additional primitive logical constants may be added in the obvious ways.

(6*) If $n > 0$, $\tau$ is a $\kappa$-pre-quasi-term, $t_1, \ldots, t_n \in \mathcal{T}^r$, and $\nu_1, \ldots, \nu_n \in FV(\tau)$ are distinct, then $(\lambda \nu_1 : t_1, \ldots, \nu_n : t_n.\tau)$ is a $\kappa$-pre-quasi-term.

For any $\kappa$-pre-quasi-term $\tau$, form $\text{smp}(\tau)$ by replacing each occurrence of any $t \in \mathcal{T}^r$ in $\tau$ by $\text{smp}(t)$.

For an r-context $C$ and $t \in \mathcal{T}^r$, let

$$C \Rightarrow_{\mathcal{S}}^{q} \tau : t \quad \text{iff} \quad \text{smp}(C) \Rightarrow_{\mathcal{S}}^{s} \text{smp}(\tau) : \text{smp}(t).$$

Let $\tau$ be a $\kappa$-quasi-term based on $\mathcal{S}$ and relative to $C$ if and only if for some $t \in \mathcal{T}^r$, $C \Rightarrow_{\mathcal{S}}^{q} \tau : t$. Clearly for such a $\tau$, $FV(\tau) = FV(\text{smp}(\tau))$.

For $\kappa$-quasi-terms, define being a $\beta$-redex and $\beta$-conversion of a $\beta$-redex as usual; Define a 1-level reduction $\rightarrow_{\beta\eta}$ as in the simple case. Define $\beta\eta$-reduction, being $\beta\eta$-normal, and being a $\beta\eta$-normalization of a $\kappa$-quasi-term as usual.

**Theorem 6.3**     *Each $\kappa$-quasi-term has a unique (up to choice of bound variables) $\beta\eta$-normalization.*

This follows immediately from Theorem 6.2 by replacing quasi-terms by their simplifications.

For any $\kappa$-quasi-term $\tau$ let $N(\tau)$ be the $\beta\eta$-normalization of $\tau$.

We are almost ready to define the r-type assignment system $\Rightarrow_{\mathcal{S}}^{r\kappa}$. We will omit mention of $\mathcal{S}$. This definition comes in two stages. First we define $\Rightarrow_{\mathcal{S}}^{r\kappa,\text{nf}}$.

**Definition**     Let $C$ be an r-type context.

(1$^{r\kappa}$) If $\nu : t \in C$, then $C \Rightarrow^{r\kappa,\text{nf}} \nu : 0, t$.

(2$^{r\kappa}$) If $\tau : t \in \mathcal{S}$, then $C \Rightarrow^{r\kappa,\text{nf}} \tau : 0, t$.

(3$^{r\kappa}$) If $\tau : \langle t_1, \ldots, t_n, m \rangle \in C \cup \mathcal{S}$ and for each $i \in (n)$ either $\tau_i : t_i \in C$ and $l_i = 0$ or $\Rightarrow^{r\kappa,\text{nf}} \tau_i : l_i, t_i$, then $C \Rightarrow^{r,\text{nf}} \tau(\tau_1, \ldots, \tau_n) : \max_{i \in (n)} l_i, \langle m \rangle$.

(4.1$^{r\kappa}$) $C \Rightarrow^{r\kappa,\text{nf}} \perp : 0, \langle 1 \rangle$.

(4.2$^{r\kappa}$) If $C \Rightarrow^{r\kappa,\text{nf}} \tau_i : l_i, \langle m_i \rangle$ for $i = 0, 1$, then

$$C \Rightarrow^{r\kappa,\text{nf}} (\tau_0 \supset \tau_1) : \max\{l_0, l_1\}, \langle \max_{i \in 2} m_i \rangle.$$

(5.1$^{r\kappa}$) If $C, \nu : t \Rightarrow^{r\kappa,\text{nf}} \tau : l, \langle m \rangle$ (with $t \in \mathcal{T}^s$), then

$$C \Rightarrow^{r\kappa,\text{nf}} \exists \nu : \mathbf{i}.\tau : l, \langle m \rangle.$$

($5.2^{r\kappa}$) If $l \in \omega$, $\kappa'$ is an infinite cardinal and $\kappa' < \kappa$, for each $i < \kappa'$, $C \Rightarrow^{r\kappa,\mathrm{nf}} \tau_i : l_i, \langle m_i \rangle$, $l = \max_{i<\kappa'} l_i$, $m = \max_{i<\kappa'} m_i$, and $\mathrm{Card}(\mathrm{Var} - \bigcup_{i<\kappa'} FV(\tau_i)) = \kappa$, then (.1) if for some $v \in \mathrm{Var} - \mathrm{dom}(C)$, $\sigma$, $k$, and $s$, $\mathrm{ord}(s) < m$, $C, v : s \Rightarrow^{r\kappa,\mathrm{nf}} \sigma : k, \langle m \rangle$, and for every $i < \kappa'$ there are a $\sigma_i$ and $l'_i$ so that $\tau_i$ is $\sigma[v := \sigma_i]$ and $C \Rightarrow^{r\kappa,\mathrm{nf}} \sigma_i : l'_i, s$, then $C \Rightarrow^{\mathrm{nf},s\kappa} (\bigvee_{i<\kappa'} \tau_i) : l + 1, \langle m \rangle$; (.2) otherwise $C \Rightarrow^{\mathrm{nf},s\kappa} (\bigvee_{i<\kappa'} \tau_i) : l + 1, \langle m + 1 \rangle$.

Define other connectives, including $(\bigwedge_{i<\kappa'} \tau_i)$ and $\forall$ restricted to $\mathbf{i}$, as usual, or add them as primitives.

($6^{r\kappa}$) If $n > 0$, $C, v_1 : t_1, \ldots, v_n : t_n \Rightarrow^{r\kappa,\mathrm{nf}} \tau : l, \langle m \rangle$ for distinct $v_1, \ldots, v_n \in FV(\tau)$ and $(\lambda v_1 : t_1, \ldots, v_n : t_n.\tau)$ is not an $\eta$-redex, then

$$C \Rightarrow^{r\kappa,\mathrm{nf}} (\lambda v_1 : t_1, \ldots, v_n : t_n.\tau) : l, \langle t_1, \ldots, t_n, \max_{i \in (n)} \{m, \mathrm{ord}(t_i) + 1\} \rangle.$$

The case division in ($5.2^{r\kappa}$) yields the following.

**Observation** If $C \Rightarrow^{r\kappa,\mathrm{nf}} (\bigvee_{i<\kappa'} \tau_i) : l_0 + 1, \langle m + 1 \rangle$, $C \Rightarrow^{r\kappa,\mathrm{nf}} (\bigvee_{i<\kappa'} \sigma_i) : l_1 + 1, \langle m \rangle$, and $\{\theta_i\}_{i<\kappa'} = \{\tau_i\}_{i<\kappa'} \cup \{\sigma_i\}_{i<\kappa'}$, then $C \Rightarrow^{r\kappa,\mathrm{nf}} (\bigvee_{i<\kappa'} \theta_i) : \max\{l_0, \ell 1\} + 1, \langle m + 1 \rangle$.

Define $\Rightarrow^{r\kappa}_{\mathscr{S}}$ thus:

($7^{r\kappa}$) if $\tau$ is a $\kappa$-quasi-term and $C \Rightarrow^{r\kappa,\mathrm{nf}} N(\tau) : t$, then $C \Rightarrow^{r\kappa} \tau : t$.

$\Rightarrow^{r\kappa}_{\mathscr{S}}$ determines its corresponding ramified-typed language thus:

$$L^{r\kappa}(\mathscr{S}) = \left\{ \langle C, \tau \rangle : \text{for some } \tau \in T^r \text{ and } l \in \omega, \ C \Rightarrow^{r\kappa}_{\mathscr{S}} \tau : l, t \right\}.$$

We can define being a term of $L^{r\kappa}$, and being a formula of $L^{r\kappa}$, both relative to $C$, in the obvious ways.

Observations from Section 1 regarding $\Rightarrow^{s\kappa}$ carried over to $\Rightarrow^{r\kappa}$.

Let $\kappa = \max\{\omega, \mathrm{card}(\mathscr{S})\}^{+}$.[75] For a ramified-type context $C$ and $t \in \mathscr{T}^r$, let $\kappa_{C,t} = \mathrm{card}(S_{C,t})$ for $S_{C,t} = \{\sigma : C \Rightarrow^{r\kappa} \sigma : t\}$; note that $\kappa_{C,t} < \kappa$. Fix a $\kappa_{C,t}$-ordering $\langle \sigma_i \rangle_{i<\kappa_{C,t}}$ of $S_{C,t}$.

We will now define a canonical translation trn from $\Rightarrow^r$ into $\Rightarrow^{r\kappa}$. This definition is inductive, using the ordinal indexing $\| \ \|$. All clauses except that for quantification over types other than $\mathbf{i}$ are homomorphic. If $C, v : t \Rightarrow^{r\kappa} \varphi : \langle m \rangle$, let

$$\mathrm{trn}(\exists v : t.\varphi) = \bigvee_{i<\kappa_{C,t}} \mathrm{trn}(\varphi[v := \sigma_i]).$$

To see that trn is well defined, note that, in this last clause, for each $i < \kappa_{C,t}$, $\|\tau[v := \sigma_i]\| < \|\exists v : t.\tau\|$.

**Observation** If $C \Rightarrow^r \tau : t$, then $C \Rightarrow^{r\kappa} \mathrm{trn}(\tau) : |\tau|, t$, that is, $|\tau| = |\mathrm{trn}(\tau)|^\infty$. Proof is by induction on the construction of $\tau$. Note that if $\tau$ is $\exists v : t.\varphi$, the argument uses the case division in clause ($5.2$)$^{r\kappa}$.

Informally, trn preserves meaning. Making this precise would require model-theoretic discussion that it seems best to skip in this paper.

I do not mean to suggest that Russell's hints at a substitutional interpretation of quantification over non-$\mathbf{i}$-types were hints at the strongest construal of such an interpretation. But were we, in Russell's place, to adopt the strongest construal, trn would allow us to circumvent the Quinean objection to the repudiation of propositions that accompanied Russell's conversion to the MRT: for $\tau$ any sentence of $L^r$, we could

construe a sentence of the form "S judges that $\tau$" as a shorthand for "S judges that trn($\tau$)," and similarly for other propositional attitude verbs. The step from $L^r$ to $L^{r\kappa}$ would then be a step toward optimality: unlike sentences of $L^r$, propositions would contain no quantification of nonzero order.

On the suggestion we are considering, for S to judge that $\tau$, presumably S would have to understand the proposition signified by trn($\tau$), which would be infinitely long if $\tau$ is of order greater than one. Of course for many such sentences $\tau$, judging that $\tau$ might be beyond human capacities. But there are higher-order choices of $\tau$ that figure in the Russellian foundations of mathematics. The approach we are considering would require that human beings be able to understand trn($\tau$) for such sentences.

Could Russell have adopted the strongest construal consistently with his other views? This question reduces to the following: could he have thought it possible to understand a proposition signified by an infinitely long formula? In [44, Section 141], Russell asks whether there are infinite "unities"—understanding a unity to be a proposition. He denies that we can know such propositions but leaves open the possibility that they exist. But his infinite unities are infinite by virtue of infinite complexity, which might better be called infinite depth: "it is possible to find a constituent unity, which again contains a constituent unity, and so on without end" (p. 145). In this sense, a proposition whose structure is reflected by an infinite formula of $L^{r\kappa}$ should not count as infinitely complex, as is shown by the fact that for any term $\tau$ of $L^{r\kappa}$, $|\tau|^\infty < \omega$. Still, it is hard to believe that this twist has much promise: even if we do not consider an infinite sentence $\tau$ of $L^{r\kappa}$ to be infinitely complex, it can contain infinitely many nonlogical constants. Russell's principle of acquaintance is usually understood to require that if one understands $\tau$, one is acquainted with each of the genuine constituents of trn($\tau$), which could be infinite in number. If this understanding is correct, the suggestion on the table runs into conflict with the plausible idea that a human being can only be acquainted with finitely many entities.

Although I know of no reconstruction of Russell's repudiation of propositions that does better, in the end it is far from clear that translation into $L^{r\kappa}$ supports a satisfactory interpretation of the repudiation of propositions. I am inclined to conclude that there is an unresolvable tension between the repudiation and Russell's thesis that human understanding must be finitary.

Setting aside the worry posed by our finite capacity for acquaintance, where would a language of the form $L^{r\kappa}$ leave the Russellian who wants to live with the MRT? As a matter of surface syntax, the MRT requires us to consider phrases like "S judges that" to be sentential (or more generally, formula) operators that form sentences (or more generally, formulas); furthermore, these sentences (formulas) are semantically atomic, with the underlying predicate (e.g., "Judges*") provided by the operator. For example, we can translate "Russell judges that Socrates is wise" as "Judges(russell, Socrates is wise)," with "Judges": $\langle i, \langle 1 \rangle \rangle$; but the semantic analysis provided by MRT requires that this be refined to "Judges*(russell,socrates,wisdom)" if we are to get closer to the underlying logical form. What about a sentence of the superficial form "Judges(russell, $\varphi$)" if $\varphi$ is infinitary, perhaps containing infinitely many individual constants? To apply the MRT to such a sentence would require going beyond $L^{r\kappa}$. It would seem to require infinite r-types and predicate constants with infinitely many places. So even a language of the form $L^{r\kappa}$ is not optimal for Russellian

purposes. Describing an optimal such language in detail would require a careful formulation of the MRT, a project that I will not tackle here.[76]

One last point regarding the above discussion. Why not go all the way and eliminate quantification over individuals? Doing this would be contrary to some straightforward remarks about such quantification. On the assumption that "Socrates is mortal" expresses an elementary judgment, Russell wrote this:

> Our judgment that all mean are mortal collects together a number of elementary judgments. It is not, however, composed of these, since (e.g.) the fact that Socrates is mortal is no part of what we assert, as may be seen by considering the fact that our assertion can be understood by a person who has never heard of Socrates. In order to understand the judgment "all men are mortal," it is not necessary to know what men there are. We must admit, therefore, as a radically new kind of judgment, such general assertions as "all men are mortal" [55, p. 45].[77]

Russell's point seems to be this: understanding a judgement with content signified by

$$\bigwedge\left\{\left(\text{Man}(\iota) \supset \text{Mortal}(\iota)\right) : \iota : \mathbf{i} \in \mathscr{S}\right\}$$

would require acquaintance with every individual signified by a member of $\mathscr{S}$, and understanding a judgment with content signified by $\forall \nu : \mathbf{i}(\text{Man}(\nu) \supset \text{Mortal}(\nu))$ does not. Thus we should not translate the latter into the former.

## Notes

1. In [4], Chihara says of Russell's brief sketch of a typed hierarchy of classes in [44, Appendix B, pp. 523–24] that it "would now be roughly classified as a simple-type theory." This is not how I would understand type theories: the view of classes that Russell there sketched was not based on an ontology of propositions and propositional functions.

   Ramsey's *The Foundations of Mathematics* (read in 1925, published in 1926), reprinted in [40], is well known. In a paper published in Polish in 1921, Chwistek had already suggested a simple theory of types; see the translation "Antimonies of formal logic" in McCall [33, p. 343]. There was more in Chwistek [9], where he said that simple types were in some respect inadequate ("inconsistent with certain fundamental problems of Logic and Semiotics").

2. The history of Russell's thinking on paradoxes as a reason for ramification was explored by Warren Goldfarb in [18].

3. See also Chihara [4, pp. 14, 16].

4. Perhaps Myhill's phrase "type signatures" would be better (see Myhill [34]). But I will defer to the more common usage.

5. Church [6] seems to be the original source. See also Church [5].

6. See Curry, Feys, and Craig [12, pp. 315–43] for Curry's most-cited presentation. I gather that Curry's earliest presentation in print is [10]. Jonathan Seldin informed me that Curry wrote to Hilbert about external typing in 1929. Curry's original notation is no longer used. I have followed the standard contemporary notation; see [19].

7. In [55], the authors announce that what they actually put on the page will lack explicit typing and should be understood as "typical ambiguous," correctly calling this a matter of "convenience." What follows meets them halfway: occurrences of variables free in a term do not as such have types. This approach reflects the discussion of "any" and "all" in [41] (see also pp. 67–68 in Russell [46]), in which Russell construes "any" to indicate free occurrences of variables subject to a universality interpretation (an astute observation about the semantics of English).

8. Sometimes the notion of a type context is defined simply as a set of type assignments for variables, in which case what we here call a context would be called a consistent context.

9. Alternatively, add them as new primitive expressions.

10. Alternatively, add $\forall$ as a new primitive. Where its scope is clear, I will omit the period after an occurrence of a quantifier prefix.

11. This convention lets us avoid reliance on $\alpha$-conversion and $\alpha$-reduction.

12. Whitehead and Russell used the circumflex "in situ" in place of $\lambda$ prefixing. As is well known, this leads to ambiguities (see Curry [11]).

13. What Church [6] calls a $\lambda$-term is what Barendregt [2] calls a $\lambda I$-term; of course in both texts every $\lambda$-prefix contains a single variable. For a nice introduction to the $\lambda I$-calculus, see Anderson [1, Section 71].

14. Strictly speaking, the definition is by an obvious induction on the construction of $\tau$, as follows. If $\tau$ is a $\beta$-redex and $\beta$-converts to $\tau'$, $\tau \to_\beta \tau'$. If $\tau \equiv (\lambda v_1 : t_1, \ldots, v_n : t_n.\sigma)$ and $\sigma \to_\beta \sigma'$, then $\tau \to_\beta (\lambda v_1 : t_1, \ldots, v_n : t_n.\sigma')$; if $\tau \equiv \sigma(\tau_1, \ldots, \tau_n)$ and $\sigma \to_\beta \sigma'$, then $\tau \to_\beta \sigma'(\tau_1, \ldots, \tau_n)$; if $\tau \equiv \sigma(\tau_1, \ldots, \tau_n)$, $j \in (n)$, $\tau_j \to_\beta \tau'_j$, and for each $i \in (n) - \{j\}$ $\tau_i \equiv \tau'_i$, then let $\tau \to_\beta \sigma(\tau'_1, \ldots, \tau'_n)$; the remaining clauses for $\neg, \supset, \exists$, and any other primitive logical constants are straightforward.

15. The strict definition is by the obvious induction; I will omit details.

16. That is to say, if $C \Rightarrow \tau : t$ and $\tau \twoheadrightarrow_{\beta\eta} \tau'$, then $C \Rightarrow \tau' : t$ and $FV(\tau) = FV(\tau')$.

17. See, e.g., Troelstra and Schwichtenberg [54, pp. 210–12]; also, see [19].

18. See [19].

19. For such a system, see Hodes [21].

20. The following remark from Russell bears on this distinction (see "The theory of logical types" in [50, p. 221]):

> "Thus a convenient way to read $(x).\varphi x$ is "$\varphi x$ is true with all possible values of $x$." This is, however, a less accurate reading than "$\varphi x$ always," because the notion of *truth* is not part of the content of what is judged. When we judge "all men are mortal," we judge truly, but the notion of truth is not necessarily in our minds, any more that it need be when we judge "Socrates is mortal."

21. This argument bears comparison with the one that Russell found in 1907 and that led him to give up on his so-called substitutional theory and commit to a theory of types (see Linsky [30]).

22. Of course Russell did not recognize this when he wrote Appendix B to [44]. In [16], André Fuhrmann suggested that this recognition came in his letter to Frege of May 24, 1903 (see Frege [14, pp. 158–60]), in which Russell first proposed the "no class" reduction of classes to propositional functions. Fuhrmann argued that along with this recognition, Russell also, (1) in effect, recognized the falsehood of what below I call Russell's dictum (see a later note), and with this he recognized (2) the fallaciousness of his propositional paradox. Fuhrmann concluded that the purported propositional paradox played no role in motivating ramification.

23. By the above remarks, $\forall v : \langle\rangle(m(v) \supset T(v))$ would be slightly better; but the difference will not matter for what follows. To facilitate comparison with the passage quoted above, I here use $m$ and $n$ instead of Greek letters.

24. In his appendix to Volume II of the *Grundgesetze*, Frege proved what I will call his appendix theorem: "For every second-level function of one argument ... there are concepts which if taken as arguments of this function determine the same value, although not all objects falling under one of these concepts also fall under the other" (see Frege [15, p. 136]). Slightly modifying Michael Potter's formulation (see [35, p. 134]) and using $\exists^{(o)}$ for existential quantification over Fregean concepts of level 1, we can express it thus:

$$\exists^{(o)}\varphi\exists^{(o)}\psi\big(f(\varphi) = f(\psi) \;\&\; \varphi\big(f(\varphi)\big) \;\&\; \neg\psi\big(f(\psi)\big)\big).$$

This is a Fregean version of the set-theoretic fact that for any set $d$ and any $f$ : Power$(d) \to d$, $f$ is not one–one. It can be proved by a variation on the above purported paradox. (Let $\varphi(x)$ be $\exists^{(o)}\xi(x = f(\xi) \;\&\; \neg\xi(x))$. Assume $\neg\varphi(f(\varphi))$; since $f(\varphi) = f(\varphi)$, $\varphi(f(\varphi))$ follows for a contradiction. By excluded middle, $\varphi(f(\varphi))$. So we may fix a $\psi$ so that $f(\varphi) = f(\psi) \;\&\; \neg\psi(f(\varphi))$.)

In his letter to Frege of May 24, 1903, Russell clearly refers to his purported propositional paradox and to the instance of Frege's appendix theorem obtained by taking $f(\xi)$ to be $\forall v : \langle\rangle(\xi(v) \supset v)$, saying "now these difficulties have been overcome by means of the theorem in your appendix..." (see [14, p. 160]). I take Fuhrmann to have proposed that this amounts to Russell's rejection of weak RD (see [16, p. 209]).

25. What follows is close to Church's [8] reconstruction of the original paradox. In particular, (1b) on p. 517 is almost weak RD, except that Church uses a special symbol for propositional identity. For an even more important difference, see a note in Section 4.

26. The subscripted $R$ refers to Russell.

27. This was pointed out in Potter [35, pp. 132–33] and even earlier in [8].

28. One might conceive of propositions as mental entities to whose existence we only have empirical access. If so, perhaps Ramsey would have put these propositional paradoxes in his group B. But this mentalistic conception of propositions is foreign to Russell's writings. Had Ramsey, like Quine, thought that all legitimate quantification is into argument positions, he would have rejected the very sort of formal language within which

the above arguments are formulated. But Ramsey liked simple-type languages. I cannot see any grounds on which he could have considered the above arguments to rely on "empirical terms."

29. Russell also did not draw the now-standard distinction between a proposition and a formula that might signify that proposition (relative to an appropriate assignment to the variables free in that formula).

30. See Irwin [23] for some background on Aristotle's use of the Greek word standardly translated as "signify."

31. See "On the nature of truth and falsehood" in Russell [42]. Bernard Linsky has suggested that this change in Russell's view occurred in 1908.

32. For an intriguing step towards spelling out such a theory, see Wrinch [57]. I thank Bernard Linsky for this citation.

33. This example assumes that Ann and Socrates are individuals and that wisdom is a quality (and thus, in pre-MRT terms, the proposition that Socrates is wise is elementary). In several places Russell treats these as idealizing assumptions.

34. See in particular the "metaphysical paragraph" in Whitehead and Russell [55, p. 43]. For some useful discussion, see Linsky [29]. With the multigrade-relation theory, commitment to facts crowded out commitment to propositions (see the quotation below). For more on this shift, see Proops [37].

35. The relevant sentence ("They seem to be aware that this fragmenting of propositions requires a similar fragmenting of propositional functions") is omitted in the reprinted version in Martin [32]. The mentioned fragmenting presumably refers to attributions of propositional attitudes. Perhaps Church became unhappy with his lack of a textual basis for this remark. Be that as it may, Whitehead and Russell should have gone along with Church's presumption.

36. So Russell could not consistently think that qualities were propositional functions. This has not been universally appreciated. In [51], Sainsbury construes qualities to be "properties with which we are acquainted in perception" (p. 268). Perhaps this is right. But he goes on to take seriously the suggestion that Russell considered propositional functions to be properties (see pp. 285–92). And the converse inclusion has some currency (e.g., "For if we go with Russell in assuming that properties are propositional functions..."; see [1, p. 392]).

    In [28, pp. 453–54], Linsky surveys Russell's equivocal use of "property." At [55, p. 57], Russell uses "property" to apply to propositional functions; elsewhere it is pretty clear that he does not (see [49, p. 94]).

37. See, for example, Russell [44, p. 356].

38. Also see [55, p. 39]: "A function is not a well-defined function unless all its values are already well-defined." See Hylton [22, p. 289, n. 7] and Linsky [28, p. 448].

39. Again, such examples are to be understood as based on idealizing assumptions about what entities are genuine individuals or genuine universals, or (speaking outside of the MRT) what propositions are genuinely elementary.

40. "...genuine constituents in the sense that they do not disappear on analysis" [55, p. 51]. In [44], Russell allows that a proposition can be a constituent of another proposition. I take it that, after repudiating propositions, Russell would not allow that one proposition could be a genuine constituent of another.

41. In [28], Linsky independently made a similar suggestion: "Atomic predicates...stand directly for both a universal *and* a propositional function" (p. 454). I think it clearer not to use "stand for" for the latter role. He goes on to remark, with complete justice, "Russell did not systematically work out the place of universals in his logic."

42. Even so, the option of having universals take only simple types in the metaphysical hierarchy would remain open; in fact, I think that it makes more sense than saying that they take ramified types.

43. From [55, p. 44]: "We will give the name of 'a *complex*' to any such object as 'a standing in relation $R$ to $b$' or 'a having quality $q$,' or 'a and b and c standing in relation $S$.' Broadly speaking, a complex is anything which occurs in the universe and is not simple." Russell and Whitehead clearly understand these to exist if and only if $a$ stands in relation $R$ to $b$, and so forth.

44. In Russell [47]; see [46, p. 208].

45. I take qualities in Russell's sense to be 1-place relations.

46. If $\tau$ is closed, we may take **a** to be empty and write $[\tau]$ for $[\tau]^{\mathbf{a}}$.

47. One might object that in 3.2(4.1) and (5), I use the concept of truth in an illegitimate way and that the following would be improvements: $(4.1')$ $[(\tau_0 \supset \tau_1)]^{\mathbf{a}}$ is the result of applying the material-conditional propositional function to $[\tau_0]^{\mathbf{a}}$ and $[\tau_1]^{\mathbf{a}}$ in that order; $(5')$ $[\exists \nu : t.\tau]^{\mathbf{a}}$ is the result of applying the existential-quantifier-of-type-$t$ propositional function to $[(\lambda \nu : t.\tau)]^{\mathbf{a}}$. But it is unclear that there is a real difference between (4.1) and $(4.1')$, or between (5) and $(5')$, especially in light of the use of "true" and "false" in the explanations of the meanings of connectives and quantifiers in [55, pp. 93, 127].

48. Russell was thinking in these terms by 1903. "There is, for each propositional function, an indefinable relation between propositions [that are values of this function] and entities, which may be expressed by saying that all the propositions have the same form, but different entities enter into them" ([44, p. 29]; see also p. 510). One's first inclination is to think of this "punching out" as a removal of constituents from a proposition. But this fits uncomfortably with Russell's metaphysics after his conversion to the MRT. Assuming that the constituents of a propostion are real, the MRT would dictate that propositions are not constutents of other propositions, in which case the arguments of a propositional function could not have propositional types. This suggests that no language of the form $L^s$ could be optimal. (Similarly for languages of the form $L^r$, to be introduced in Section 6. But a language of the form $L^{r\kappa}$, also to be introduced in Section 6, would not face this difficulty.)

49. I call this a picture rather than a theory because it leaves open in what reflection of structure consists.

50. For further discussion, see [21].

51. In Section 6 I will further consider the shape of an optimal formal language.

52. Michael Potter thought so: "If Russell had analyzed the reason for the apparent resolution of his propositional paradox by this result (viz. the "appendix theorem" mentioned in a previous note), he would surely have realized that Frege's result is itself in direct opposition to Russell's conception of propositions" ([35, p. 134]). This "opposition" would require that Russell's conception commit him to at least RD and presumably PQ as well. Furhmann apparently agreed with this claim of "opposition," though he concluded that Russell's resolution of his purported propositional paradox "heralds the demise of Russell's early theory of propositions as certain complexes—at least in the naive version that permeates his writings until and including *The Principles*" ([16, p. 210]). If this were correct, there should be evidence of this demise in Russell's writings between May 24, 1903, and his adoption of the MRT in 1908 (at the earliest). I would like to see such evidence.

53. In [8, p. 518], Church referred to his **n** and **m** as names, which I take to mean that they are nonlogical constants. But in the principle he actually used in his reconstruction of Russell's paradox, (1b) from p. 517, $n$ and $m$ are not in boldface and so presumably are variables (which the argument requires, since the use of existential elimination requires an eigenvariable). Since he missed the question of what a predicational occurrence of a term in a formula can indicate about a constituent of the proposition expressed by that formula, it seems that he was too quick to find RD plausible.

54. See Godel [17]. For example, there is evidence of a reading that requires the well-foundedness of class membership.

55. In places, Russell uses "defined" rather than "specified." I avoid that usage: we want a name to be a predicative specification of its bearer, but it would be odd to say that a name defines its bearer.

56. Let MR abbreviate "is a man in the room," and let T abbreviate "is at least as tall as." With some simplification, we could regiment "the tallest man in the room" into $L^s$ as follows:

$$\text{the } \nu : \mathbf{i}.\big(MR(\nu) \,\&\, \forall \mu : \mathbf{i}\big((MR(\mu) \,\&\, T(\nu, \mu)) \supset (\mu =^{\mathbf{i}} \nu)\big)\big).$$

57. This is from the manuscript "On Insolubilia" from 1906: "t is important to observe that the vicious-circle principle is not itself the solution of vicious-circle paradoxes, but merely the result which a theory must yield if it is to afford a solution of them" ([50, p. 205]).

58. See Goldfarb [18, pp. 24–25]. This usage goes back at least to Gödel [17], who refers to "the constructivistic (or nominalistic) standpoint" (see Benacerraf and Putnam [3, p. 456]). Gödel finds this cast of mind but says, "In the first edition of *Principia* ...the constructivistic attitude was, for the most part, abandoned..." ([3, p. 461]). Note that in this context "constructivist" has nothing to do with avoiding use of excluded middle.

59. From Russell [45, p. 254].

60. From Russell [48, p. 124].

61. Apparently, no one explicitly recognized the difference between a substitutional inter-
pretation and a objectual interpretation of a quantifier until the early 1960s. To my
knowledge, the distinction was first clearly drawn in Marcus [31]. A substitutional inter-
pretation of quantification over types other than **i** would not imply that propositions and
propositional functions are linguistic expressions; in fact, on such an interpretation the
latter thesis is ill formed. On such an interpretation, a variable restricted to a type other
than **i** would not have a range of values; it would merely have a range of permissible
substitutends. Unclarity about these points has outlived Russell; see, e.g., Landini [26]
and Klement [24].

62. See Dummett [13, pp. 217–18] for discussion. It is not clear that this indicates a sub-
stitutional interpretation of such quantification. In [56], Wittgenstein's understanding of
quantification is, crucially, substitutional.

63. If we insist that the disjuncts in an infinite disjunction are well ordered, we can fix in
advance a well-ordering of such terms to drop the scare quotes on "the." But we could
as well take the disjuncts of an infinitary disjunctive proposition to be unordered.

64. On a Fregean view of the semantic values of sentences, $\exists \nu : \langle\rangle.\nu$ is intelligible with
$\exists \nu : \langle\rangle$ interpreted objectually; but on non-Fregean views, it would be less clear how
interpret this (and related) sentences.

65. Another con from the viewpoint of the Whitehead–Russell project: when applied to
ramified-type formulas, substitutional interpretation of quantification over propositional
functions makes the axioms of reducibility even less plausible than they are under an
objectual interpretation.

66. For a useful discussion, see Soames [52, Chapter 2].

67. In the proof sketched in a footnote in Section 2, $\psi$ was an arbitrary witness to an exis-
tential truth. Were we to interpret type $\langle\rangle$ to consist of sentences of $L^s(\mathcal{S})$ (for a given
$\mathcal{S}$), that instance of Frege's theorem would be false.

68. The following material on $\Rightarrow^r$ is presented with further discussion and detail in [21].

69. The superscripted nf is for "normal form."

70. Thus $\tau(\tau_1, \ldots, \tau_n)$ is not a $\beta$-redex.

71. According to the revision being considered, we would have to say of

$$\left(\lambda\mu : \langle\langle 2\rangle, 3\rangle, \nu : \langle\langle \mathbf{i}, 1\rangle, 2\rangle.\mu\big(\nu(\pi)\big)\right)$$

either that it fails to signify at all, or that it signifies a partial propositional function.
Neither is appealing. I have no idea of whether this concern motivated Whitehead and
Russell in effect to count this and similar $\lambda l$-terms as ill formed. But that ruling, captured
by $(3^r)$, suggests that $\Rightarrow^r_{\mathcal{S}}$ should not be the last word in ramified-type assignments. I
discuss this further in [21].

72. For ordinals below $\omega^\omega$, # is defined as follows; for $j \in 2$ and $\alpha_j = \sum_{i=0}^{n_j} \omega^i \cdot x_{j,i}$ for $x_{j,0}, \ldots, x_{j,n_j} \in \omega$, $\alpha_0 \# \alpha_1 = \sum_{i=0}^{n} \omega^i \cdot x_i$ where $n = \max(n_0, n_1)$, and setting $x_{j,i} = 0$ for all $n_j < i \leq n$, for all $i \leq n$, $x_i = x_{0,i} + x_{1,i}$. Note that # is associative and commutative.

73. Similarly for any other primitive quantifier constants.

74. See [21] for a detailed model-theoretic exposition of this idea.

75. Here $\alpha^+$ is the successor cardinal for a cardinal $\alpha$.

76. If a universal can only be instanced by a particular or a tuple thereof, a vocabulary set $\mathcal{S}$ need only contain predicate constants of order 1. In this unlikely case, we could have $L^{r\kappa}$ be optimal: the $\beta\eta$-normal sentences of $L^{r\kappa}$ carry no commitment to propositions or to propositional functions; not only do they contain no quantification over such shadowy entities, they also contain no predications of them either, as shown by the following.

    *Observation.* If all predicate constants in $\mathcal{S}$ are of order 1 (i.e., for any $\pi : \langle t_1, \ldots, t_n, m \rangle \in \mathcal{S}$, $t_i = \mathbf{i}$ for $i \in (n)$ and $m = 1$), $\varphi$ is a $\beta\eta$-normal sentence of $L^\kappa$, and $\tau(\tau_1, \ldots, \tau_n)$ is an occurrence of a subformula of $\varphi$, then for each $i \in (n)$ either $\tau_i : \mathbf{i} \in \mathcal{S}$ or $\tau_i \in \text{Var}$ and is bound by a prefix in $\varphi$ of the form $\exists \tau_i : \mathbf{i}$. Proof: assume otherwise. By assumption about $\mathcal{S}$, $\tau$ is not a predicate constant. If $\tau$ is a $\lambda$-term, then $\varphi$ is not $\beta$-normal, and so not $\beta\eta$-normal. If $\tau$ is a variable, it is not bound in $\varphi$ since the occurrence in question cannot be bound by an occurrence of $\exists \tau : \mathbf{i}$, but it cannot be free in $\varphi$ since $\varphi$ is a sentence.

77. There is an almost identical passage in Russell's "The theory of logical types"; see [50, p. 226].

## References

[1] Anderson, A. R., N. D. Belnap, and J. M. Dunn, *Entailment: The Logic of Relevance and Necessity, Vol. II*, Princeton University Press, Princeton, 1992. MR 1223997. 406, 408

[2] Barendregt, H. P., *The Lambda Calculus*, vol. 103 of *Studies in Logic and the Foundations of Mathematics*, North-Holland, Amsterdam, 1984. MR 0774952. 382, 406

[3] Benacerraf, P., and H. Putnam, editors, *Philosophy of Mathematics: Selected Readings*, 2nd edition, Cambridge University Press, Cambridge, 1983. Zbl 0548.03002. MR 0742474. 410

[4] Chihara, C., *Ontology and the Vicious Circle Principle*, Cornell University Press, Ithaca, N.Y., 1973. MR 0469678. 405

[5] Church, A., "A formulation of the simple theory of types," *Journal of Symbolic Logic*, vol. 5 (1940), pp. 56–68. MR 0001931. 405

[6] Church, A., *The Calculi of Lambda-Conversion*, vol. 6 of *Annals of Mathematical Studies*, Princeton University Press, Princeton, 1941. MR 0005274. 405, 406

[7] Church, A., "Comparison of Russell's resolution of the semantical antinomies with that of Tarski," *Journal of Symbolic Logic*, vol. 41 (1976), pp. 747–60; reprinted with corrections in [32], pp. 289–306. MR 0441682. 387

[8] Church, A., "Russell's theory of identity of propositions," *Philosophia Naturalis*, vol. 21 (1984), pp. 513–522. MR 0798506. 407, 410

[9] Chwistek, L., "The theory of constructive types," extracted from *Annales de la Societe Polonaise de Mathematiques*, Cracow University Press, Crakow, Poland, 1925; reprint,

University of Michigan Library, Ann Arbor, 2005. 405

[10] Curry, H. B., "Functionality in combinatory logic," *Proceedings of the National Academy of Science U.S.A.*, vol. 20 (1934), pp. 584–90. 405

[11] Curry, H. B., *Foundations of Mathematical Logic*, McGraw-Hill Book Company, New York, 1963. MR 0148529. 406

[12] Curry, H. B., R. Feys, and W. Craig, *Combinatory Logic, Vol. I*, North-Holland Publishing Company, Amsterdam, 1958. MR 0094298. 405

[13] Dummett, M., *Frege: The Philosophy of Mathematics*, Harvard University Press, Cambridge, Mass., 1991. MR 1154309. 411

[14] Frege, G., *Philosophical and Mathematical Correspondence*, edited and with an introduction by Gottfried Gabriel et al., abridged from the German edition and with a preface by Brian McGuinness, with an appendix by Philip E. B. Jourdain, University of Chicago Press, Chicago, 1980. Zbl 0502.03002. MR 0568066. 407

[15] Frege, G., *The Basic Laws of Arithmetic*, translated, edited, and introduced by Montgomery Furth, reprint of the 1964 original, University California Press, Berkeley, 1982. MR 0675436. 395, 407

[16] Fuhrmann, A., "Russell's way out of the paradox of propositions," *History and Philosophy of Logic*, vol. 23 (2002), pp. 197–213. Zbl 1034.01014. MR 1951813. DOI 10.1080/01445340210161017. 407, 410

[17] Gödel, K., "Russell's mathematical logic," pp. 125–153 in *The Philosophy of Bertrand Russell*, edited by Paul A. Schlipp, Northwestern Univ. Press, Evanston, Illinois, 1944; reprint, pp. 447–469 in *Philosophy of Mathematics: Selected Readings*, edited by P. Benacerraf and H. Putnam, Cambridge University Press, Cambridge, 1983. 410

[18] Goldfarb, W., "Russell's reasons for ramification," pp. 24–40 in *Rereading Russell*, edited by C. W. Savage and C. A. Anderson, vol. XII in *Minnesota Studies in the Philosophy of Science*, University of Minnesota Press, Minneapolis, 1989. MR 1006456. 405, 410

[19] Hindley, J. R., *Basic Simple Type Theory*, vol. 42 of *Cambridge Tracts in Theoretical Computer Science*, Cambridge University Press, Cambridge, 1997. MR 1466699. DOI 10.1017/CBO9780511608865. 381, 382, 405, 406

[20] Hodes, H. T., "A report on some ramified-type assignment systems and their model-theoretic semantics", pp. 305–36 in *The Palgrave Centenary Companion to Principia Mathematica*, edited by N. Griffin and B. Linsky, Palgrave Macmillan, London, 2013. 381

[21] Hodes, H. T., "Some ramified-type assignment systems and their logics," in preparation. 399, 406, 410, 411, 412

[22] Hylton, P., *Russell, Idealism, and the Emergence of Analytic Philosophy*, Clarendon Press, Oxford University Press, Oxford, 1990. 389, 408

[23] Irwin, T., "Aristotle's concept of signification," pp. 241–66 in *Language and Logos*, edited by M. Schofield and M. Nussbaum, Cambridge University Press, Cambridge, 1982. 408

[24] Klement, K., "The functions of Russell's no class theory," *Review of Symbolic Logic*, vol. 3 (2010), pp. 633–64. MR 2759283. DOI 10.1017/S1755020310000225. 411

[25] Laan, T., and R. Nederpelt, "A modern elaboration of the ramified theory of types," *Studia Logica*, vol. 57 (1996), pp. 243–78. MR 1416736. DOI 10.1007/BF00370835. 399

[26] Landini, G., *Russell's Hidden Substitutional Theory*, Oxford University Press, New York, 1998. Zbl 0933.03002. MR 1694614. 411

[27] Levine, J., "Review of Landini's Russell's hidden substitutional theory," *Philosophical Review*, vol. 110 (2001), pp. 138–41.

[28] Linsky, B., "Propositional functions and universals in *Principia Mathematica*," *Australasian Journal of Philosophy*, vol. 66 (1988), pp. 447–60. 388, 408, 409

[29] Linsky, B., *Russell's Metaphysical Logic*, vol. 101 of *CSLI Lecture Notes*, CSLI Publications, Stanford, 1999. MR 1833170. 408

[30] Linsky, B., "The substitutional paradox in Russell's 1907 letter to Hawtrey," *Russell: The Journal of Bertrand Russell Studies*, vol. 22 (2002), pp. 47–55. 407

[31] Marcus, R. B., "Interpreting quantifiers," *Inquiry*, vol. 5 (1962), pp. 252–59. 411

[32] Martin, R. L., editor, *Recent Essays on Truth and the Liar Paradox*, Clarendon Press, Oxford University Press, Oxford, 1984. Zbl 0623.03001. 408, 412

[33] McCall, S., editor, *Polish Logic (1920–1939)*, translated by B. Gruchman et al., Clarendon Press, Oxford University Press, Oxford, 1967. Zbl 0168.24703. MR 0218199. 405

[34] Myhill, J., "A refutation of an unjustified attack on the axiom of reducibility," pp. 81–90 in *Bertrand Russell Memorial Volume*, edited by George W. Roberts, Muirhead Library of Philosophy, George Allen & Unwin, London, and Humanities Press, New York, 1979. 405

[35] Potter, M., *Reason's Nearest Kin: Philosophies of Arithmetic from Kant to Carnap*, Oxford University Press, New York, 2000. MR 1859099. 407, 410

[36] Pottinger, G., "A tour of the multivariate lambda calculus," pp. 209–30 in *Truth or Consequences: Essays in Honor of Nuel Belnap*, edited by M. Dunn and A. Gupta, Kluwer Academic Publishers, Dordrecht, 1990. 382

[37] Proops, I., "Russell on substitutivity and the abandonment of propositions," *Philosophical Review*, vol. 120, (2011), pp. 151–205. DOI 10.1215/00318108-2010-027. 408

[38] Quine, W. V. O., *From a Logical Point of View: Nine Logico-Philosophical Essays*, 2nd edition, Harvard University Press, Cambridge, Mass., 1994. MR 1445774. 388

[39] Quine, W. V. O., *Set Theory and Its Logic*, revised edition, Harvard University Press, Cambridge, Mass., 1969. MR 0274272. 380

[40] Ramsey, F. P., *The Foundations of Mathematics*, edited by R. B. Braithwaite, Harcourt, Brace and Company, New York, 1931. 380, 405

[41] Russell, B., "Mathematical logic as based on the theory of types," *American Journal of Mathematics*, vol. 30 (1908), pp. 222–62. Zbl 39.0085.03. MR 1506041. DOI 10.2307/2369948. 381, 391, 394, 406

[42] Russell, B., *Philosophical Essays*, Longmans, Green, and Company, London, 1910. Zbl 41.0084.01. 408

[43] Russell, B., *Introduction to Mathematical Philosophy*, George Allen & Unwin, London, 1919. 397

[44] Russell, B., *Principles of Mathematics*, 2nd edition, George Allen & Unwin, London, 1937. 384, 388, 404, 405, 407, 408, 409

[45] Russell, B., *An Inquiry into Meaning and Truth*, W. W. Norton & Company, New York, 1940. 411

[46] Russell, B., *Logic and Knowledge*, edited by Robert C. Marsh, Capricorn Books, G. P. Putnam's Sons, New York, 1956. 406, 409

[47] Russell, B., "Philosophy of logical atomisim," pp. 175–282 in *Logic and Knowledge*, Capricorn Books, G. P. Putnam's Sons, New York, 1956. 387, 409

[48] Russell, B., *My Philosophical Development*, George Allen & Unwin, London, 1959. 411

[49] Russell, B., *The Problems of Philosophy*, Oxford University Press, London, 1959. 389, 408

[50] Russell, B., *Essays in Analysis*, edited by D. P. Lackey, George Allen & Unwin, London, 1973. 388, 406, 410, 412

[51] Sainsbury, R. M., *Russell*, Routledge & Kegan Paul, London, 1979. 395, 408

[52] Soames, S., *Understanding Truth*, Oxford University Press, New York, 1999. Zbl 0978.03007. MR 1824687. DOI 10.1093/0195123352.001.0001. 411

[53] Schutte, K., *Beweistheorie*, vol. 103 of *Die Grundlehren der mathematischen Wissenschaften*, Springer, Berlin, 1960. MR 0118665.

[54] Troelstra, A. S., and H. Schwichtenberg, *Basic Proof Theory*, vol. 43 of *Cambridge*

*Tracts in Theoretical Computer Science*, Cambridge University Press, Cambridge, 1996. MR 1409368. 406

[55] Whitehead, A. N., and B. Russell, *Principia Mathematica, Vol. 1*, Cambridge University Press, Cambridge, 1910; 2nd edition, 1927. 381, 387, 388, 389, 391, 394, 399, 405, 406, 408, 409

[56] Wittgenstein, L., *Tractatus Logico-Philosophicus*, translated by D. F. Pears and B. F. McGuiness, Routledge & Kegan Paul, London, 1961. 411

[57] Wrinch, D., "On the nature of judgment," *Mind*, New Series, vol. 28 (1919), pp. 319–29. 408

## Acknowledgments

Sage School of Philosophy
Cornell University
Ithaca, New York 14853
hth3@cornell.edu