The Logic of *Principia Mathematica*

by

Darryl Jung

B.Eng., McGill University (1985)
M.A., McGill University (1990)

Submitted to the Department of
Linguistics and Philosophy
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

at the Massachusetts Institute of Technology

July 1994 [Feb 1995]

Signature of Author...
Darryl Jung
Department of Linguistics and Philosophy

Certified by...
George Boolos
Professor, Department of Linguistics and Philosophy
Thesis Supervisor

Accepted by...
George Boolos
Chairman, Departmental Committee on Graduate Students

The Logic of *Principia Mathematica*

by

Darryl Jung

Submitted to the Department of Linguistics and Philosophy in July 1994
in partial fulfillment of the requirements for the
Degree of Doctor of Philosophy

ABSTRACT

Bertrand Russell's *Principia Mathematica* has played a crucial role in the development of modern logic and philosophy. Notwithstanding its considerable influence, *Principia*'s contents are not well known and, in important respects, not well understood today. Although several commentaries about *Principia* have appeared since its publication, most of these mischaracterise several of its gross structural features -- for instance, those by Copi, Kleene, Kneale, and Ramsey -- and even some of the better commentaries remain unclear about its details -- for instance, those by Chihara, Hylton, and Quine.

This dissertation examines *Principia Mathematica* in detail. The examination is carried out in light of four general aims. The first is to explain the nature of the logical theory that Russell endeavoured to put forward in *Principia*. The second is to examine how Russell put *Principia*'s logical theory to work to serve his various ends. The third is to show how this logical theory's general structural features and several of its details relate to the wealth of ideas that Russell seriously entertained and pursued in the period in which he was working on *Principia*. Finally, the fourth aim is to compare Russell's conception of his logical theory with contemporary conceptions of logic.

The dissertation is divided into five chapters. Chapter 1 discusses some of the ends that Russell set for himself in his 1903 *Principles* -- ends which he intended the logical theory of *Principia* to accomplish. The next two chapters look at the antecedents to this logical theory: Chapter 2 examines the logical theory of *Principles* and Chapter 3 examines Russell's celebrated vicious-circle principle. The last two chapters describe the logical theory proper of *Principia*: Chapter 4 explains the theory's propositional fragment and Chapter 5 explains its quantificational fragment. This last chapter also details the theory's type-theoretic features.

Thesis Committee: Professor George Boolos
Professor Richard Cartwright
Professor Robert Stalnaker

Acknowledgements

I wish, first of all, to express my deep gratitude to George Boolos, who suggested the project to me, and gently and patiently encouraged me throughout its development. He very carefully read through various drafts of this manuscript, offered detailed comments and insightful recommendations, and provided judicious guidance and counsel. More generally, I owe much to George for my overall philosophical and technical formation. He has been a wonderful and generous teacher and supervisor.

I wish to thank Richard Cartwright, from whose subtle knowledge of Bertrand Russell's work I was able greatly to benefit. His effective criticism of various drafts allowed me to clear up several confusions that I had at different stages of the project.

I also wish to thank Robert Stalnaker who has contributed much to shaping my general philosophical views and work. He has been a patient and generous advisor and has helped me in more ways than I can possibly mention.

I owe Michael Hallett a debt of gratitude for originally taking me under his wing, introducing me to the wonder of mathematical logic, and putting me on my present course. He read through and scrupulously annotated early drafts.

I am indebted to Warren Goldfarb for his lectures and writings on early analytic philosophy. They provided points upon which my project could focus. His conversations with me have been very instructive.

I owe much to Emily Carson, Michael Glanzberg, David Hunter, and Robert Kermode for their conversation, support, and friendship. Because of them, my time at MIT has been not only rewarding, but also very enjoyable.

I am very grateful to my parents, Douglas and Maureen Jung, who supported and encouraged me throughout this undertaking. It would not have been possible without them.

My greatest debt is to my best friend and wife, Sally Archdeacon. Forever patient and understanding, she has been able to inspire in me the courage and confidence that I needed to face each stage of the project. Her care sustains me; she is a constant source of joy in my life. For all this and much more, to Sally, my thanks and my love.

The generous financial support of SSHRC of Canada and FCAR of Québec is gratefully acknowledged.

Table of Contents

Introduction

Bertrand Russell's *Principia Mathematica* stands as a monumental work which has played a pivotal role in the history of logic.[1] Among other things, it contributed one of the first full presentations of modern logic;[2] it provided the first unified solution to all of the modern paradoxes, both logical and semantic; it carried out the first *consistent* reduction of the concepts of mathematics -- more precisely, number theory, transfinite ordinal and cardinal arithmetic, and real analysis -- to those of class theory; and, by means of this reduction, it produced the first *consistent* deduction of the theorems of these branches of mathematics from the axioms of class theory.[3] Accordingly, *Principia* went a long way toward accomplishing the logicistic aims that Russell had earlier set for himself in his 1903 *Principles of Mathematics* (henceforth *Principles*).[4]

Not surprisingly, *Principia*'s influence has been considerable. Philosophically, it was instrumental in forming the views of Wittgenstein, Ramsey, and the logical positivists and, as a result, directly and indirectly affected the development of analytic philosophy in general. Technically, it did no less than co-found modern logic, along with Frege's *Begriffsschrift*. Because Frege's work was not widely circulated during the first half of the century, *Principia*

---

[1]A.N. Whitehead and B. Russell, *Principia Mathematica* (henceforth, *Principia*), vols. 1-3 (Cambridge: CUP, 1910-12; 2nd edn. 1927) Although Whitehead's name appears first, it is Russell who is responsible for those parts that are of logical and philosophical interest.

[2]Frege's *Begriffsschrift* gave the first presentation of modern logic.

[3]Note that, rather than saying 'consistent' here, it would be more accurate to say 'consistent insofar as one has reason to believe'.

[4]*The Principles of Mathematics*. (London: CUP, 1903). (I shall refer to the second edition: New York: Norton, 1938.)

was *the* source-book for students of this new subject, among whom were Carnap, Tarski, Gödel and Quine.

Notwithstanding its considerable influence, however, *Principia*'s contents are not well known and, in important respects, not well understood today. Although several commentaries about *Principia* have appeared since its publication, most of these mischaracterise several of its gross structural features -- for instance, those by Copi, Kleene, Kneale, and Ramsey -- and even some of the better commentaries remain unclear about its details -- for instance, those by Chihara, Hylton, and Quine. There are several reasons for these deficiencies. First, the conception of logic that underlies *Principia*'s technical development is, in many respects, at variance with more contemporary conceptions. Secondly, *Principia*'s discursive explanations and its formal work at times fall short of contemporary standards of rigour and clarity. Thus, for instance, these do not explicitly distinguish the sundry roles played by its axioms, rules of inference, and even rules of formation, usually simply calling them all "principles." *Principia* contrasts notably with Frege's *Grundgesetze der Arithmetik* in this respect. Thirdly, *Principia*'s discursive explanations and formal work contain several peculiarities which may even strike one as incongruous at first blush. Thus, its introduction and Chapter *12 present mutually incompatible definitions of the notion of predicative function. Chapter *9 specifies a very fine-grained notion of type for propositions and propositional functions, while *Principia*'s actual deductions may be taken to appeal to a rather coarse-grained notion -- yet a third notion is employed to carry out class theory. *Principia*'s various explanations of its notion of order have caused scholars to be confused about the nature of its ramified theory of types. And, most notably, its Chapters *9 and *10 offer two quite different and entirely independent expositions of quantification theory.

This dissertation, therefore, examines *Principia Mathematica* in detail. The examination is carried out in light of four general aims. The first is to explain the nature of the logical theory that Russell endeavoured to put forward in *Principia*. The second is to examine how Russell put *Principia*'s logical theory to work to serve (some of) his various ends. The third is to show how this logical theory's general structural features and several of its details relate to the wealth of ideas that Russell seriously entertained and pursued in the period in which he was working on *Principia*. Finally, the fourth aim is to compare Russell's conception of his logical theory with contemporary conceptions of logic.

The dissertation is divided into five chapters. Chapter 1 discusses some of the ends that Russell set for himself in his 1903 *Principles* -- ends which he intended the logical theory of *Principia* to accomplish. The next two chapters look at the antecedents to this logical theory: Chapter 2 examines the logical theory of *Principles* and Chapter 3 examines Russell's celebrated vicious-circle principle. The last two chapters describe the logical theory proper of *Principia*: Chapter 4 explains this logical theory's propositional fragment and Chapter 5 explains its quantificational fragment. This last chapter also details the theory's type-theoretic features.

Chapter 1

The Aims of *Principia Mathematica*

Russell intended *Principia* to accomplish three separate, although interrelated, aims: *Principia* should give a full presentation of modern logic by means of articulating an interpreted formal system. It should provide a unified solution to his and other recently discovered paradoxes, both logical and semantic.[5] It should show by means of the formal system that mathematics can in a precise sense be reduced to logic. Russell had explicitly set these aims for himself in *Principles* and they occupied his attention during the decade that he spent writing *Principia*.[6] I discuss the first and second of these aims separately below. Because I am specifically concerned with *Principia*'s logic and not with its reduction of mathematics to such logic, I forego any discussion of the third aim in what follows.

1. The Formal System

In order to understand the significance of Russell's aim of presenting an interpreted formal system in *Principia*, one must see both what he takes a formal system to be, and how he considers it to be interpreted.[7] I begin by looking at the former concern. Following Frege and to a certain extent Peano, Russell takes such a system to be a formal object consisting of

---

[5]As we shall see, Russell did not think that this distinction had any relevance to the solution of the paradoxes.

[6]Although Russell explicitly states these aims in *Principles*, he never presents them together. They are rather individually repeated throughout the text.

[7]For expository purposes, I shall in what follows consider a formal system proper to be a purely formal object -- that is, something individuated by purely syntactic criteria. I do not intend such individuative criteria, however, to exhaust the important features that Russell ascribes to such an object.

various elements, including crucially its primitive vocabulary, its rules of formation, its axioms, its rules of inference, and its notions of deduction and theoremhood. Russell, in turn, sees these elements as effectively determining both (what we would now call) a formal language and a formal calculus. The primitive vocabulary and rules of formation in effect determine the formal language by specifying the terms, predicates, and open and closed formulae. The notions of deduction and theoremhood in turn determine the calculus by specifying the objects that are the deductions and theorems. Thus, where P is a set of formulae and C is a formula, a deduction from P -- its premise set -- to C -- its conclusion -- is a sequence of formulae that are related to P, C, and each other in certain effectively specifiable ways. A theorem is simply a conclusion of a deduction from an *empty* premise set. When I examine *Principia* in detail, I shall concentrate on Russell's particular definition of these items. From what I have already said, however, it is clear that Russell's notion of formal system, *qua* formal object, is very much like the modern notion.

I now move on to examine how Russell takes a formal system to be interpreted. In contrast to the first concern, Russell's conception in this regard differs considerably from the modern one, although not completely. I divide the examination into two parts. In the first part, I examine how Russell interprets a formal system's formal language and, in the second part, how he interprets its formal calculus. In both these parts, I shall focus exclusively on the formal system that Russell presents in *Principia*.

There are two important aspects of Russell's interpretation of *Principia*'s formal language. For convenience, I dub these *fixity* and *universality*. Fixity: Like Peano and Frege and unlike the contemporary algebraic logicians Boole, Schröder, and Löwenheim, Russell takes *Principia*'s formal language to enjoy a fixed interpretation. He takes this interpretation, moreover, to be fixed in a *compositional* way. Each particle of *Principia*'s primitive vocabulary enjoys a fixed semantics in the sense that it *means* some specific object in his ontology -- which he calls *the realm of being*. Likewise, each formula constructed out of these particles according to *Principia*'s rules of formation enjoys a fixed semantics in the sense that it means what Russell calls a *proposition*, which is simply the ordered complex of objects that are meant by the constituent particles. These claims will be spelled out more fully when we investigate his *Principles* and other pre-*Principia* writings. At this point, however, we can already say that this aspect of how Russell interprets *Principia*'s formal language is very much at variance with the contemporary model-theoretic conception. According to this conception, we may treat the terms and formulae of any formal language as *de-interpreted* and assign them whatever mathematically tractable interpretation we may choose.

Universality: As is well-known, Russell takes *Principia*'s formal language to be universal and, indeed, he does so in two separate but interrelated ways. First, he construes its variables to range over everything there is -- that is, everything there is in his realm of being. It is noteworthy that Russell had, in fact, construed his formal language's variables to range over everything there is long before he wrote *Principia* but the particular way in which he did so contrasts starkly from the way in which he does in *Principia*. In *Principles* and other pre-*Principia* writings, he meant this in the sense that each variable was understood to range over a completely unrestricted universe of discourse.[8] In these earlier writings, he more than once offered a certain argument for this construal:

> It is customary in mathematics to regard our variables as restricted to certain classes: in Arithmetic, for instance, they are supposed to stand for numbers. But this only means that *if* they stand for numbers they satisfy some formula, i.e., the hypothesis that they stand for numbers implies the formula. This, then, is what is really essential, and in this proposition it is no longer necessary that

---

[8]This position, however, seems to be rejected in one of *Principles*'s appendices in which a version of the simple theory of types is put forward.

> our variables should be numbers: the implication holds equally when they are not so. ... Thus in every proposition of pure mathematics, when fully stated, the variables have an absolutely unrestricted field.[*Principles*, 7]

In *Principia*, by contrast, Russell introduces his ramified theory of types specifically in order to forbid this particular construal because he holds it responsible for the paradoxes. Nevertheless, he construes *Principia*'s variables to range over everything there is in the sense that anything there is is understood to fall under some particular type and each type is understood to have a class of variables that range over and only over all of the objects that fall under it. Note that this construal is still similar to his earlier construal in that it entails that there can never be anything in the realm of being that falls outside the scope with which his formal language's variables are concerned.

Secondly, and more generally, Russell takes *Principia*'s formal language to be universal in the sense that he takes any proposition (which is expressible by any means) to be expressible by means of some one of its formulae. If we consider Russell's theory of propositions -- we shall discuss this in detail when we look at *Principles* -- and his ramified theory of types, we may understand this position to be an unsurprising and natural outcome. According to his theory of propositions, any given proposition P is simply an ordered complex of objects each of which belongs to the realm of being. According to his ramified theory of types, each object in the realm of being falls under a type and each type has a class of particles -- its variables -- belonging to *Principia*'s primitive vocabulary which range over all and only the objects falling under that type. Any particular particle of such a class of particles, therefore, may be taken at any one time to refer to any object falling under the type.[9] In this light, if for each constituent

[9]For the moment, I am saying that a variable may be taken to refer to any object belonging to its type. Russell's account is actually more complicated than this construal and I shall discuss it when I examine his infamous any/all distinction.

object o of the given proposition P, we construe a particle of its type as referring to o, then if we arrange all such particles according to the rules of formation in a way that is analogous to the way that P's constituent objects are arranged, these particles will constitute a formula of *Principia*'s formal language that means the proposition P.

Russell is most forthright about this position when he asserts its contrapositive, *viz.*, if there is no formula of *Principia* that expresses some putative proposition, then *it* cannot be expressed by any means -- in particular, *it* cannot be expressed by any sentence of ordinary language even if some such sentence may appear to do so. Russell very explicitly asserts this contraposition in various passages where he discusses the consequences of the theory of types. There he says of any sentence of ordinary language that violates the grammatical restrictions dictated by the theory of types that even if it may appear to express a proposition, it must fail to do so. Thus, he writes in his 1937 Introduction to *Principles*:

> The technical essence of the theory of types is merely this: Given a propositional function "ψx" of which all values are true, there are expressions which it is not legitimate to substitute for "x." For example: All values of "if x is a man x is a mortal" are true, and we can infer "if Socrates is a man, Socrates is mortal"; but we cannot infer "if the law of contradiction is a man, the law of contradiction is a mortal." The theory of types declares this latter set of words to be nonsense, and gives rules as to permissible values of "x" in "ψx." In the detail there are difficulties and complications, but the general principle is merely a more precise form of one that has always been recognized. In the older conventional logic, it was customary to point out that such a form of words as "virtue is triangular" is neither true nor false, but no attempt was made to arrive at definite set of rules for deciding whether a given series of words was or was not significant. This the theory of types achieves.[*Principles*, p. xiv]

Russell actually holds a stronger version of the position described above. Namely, not only does he take any proposition that is expressible by any means to be expressible by means of some formula belonging to *Principia*'s formal language, but he also takes it that insofar as any proposition is expressible by some means -- say by means of some sentence of ordinary language -- it is so expressible in virtue of already being expressible by a formula of *Principia*. The reason that Russell holds this stronger position stems from his view of logical analysis, which he developed in the decade before *Principia* was published.[10] Briefly, according to this view, although a given sentence S of ordinary language may express a given proposition P, S may have on the surface a grammatical form that is radically different from the "logical form" of P. In order to discover P's logical form, we should start by analysing S. Such an analysis should lead to another sentence S_1 which also expresses P and whose grammatical form is in some sense closer to the logical form of P. We should then analyse S_1 and this process should lead successively to the sentences S_2, S_3, etc., each of which expresses P and each of which has a grammatical form that is closer to the logical form of P than is the grammatical form of its predecessor. At the end, this process should lead to some sentence S_p whose grammatical form is the logical form of P. Now, according to this view of analysis, this sentence S_p exhibits the true logical form of S in the sense that it is in virtue of having this form that S may behave in the logical way that it is so licensed. Moreover, S expresses the proposition P in virtue of the fact that S_p does, for S is really a kind of disguised abbreviation for S_p. As one might guess, S_p is supposed to be a formula belonging to the formal language of *Principia*.[11,12]

---

[10]Russell's view of analysis originates from Moore. Indeed, except for a few complications, it is Moore's view of analysis that is articulated in *Principles*. However, Russell extends and modifies Moore's view considerably in various papers after 1903. His "On Denoting" is significant in this respect because in it, he introduces the important notions of incomplete symbol and contextual definition.

[11]Here, I am distinguishing sentences according to some suitable notion of type.

[12]I should note here that Russell considers every formula of *Principia*'s formal language to exhibit its logical form in a completely perspicuous way and, for this reason, he calls it a *logically perfect language*. Cf. *The Philosophy of Logical Atomism* (LA), (La Salle, Illinois: Open Court, 1985), p. 58. Russell also notes there that we may add whatever nonlogical particles we may like -- as long as we do not modify the syntax -- to *Principia*'s formal language without affecting its logical perfection.

a result of this acquaintance.

To summarise the last few paragraphs, we have seen that Russell takes *Principia*'s formal language to be universal both in the sense that there is nothing which lies outside the scope of *Principia*'s variables and in the sense that any proposition which is expressible by some means is expressible by some formula of *Principia*. These two aspects of how Russell interprets *Principia*'s formal language are to a certain extent at variance with contemporary conceptions. Concerning the first aspect, some contemporary writers -- for instance, Dummett and Parsons -- argue that it makes no sense to conceive of the variables of any formal system as enjoying a completely unrestricted domain of discourse.[17] Others -- most notably Boolos, Cartwright, and Quine -- argue that this possibility makes perfect sense.[18] Concerning the second aspect, perhaps no one today takes the formal language of any formal system to be a universal framework inside of which any proposition that is expressible by any means is also expressible. However, there are contemporary twists to this theme. Quine, for instance, recommends on methodological grounds that our discourse be regimented according to the strictures of first-order logic and claims that it makes sense to discuss what the ontological commitments of a given discourse are only after such a regimentation of the discourse in question is achieved.[19] When discussing "radical interpretation," Davidson says that first-order logic provides the most useful structure in terms of which a language that is radically different

---

[17]M. Dummett, *Frege: Philosophy of Language*, 2nd edn. (Cambridge, Mass.: HUP, 1981), ch. 16. C. Parsons, "Sets and Classes" in *Mathematics in Philosophy*, (Ithaca, New York: Cornell University Press, 1983).

[18]G. Boolos, "Whence the Contradiction?" in *Aristotelian Society*, Suppl. vol. 67 (1993) pp. 213-234. R. Cartwright, "Speaking of Everything", 1989 ASL lecture. W.V. Quine, "On What There Is", "Existence and Quantification", among others.

[19]Cf. "On What There Is", "Existence and Quantification", *Word and Object*: chs. 3 and 5, among others.

Russell first puts forth this stronger position with respect to mathematical discourse. In the opening pages of *Principles*, he claims that sentences that express mathematical propositions really express propositions of pure logic.[13] Such a claim, indeed, follows from his adherence to logicism. Later, when Russell directs his attention to epistemological concerns, he makes a claim that is somewhat similar about sentences that express material-object propositions and propositions of natural science. That is, he claims that such sentences are really disguised abbreviations for sentences about sense-data and logical constructions of sense-data. In "The Relation of Sense-Data to Physics"(1914),[14] he hints at how such disabbreviation may be carried out and, in *Our Knowledge of the External World*(1914, the Lowell Lectures), he carries it out in some detail.[15]

Concerning the claim that a sentence S of ordinary language may express a proposition P only in virtue of the fact that it is an abbreviation of some formula S_p of *Principia* and that S_p expresses P, one may ask in virtue of what does S_p express P. Russell answers this question by appealing to his celebrated notion of acquaintance. He says in "Knowledge by Acquaintance and Knowledge by Description"(1910), and elsewhere, that a disabbreviated sentence such as S_p expresses its proposition P in virtue of the fact that we are acquainted with the constituents of P.[16] Presumably, he considers the particles of such a sentence to mean these constituents as

---

[13]Here, as elsewhere, Russell actually talks at the level of propositions and not at the level of their linguistic vehicles. My discussion in the following, therefore, is somewhat of a reconstruction.

[14]"The Relation of Sense-Data to Physics" in *Mysticism and Logic*, pp. 108-131.

[15]Carnap carries this disabbreviation project even further in his *Der Logische Aufbau der Welt* (1928).

[16]According to his notion of acquaintance, the only items that we are acquainted with are logical objects -- his *indefinables* of *Principles* -- and sense-data.

from one's home language should be regimented in order to interpret it.[20]

I now move on to discuss the second part of the examination of how Russell takes the formal system presented in *Principia* to be interpreted -- that is, the part having to do with how he interprets the formal system's formal calculus. In this part, I first consider how Russell interprets the formal calculus's deductions and then consider how he interprets its theorems.

There are three points to note about how Russell interprets the deductions in *Principia*.[21] Recall that a deduction belonging to *Principia*'s formal calculus is a purely formal object. More precisely, if D is a deduction from a set of premises P to a conclusion C, then D is simply a sequence of formulae that relate to P, C, and each other in certain effectively specifiable ways, where these are described by the formal calculus's rules of inference.[22] Recall also that the formulae of *Principia* enjoy a fixed interpretation in the sense that any such formula expresses a particular proposition. It follows from these considerations that for any deduction D, there exists a sequence of propositions A such that to any formula F of D there corresponds in a one-to-one fashion a proposition Q of A which F expresses. For the purpose of exposition, we may call any such sequence of propositions A that so corresponds to some deduction D of the formal calculus a *PM-argument* (*Principia*-argument). The first of the three points about how Russell interprets *Principia*'s deductions, then, is simply that he

---

[20]Cf. "Radical Interpretation" in *Truth and Interpretation*, (Oxford: Clarendon, 1985).

[21]It should be noted that, in the following, I distinguish carefully between various formal or syntactic items and their semantic analogues. Russell in his writing is not always so careful. For the most part, he only talks explicitly about the semantic items, although he clearly presupposes an intimate connection between these and their formal analogues. Thus, where I distinguish between deduction, *qua* formal object, and argument, *qua* semantic analogue, Russell uses the word "deduction" to refer to both.

[22]Among these rules of inference are *modus ponens*, substitution, definitional interchange, and universal generalisation.

takes any such PM-argument to be logically valid. In saying this, I should note here that Russell does not construe logical validity in the way that Tarski later famously explicated this notion.[23,24] Rather, Russell construes logical validity to mean something like validity in virtue of general reasoning.[25] This construal is, of course, vague and unsatisfactory; but the point is that Russell takes the traditional notion of logical validity to be an important and completely legitimate notion that is perhaps primitive in the sense of not being definable or explicable in simpler terms.

The second point about how Russell interprets *Principia*'s deductions is that he holds the following contemporary-sounding position: not only is every PM-argument logically valid, but insofar as any argument expressed in ordinary language is logically valid, it is so in virtue of there being a PM-argument which corresponds to it.[26] More precisely, where P is a set of premises expressed by ordinary language sentences, C is a putative conclusion from P expressed by an ordinary language sentence, and B is an argument from P to C consisting of propositions expressed by ordinary language sentences, then insofar as B is logically valid, there exists a PM-argument from P to C expressed by formulae of *Principia*. It follows from this position that an argument is logically valid iff it is a PM-argument or corresponds to a PM-

---

[23]Tarski, Alfred. "On the Concept of Logical Consequence" 1936. Cf. W. Hodge's "Truth in a Structure".

[24]However, supposedly Bolzano had already articulated a similar explication.

[25]Define a *truth-preserving relation* as follows. A relation R between propositions $p_1, p_2, \ldots, p_n$, is truth-preserving whenever: R holds of $p_1, p_2, \ldots, p_n$ iff both (a) R holds in virtue of the logical structures which propositions $p_1, p_2, \ldots, p_n$ exhibit and (b) if $R(p_1, p_2, \ldots, p_n)$ and $p_1, p_2, \ldots, p_{n-1}$ are all true, then p_n must be true also. It is perhaps possible to say that Russell takes an argument to be logically valid when and only when truth-preserving relations hold in suitable ways between various propositions of the argument.

[26]Note the similarity between this position and the position described above concerning the possibility of there being an ordinary language sentence expressing a proposition.

argument. Since something is a PM-argument iff there is a corresponding deduction in the formal calculus, it also follows that any argument is logically valid iff it has a corresponding deduction in the calculus. In this respect, although Russell takes the notion of logical validity to be indefinable, he nevertheless gives a characterisation of the notion in proof-theoretic terms.

Russell's adherence to this position is most apparent when he advances the claim that all correct mathematical reasoning can be formalised in *Principia*'s formal calculus. Indeed, this claim is entailed by his claim of logicism. He writes in a number of places that where mathematicians have, in fact, employed methods that are not amenable to such formalisation, these methods have led to results which, although they may be obvious, do not *follow* from their premises. For instance, in 1901 he writes this about the use of figures in geometry:

> In Geometry, as in other parts of mathematics, Peano and his disciples have done work of the very greatest merit as regards principles. Formerly it was held by philosophers and mathematicians alike that the proofs in Geometry depended on the figure; nowadays, this is known to be false. In the best books there are no figures at all. The reasoning proceeds by the strict rules of formal logic from a set of axioms laid down to begin with. If a figure is used, all sorts of things seem obviously to follow, which no formal reasoning can prove from the explicit axioms, and which, as a matter of fact, are only accepted because they are obvious.[27]
>
> The rigid methods employed by modern geometers have deposed Euclid from his pinnacle of correctness. ... Countless errors are involved in his first eight propositions. That is to say, not only is it doubtful whether his axioms are true, which is a comparatively trivial matter, but it is certain that his propositions do not *follow* for the axioms which he enunciates. A vastly greater number of axioms, which Euclid unconsciously employs, are required for the proof of his propositions. Even the first proposition of all, where he constructs an equilateral triangle on a given base, he uses two circles which are assumed to intersect. But no explicit axiom assures us that they do so, and in some kinds of spaces they do not always intersect. ... Thus Euclid *fails entirely to prove his point* in the very first proposition.[28]

---

[27] "Mathematics and the Metaphysicians," (MM) in *Mysticism and Logic*, Barnes and Noble, p. 72.

[28] *ibid.* pp. 72-3, my italics.

Moreover, Russell writes that not only has the employment of methods that are not amenable to formalisation led to results that do not follow, but such employment has actually led to problematic and even false results. He cites the foundational problems of the infinitesimal calculus as an instance. In the same essay, in a somewhat later passage, he also claims that formalisation is responsible for resolving these problems:

> The most remarkable result of modern methods in mathematics is the importance of *symbolic logic and of rigid formalism*. Mathematicians, under the influence of Weierstrass, have shown in modern times a care for accuracy, and an aversion to slipshod reasoning, such as had not been known among them previously since the time of the Greeks. The great inventions of the seventeenth century -- Analytic Geometry and the Infinitesimal Calculus -- were so fruitful in new results that mathematicians had neither time nor inclination to examine their foundations. Philosophers, who should have taken up the task, had too little mathematical ability to invent the new branches of mathematics which have now been found necessary for any adequate discussion. Thus mathematicians were only awakened from their 'dogmatic slumbers' when Weierstrass and his followers showed that many of their most cherished propositions are in general *false*. Macaulay, contrasting the certainty of mathematics with the uncertainty of philosophy, asks who ever heard of a reaction against Taylor's theorem? If he had lived now, he himself might have heard such a reaction, for this is precisely one of the theorems which modern investigations have overthrown. Such rude shocks to mathematical faith have produced that love of formalism which appears, to those who are ignorant of its motive, to be mere outrageous pedantry.[29]

It should be mentioned that it is only because of the success of the new logic that Russell can plausibly adhere to the stronger position above described. Until 1879, there remained many kinds of argument in mathematics that resisted formalisation according to the existing principles of logic yet that were taken to be instances of correct reasoning. With the new logic, however, one could show that any of the inferential steps of these recalcitrant kinds of argument could be understood to be composed of several "smaller" inferential steps such that each of these is effectively described by one of the new logic's few rules of inference. Both

---

[29]*ibid*. p. 73, my italics.

Russell and Frege were very aware of this fact:

> If we try to list all of the laws governing the inference which occur when arguments are conducted in the usual way, we find an almost unsurveyable multitude which apparently has no precise limits. The reason for this, obviously, is that these inferences are composed of simpler ones.[30]

The third point about how Russell interprets *Principia*'s deductions -- which is closely related to the above discussion -- is that he takes any PM-argument to be "gapless" in the sense that Frege takes his formalised arguments to be. Thus, Russell writes in the Preface to *Principia*:

> We have found it necessary to give very full proofs, because otherwise it is scarcely possible to see what hypotheses are really required, or whether our results follow from our explicit premisses. (It must be remembered that we are not affirming merely that such and such propositions are true, but also that the axioms stated by us are sufficient to prove them.) At the same time, though full proofs are necessary for the avoidance of errors, and for convincing those who may feel doubtful as to our correctness, yet the proofs of propositions may usually be omitted by the reader who ...[*Principia*, p. vi]

Elsewhere, Russell makes it clear that for a proof to be full in the above sense, it must be gapless. Discussing the importance of formal systems, he writes:

> It is not easy for the lay mind to realize the importance of symbolism in discussing the foundations of mathematics, ... The fact is that symbolism is useful because it makes things difficult. ... What we wish to know is, what can be deduced from what. Now, in the beginnings, everything is self-evident; and it is very hard to see whether one self-evident proposition follows from another or not. Obviousness is always the enemy of correctness. Hence we invent some new and difficult symbolism, in which nothing seems obvious. Then we set up certain rules for operating on the symbols, and the whole thing becomes *mechanical*. In this way we find out what must be taken as premise and what can be demonstrated or defined.[31]

---

[30]Frege, "Über die Begriffsschrift des Herrn Peano und meine eigene," *Berichte über die Verhandlungen der Königlich Sachsischen Gesellschaften der Wissenschaften zu Leipzig*, Mathematischephysische Classe, 48 (1897). Translated as "On Mr. Peano's Conceptual Notation and My Own," *in Collected Papers*, p. 235.

[31]MM, p. 61, my italics.

In this respect, we may say, roughly speaking, that a given argument from a premise set P to a conclusion C is gapless if and only if no extraneous premise -- that is, no proposition not belonging to P, where P is taken to be effectively decidable -- is appealed to in the course of the argument to C, in the way, for instance, that Kant insists -- when he talks about "construction of concepts" in his first Critique -- that intuition must be appealed to at any given step of a mathematical proof. Now, that any PM-argument is in this sense gapless actually follows from the specifications of what it is to be a PM-argument and what it is to be a premise to such an argument For since the set of premises in a PM-argument is effectively decidable and since a proposition can occur in a PM-argument only if it relates to previously occurring propositions and/or propositions belonging to the argument's premise set in an effectively decidable way, there can be no tacit appeal to an extraneous premise in any interesting sense.

We now move on to consider how Russell interprets the theorems of *Principia*'s formal calculus. Recall that a theorem is a formula of *Principia*'s formal language which is the conclusion of a deduction whose premise set is empty. It is perhaps unnecessary to say that Russell takes the propositions expressed by *Principia*'s theorems to be logical truths and that he intends the notion of theoremhood to capture proof-theoretically the notion of logical truth. In order to understand what this means, however, we must ask how he otherwise characterises the notion of logical truth. In his Introduction to the second edition of *Principles* in 1937 and elsewhere, Russell states two relevant characteristics:

> A logical [truth] must have *certain* characteristics which can be defined: it must have complete generality, in the sense that it mentions no particular thing or quality; and it must be true in virtue of its form.
> [*Principles*, p. xii]

For a proposition to mention no particular thing or quality, Russell claims that it must consist of nothing but variables and logical constants.[32,33] Unfortunately, he does not have a substantive theory of what it is to be a variable or a logical constant (he confesses this in his 1937 Introduction), so it is unclear how we are to explain this first characteristic of logical truth.[34] Concerning the second characteristic, Russell writes:

> I confess, however, that I am unable to give any clear account of what is meant by saying that a proposition is "true in virtue of its form."[*Principles*, p. xii]

Notwithstanding this remark, Russell clearly has some intuitive notion in mind, for he claims shortly before the remark that it is on the basis of this characteristic that the axioms of choice and infinity are or are not logical truths. Like the first characteristic, however, it is unclear how we should spell out this second one. In this light, it is difficult for us to say in virtue of what the propositions expressed by the theorems of *Principia* are supposed to be logical truths.

Let us summarise what we have seen about how Russell interprets the formal calculus of *Principia*. Concerning its formal deductions, we have seen that Russell takes any such deduction to express a gapless, logically valid PM-argument and, moreover, that insofar as any argument is logically valid, it has a corresponding PM-argument. In this respect, *Principia*'s formal calculus may be taken to capture the notion of logical validity proof-theoretically. Concerning the formal calculus's theorems, we have seen that Russell takes these to express

---

[32]*Principles*, §§ 3, 12, etc.; *Principia*, p. 93.

[33]As we shall see when we examine the *Principles*, Russell treats these items as actual objects in the realm of being, not as linguistic particles.

[34]Notwithstanding this unclarity, some have drawn interesting consequences from this claim. Peter Hylton, in his *Russell, Idealism, and the Emergence of Analytic Philosophy* (RIAP), writes:

> A consequence of this is that if, for some number *n*, there are exactly *n* entities, then it is a truth *of logic* that there are exactly *n* entities. (We can say that there are exactly *n* entities using only variables and logical constants ... [p. 200, note 39]

This is a curious consequence given that Russell never takes the axiom of infinity to be a logical truth. See the comment about the second characteristic.

logical truths and that he intends the calculus proof-theoretically to capture the notion of logical truth. To this extent, it is clear that there are substantive differences between Russell's conception of *Principia*'s deductions and theorems and the modern conception of these items. I mention only the most notable. Whereas Russell takes any *Principia* deduction to express a PM-argument, in modern terms, a deduction is a de-interpreted schema that may be appealed to as a finitary means for determining whether or not the explicated, model-theoretic relation of logical consequence holds between certain formulae -- more precisely, between a certain set of formulae and a distinguished formula. Also, whereas Russell takes any *Principia* theorem to express a proposition that is logically true, the modern conception treats any such theorem also as a de-interpreted schema that satisfies the notion of logical truth, again explicated model-theoretically. Note that since Russell's notion of logical truth and the model-theoretic one apply to different kinds of object, they themselves must certainly be different. Clearly, such differences stem partly from Russell's taking *Principia*'s formal language to have a fixed interpretation.

In an effort to make clear the significance of Russell's *aim* in *Principia* of presenting an interpreted formal system, I have in this section examined what he takes a formal system to be, and how he considers it to be interpreted. To repeat, Russell takes a formal system to be more or less the same formal object that we take it to be. But he considers its formal language to have a fixed and universal interpretation, its formal deductions to express logically valid arguments, and its theorems to express logical truths. Considering the nature of each of these claims, it is easy to appreciate the significance Russell may attach to the aim in question.[35]

Before moving on to discuss the second aim that Russell has in *Principia*, I should

---

[35]Throughout the rest of this work, I shall use the expressions "logical system" and "logic" to mean interpreted formal system as described in this section.

discuss one inference that some scholars -- for instance, Goldfarb, Hylton, and Ricketts -- have made from the premise that Russell takes *Principia*'s formal language to be universal in the above described senses.[36] From this premise, they have inferred that for Russell (and Frege) there is no room for metatheoretic considerations about logic in the sense that there is no legitimate stance from which to frame such considerations. One may, perhaps, explain their inference as follows. *Principia*'s formal language is universal in the sense that anything that can be expressed by some means must be expressible by some formula of that language. Thus, there is no "stance" *outside* that language from which to express anything in the sense that there can be no proposition which is not already expressible from *inside* the language. From this, it *should follow* that there can be no proposition *about* the language, the formal system, or logic in general -- that is, there can be no metalogical propositions at all.

The basis for this inference appears questionable. It seems to be possible to suppose the premise about the universality of *Principia*'s formal language without supposing the conclusion by merely supposing that the universal formal language be able to talk about itself and other aspects of *Principia*'s formal system. At least two considerations about Russell cohere well with this supposition. In *Principles*, and elsewhere, Russell subscribes to a noncategorial conception of logic, according to which the variables of quantification enjoy a completely unrestricted range over which to vary. Also, Russell considers and uses forms of self-referential reasoning, for instance, in the generation of the semantic paradoxes and in the generation of the propositional-function version of his own paradox -- we shall look at these in

---

[36]See W. Goldfarb, "Logic in the Twenties: The Nature of the Quantifier", in *Journal of Symbolic Logic* 44 (1979) p. 353, "Russell's Reasons for Ramification" in Savage and Anderson, p. 27; Hylton, RIAP, p. 202-3; T. Ricketts, "Frege, the *Tractatus* and the Logocentric Predicament", *Nous* 19 (1985) pp. 4-9. Also, cf. J. van Heijenoort, "Logic as Calculus and Logic as Language" in *Synthese* 17 (1967), pp. 324-30.

the next section.

In questioning the basis of this inference, one must concede that Russell does make various remarks which independently offer some support to the conclusion. The most celebrated remark of this kind is to be found in §17 of *Principles*:

> it should be observed that the method of supposing an axiom false, and deducing the consequences of this assumption, which has been found admirable in such cases as the axiom of parallels, is here not universally available. For all our axioms are principles of deduction; and if they are true, the consequences which appear to follow from the employment of an opposite principle will not really follow, so that arguments from the supposition of the falsity of an axiom are here subject to special fallacies.[37]

Moreover, it would be somewhat anachronistic to ascribe Russell a substantial metatheory since the distinction between object-theory and metatheory was only made clear in a mathematically precise way by Hilbert and then Tarski some time after Russell had stopped working on logic.

In the light of all of these considerations, I shall, as an exegetical rule of thumb, read Russell as in general working at the object-theory level. I shall, however, take a neutral stand, to the extent that this is possible, on whether or not Russell can seriously entertain metatheoretic considerations.

2. <u>The Solution of the Paradoxes</u>

Another of Russell's aims in *Principia Mathematica* is that its logical system should provide a unitary solution to his and the other modern paradoxes. In order to understand the significance of this aim, one must see both Russell's reasons for thinking that the logical system *per se* should provide their solution and his reasons for thinking that their solution should be unitary. I

---

[37]See also *Principia*, p. 91.

examine Russell's reasons in the following. Toward this end, I first review all of the relevant paradoxes that exercised his attention in the decade in which he was working on *Principia*.

(a) The Paradoxes[38]

In 1897, Cesare Burali-Forti discovered the first of the modern paradoxes.[39] The argument to this paradox is rather technical: it appeals to the resources developed within Cantor's theory of the transfinite. An informal version of the argument to the paradox goes as follows: Let x be any class of ordinal numbers. Then x is well-ordered by the notion of magnitude particular to the ordinal numbers in question. Since any well-ordered class has an order-type represented bv some ordinal number, x has an order-type represented by some ordinal number α. By a simple derivation, it can be shown that this α cannot belong to x. Now, consider the class On of all ordinal numbers. Since On is well-ordered, it has some ordinal number β representing its order-type. As in the case with α, β cannot belong to On. Contradiction.

In 1899, Cantor discovered a simpler paradox. Like the argument to Burali-Forti's paradox, the argument to this paradox appeals to resources developed within his theory of the transfinite, although these resources concern cardinal rather than ordinal numbers. An informal version of the argument to the paradox goes as follows: According to Cantor's theorem, for any class x, the class of all its subclasses $\wp(x)$ -- its power-class -- has a cardinality strictly greater than the cardinality of x, where the cardinality of a class z is in this sense strictly greater

[38]Those who are familiar with the modern paradoxes may jump ahead to the discussion of Russell's less well known propositional function paradox and proposition paradox -- these are, respectively, the fourth and fifth paradoxes examined here.

[39]In his *Cantorian Set Theory and the Limitation of Size* (Oxford: Clarendon Press, 1984), Michael Hallett points out that Cantor had known of this paradox as early as 1895. See pp. 74, 169.

than that of a class y iff there is a one-to-one map from y into z but no onto map from y to z.[40] Now, consider the universal class V -- that is, the class of all classes. By Cantor's theorem, the cardinality of $\wp(V)$ is strictly greater than that of V. Thus, there is no onto map from V to $\wp(V)$. Accordingly, there must be at least one class in $\wp(V)$ which is not in V, the class of all classes.[41]

In May of 1901, Russell discovered the simplest of all the modern paradoxes.[42,43] Unlike the arguments to the former two paradoxes, the argument to this one does not appeal to any of the resources of Cantor's theory of the transfinite; indeed, it is very easily stated. Russell was led to the paradox by examining Cantor's argument to his own paradox. In this examination, he came to consider that a class sometimes is, and sometimes is not, a member of itself. So, for example, the class of teaspoons is not a member of itself since it is not a teaspoon, whereas the class of non-teaspoons is a member of itself since it is a non-teaspoon. Call a class that is a member of itself *self-membered* and a class that is not *non-self-membered*. Now, consider the class R of all non-self-membered classes. Query: is R non-self-membered? Suppose that it is; that is, suppose that $R \notin R$. Then R satisfies the condition for belonging to itself. So $R \in R$. Conversely, suppose that R is self-membered, so $R \in R$. Then, R must satisfy the condition for belonging to itself, *viz.*, being non-self-membered. So $R \notin R$. Hence, $R \in R$ iff

---

[40]One could define 'strictly-greater-than' as follows: the cardinality of a class z is strictly greater than that of a class y iff there is a map from z onto y but no such map is one-to-one. However, the axiom of choice is required to prove that this definition is equivalent to the definition given above.

[41]Care is required to state this argument formally.

[42]In his autobiography and in his *My Philosophical Development* (henceforth, MPD), Russell writes that he discovered it in May; however, in his *My Mental Development*, he writes that he discovered it in June.

[43]Zermelo discovered this paradox independently in 1902.

$R \notin R$.

This argument may be easily formalised in order to show precisely what is required in order to obtain the contradiction. Crucially, assume a principle of unrestricted comprehension governing the existence of classes: $\forall\psi\exists y\forall x(x \in y \leftrightarrow \psi(x))$ (i). Here, the variable 'ψ' ranges over propositional functions. From (i), infer $\exists y\forall x(x \in y \leftrightarrow x \notin x)$ by universally instantiating '$\psi[1]$' with '$[1]\notin[1]$'.[44] Next, instantiate "y": $\forall x(x \in y \leftrightarrow x \notin x)$. Then instantiate "x": $y \in y \leftrightarrow y \notin y$. Thus, only four interesting items are required to obtain the contradiction: the comprehension principle, the well-formedness of '$x \notin x$', universal instantiation, and existential instantiation.[45]

In §§96 and 101 of *Principles*, Russell discusses another paradox whose argument parallels the argument to his class paradox. This argument, however, appeals to propositional functions rather than to classes. The argument may be obtained by replacing in the former argument references to classes by references to monadic propositional functions and references to membership by references to satisfaction. Briefly, call a monadic propositional function that applies to itself *self-satisfying* and one that does not *non-self-satisfying*. Now, consider the monadic propositional function Q which applies to all and only non-self-satisfying monadic propositional functions. Query: Does Q satisfy itself? Answer: Q satisfies itself iff it does not.

This argument may be put more formally as follows. Assume a principle of unrestricted comprehension governing the existence of monadic propositional functions:

---

[44]Quine's circled numerals are employed here in order correctly to substitute a higher-order variable with an open formula. See Quine's *Elementary Logic*, (Cambridge Massachusetts: HUP, 1980) §§ 40-42.

[45]Note that the rule of universal instantiation appealed to here in order to replace the propositional function variable by an open formula is not as ontologically innocent as it might at first appear to be. In particular, it entails strong existence assumptions, among which are principles of comprehension. See Quine, *Set Theory and its Logic*, pp. 257-8.

$\forall\psi\exists y\forall x(y(x) \leftrightarrow \psi(x))$, where the variable '$\psi$' ranges over propositional functions and the variables 'x' and 'y' range over monadic propositional functions. Infer $\exists y\forall x(y(x) \leftrightarrow \neg x(x))$ by universally instantiating '$\psi[1]$' with '$[1]([1])$'. Instantiate 'y': $\forall x(y(x) \leftrightarrow \neg x(x))$. Instantiate 'x': $y(y) \leftrightarrow \neg y(y)$. Thus, nothing interesting is required to obtain the contradiction but a comprehension principle, the well-formedness of '$\neg x(x)$', and universal and existential instantiation.

In §500 of *Principles*, Russell discusses yet another paradox which is similar to his class and propositional function paradoxes but which focuses on propositions. Briefly, consider that for any class M of propositions, there is a proposition which asserts that every member of M is true. Call such a proposition a *c-proposition* ('c' for class). Now, some c-propositions are members of the class of propositions which they assert to be true and others are not. Call the former c-propositions *sc-propositions* ('s' for self) and the latter c-propositions *non-sc-propositions*. Consider the class W of all and only non-sc-propositions. There is a c-proposition p which asserts that every member of W is true. Query: $p \in W$? Answer: $p \in W$ iff $p \notin W$. Formalising this argument shows that the only interesting notions which it employs are those of class, proposition, and truth, and that the only interesting principles that it appeals to are unrestricted comprehension governing class existence and an existence principle for propositions.[46]

In 1905, König discovered a paradox that involves the notion of *specifiability* -- roughly speaking, something is specifiable iff there is a finite sequence of symbols that uniquely picks it out. The argument to the paradox is simple: Since there can be only $\aleph_0$ finite sequences of symbols, there can be only $\aleph_0$ ordinal numbers that are specifiable. However,

[46]Those who are familiar with the semantic paradoxes may jump ahead a few pages to the discussion of Ramsey.

Finally, in 1907, Poincaré suggested that the ancient liar paradox -- the Epimenides paradox -- should be considered among these others.[48] Its simplest argument is the following: Consider the sentence "I am lying". Query: If it is uttered in an 'appropriate' context, is it true or false? Answer: It is true iff it is false.

Although we have only seen informal arguments to the last five paradoxes, formal analogues of these arguments can be given, albeit with some effort. These analogues show that the most interesting principles that are appealed to by them are the principles that govern the notions of specification, application and truth. Generally, these principles are either disquotational in nature, or entail others that are.

Before moving on to subsection (b), I should note here that generally today, Burali-Forti's paradox, Cantor's paradox, and Russell's class paradox are called *set-theoretic paradoxes* because their arguments employ the notion of class and appeal to principles that govern this notion; König's paradox, Richard's paradox, Berry's paradox, and the liar paradox are called *semantic paradoxes* because their arguments employ semantic notions like specifiability, applicability, and truth and appeal to principles that govern these notions. Peano first drew the distinction between these two groups of paradoxes in 1906,[49] but Ramsey is the one who is recognised for arguing for its importance, which he did in his 1923 "Foundations of Mathematics".[50] There, he writes:

> [The set-theoretic paradoxes] involve only logical or mathematical terms such as class and number, and show that there must be something wrong with

---

[48]"Les mathématiques et la logique", *Revue de Métaphysique et de Morale*, 1906, 14, p. 306.

[49]*Rivista di Mat.* 8 (1906), p. 157.

[50]"Foundations of Mathematics" (henceforth, FM), in *Foundations*, (Atlantic Highlands, NJ: Humanities Press,1978).

there are certainly more than $\aleph_0$ ordinal numbers and, since the class of all ordinal numbers is well-ordered, there must be a least ordinal number β that is not specifiable. Now, note that the description "the least ordinal number that is not specifiable" uniquely picks out β. So, β is specifiable.

In the same year, Richard discovered a paradox similar to König's. Briefly, consider the class E of all decimals between 0 and 1 that are specifiable. E has at most $\aleph_0$ many elements and, thus, these must be listable. Pick some such list L. Now, specify a number N via the Cantor diagonal method: if x is the n^{th} decimal in the list L and the numeral in its n^{th} decimal place is **i**, then if the number denoted by **i** is less than 5, the numeral in N's n^{th} decimal place is '7'; otherwise, it is '2'. Clearly, N is different from every decimal in E. Since N is a specifiable decimal between 0 and 1, however, N must be an element of E.

Berry discovered another paradox concerning specifiability which has the merit of restricting its attention to whole numbers.[47] Informally, consider the class of all whole numbers not specifiable in fewer than nineteen syllables. Since the number of nineteen-syllable expressions formed from a finite alphabet is finite, this class must be non-empty. Moreover, since the whole numbers are well-ordered, there must be a least member j of the class. In this respect, "the least whole number not nameable in fewer than nineteen syllables" specifies or names j, but this description has fewer than nineteen syllables.

In 1907, Grelling discovered a paradox that concerns predicate application. He noted that some predicates apply to themselves while others do not. The predicate "polysyllabic" applies to itself whereas the predicate "monosyllabic" do not. Call a predicate that does not apply to itself *heterological*. Query: Is "heterological" heterological? Answer: It is iff it is not.

[47]Russell reports this paradox in his 1906 "On 'Insolubilia' and their Solution in Symbolic Logic" (henceforth, "On 'Insolubilia'") in EA, p. 210.

Ramsey in his FM. On the face of it, they do not clearly fall under either of the two groups that Ramsey describes. I shall address this matter in subsection (b). For the moment, call these paradoxes *mixed paradoxes*.

(b) The Nature of the Solution

From the point that Russell first committed himself to finding a solution to the paradoxes, he held two notable positions on the nature of such a solution.[52] First, the paradoxes 'arise' from

---

[52]Russell's commitment to the aim of finding a solution to the paradoxes took some time to develop. To begin with, neither Burali-Forti's nor Cantor's paradox spawned any serious reaction in general. Because these were both results derived from principles belonging to Cantor's recently developed transfinite class theory, they were considered to be inconsequential. Indeed, some mathematicians who looked askance at Cantor's research actually welcomed the paradoxes. In his MM, Russell mentions Cantor's paradox in passing. He writes something curious indicating that, like most, he did not treat it seriously:

> There is a greatest of all infinite numbers, which is the number of all things altogether, of every sort and kind. It is obvious that there cannot be a greater number than this, because, if everything has been taken, there is nothing left to add. Cantor has a proof that there is no greatest number, and if this proof were valid, the contradictions of infinity would reappear in sublimated form. But in this one point, the master has been guilty of a very subtle fallacy, which I hope to explain in some future work.[MM, p. 69]

Ironically, Russell discovered his class paradox while he was examining the proof of Cantor's paradox. His initial reaction to this discovery was similar to his reaction to Cantor's paradox:

> At first I thought there must be some trivial error in my reasoning. I inspected each step under a logical microscope, but I could not discover anything wrong.[MPD, p. 58.]

Soon, however, Russell and many others realised that his paradox could not be as easily dismissed as Cantor's and Burali-Forti's paradox. They took the class paradox to show that some of the received principles about the notions of class and even predication were mistaken. Thus, Frege eventually gave up his logicist project altogether. Dedekind temporarily stopped the publication of his celebrated "Was Sind und Was Sollen die Zahlen?" because of its uncritical appeal to the notion of class. Several mathematicians who had begun to accept class theory as a legitimate mathematical discipline and, in fact, had begun to contribute to it now rejected it -- Poincaré was among these.

In his 1903 *Principles*, Russell manifested a disquietude over his paradox as well as a sense of commitment to finding its solution. Although no definite solution is offered in the book, much of it is devoted to discussing the paradox and its possible solution. It is noteworthy that the book ends with the following exhortation:

> What a complete solution of the difficulty may be, I have not succeeded in

> our logic or mathematics. ... [The semantic paradoxes] are not purely logical, and cannot be stated in logical terms alone; for they all contain some reference to thought, language, or symbolism, which are not formal but empirical terms. So they may be due not to faulty logic or mathematics, but to faulty ideas concerning thought and language.[FM, pp. 171-2]

According to Ramsey, therefore, whereas the set-theoretic paradoxes are solved by correcting our logic, the semantic paradoxes are solved by changing the *empirical* assumptions ('ideas') -- presumably belonging to an empirical theory which treats of the semantic notions -- that give rise to them.

As is well-known, Ramsey's purpose in FM is to criticise Russell's solution to the paradoxes, *viz.*, his ramified theory of types. Towards this end, Ramsey complained that Russell did not respect the import that he attributed to Peano's distinction. As a result of Ramsey's remarks, many have thought that Russell failed even to recognise the distinction. However, Russell certainly *did* recognise it and he vividly illustrates such recognition in his "On 'Insolubilia'", after discussing a tentative solution -- his 'Substitutional Theory' -- to the set-theoretic paradoxes.[51] There, Russell writes:

> The above doctrine solves, so far as I can discover, all the paradoxes concerning classes and relations; but in order to solve the *Epimenides* we seem to need a similar doctrine as regards propositions.["On 'Insolubilia'", p. 204.]

According to Russell, therefore, some of the paradoxes crucially employ the notions of class and relation and others employ the notions of proposition and propositional function. Contrary to Ramsey, however, Russell does not take this consideration to be relevant to the paradoxes' solution for -- as we shall see in subsection (b) -- he has several good reasons for not doing so.

I have not yet said anything about how Russell's propositional function paradox and his proposition paradox are to be classified among the other paradoxes mentioned above, nor does

---

[51]Cf. Warren Goldfarb's "Russell's Reasons for Ramification", p. 28.

'logic' and, thus, their solution must consist in logic's reform. Secondly, all of the paradoxes arise from the same error -- for some notion of sameness of error -- and, so, their solution must be unitary. We shall look at why Russell held these two positions in what follows.

To begin with, the first position should be put more precisely. It may be read as saying that, for each of the modern paradoxes, there is some definite mischaracterisation of the logical system such that the argument to the paradox in question crucially appeals to it in some way.[53] Such a mischaracterisation of the logical system must consist in a mischaracterisation of at least one of the system's elements -- that is, its rules of formation, axioms. rules of inference, etc.. The solution to each paradox, therefore, must in turn consist in discovering which of these elements is (are) mischaracterised and how it (they) is (are) to be characterised correctly.

I shall explain why Russell holds this position in terms of the three groups of paradoxes described in subsection (a). Thus, I shall first explain why he takes the set-theoretic paradoxes to arise from logic, then I shall turn to consider the semantic paradoxes, and then, finally, the mixed paradoxes.

Consider the set-theoretic paradoxes: Burali-Forti's, Cantor's, and Russell's class paradox. Recall that these paradoxes all employ the notion of class and that their arguments all appeal to principles governing the existence of classes. For this very reason, most contemporary (post-Quinean) philosophers, contrary to Russell, deny that the set-theoretic paradoxes arise from logic. According to these philosophers, the notion of class is not a logical

---

> discovering; but as it affects the very foundations of reasoning, I earnestly commend the study of it to the attention of all students of logic.[*Principles*, §500.]

[53]Each paradox is here construed as the relevant contradictory proposition that is the conclusion of its argument. Thus, Russell's class paradox is the proposition that the class of all non-self-membered classes both is and is not a member of itself.

notion in the sense that there do not exist logical particles in terms of which the notion may be expressed. Rather, the notion of class is a special mathematical notion peculiar to that domain of discourse treated by set theory, in the way that the physical notion of momentum is peculiar to that domain of discourse treated by mechanics.[54] In this respect, according to these contemporary philosophers, the salient principles that govern the notion of class are not principles of logic -- that is, roughly speaking, they are not logically valid truths essentially involving the notion of class. Rather, they are merely deductively powerful, general truths about classes. As such, from the understanding that some of these salient principles -- such as the principle of unrestricted comprehension -- are appealed to in an *essential* way by the arguments to the set-theoretic paradoxes and that the other principles which are appealed to by these arguments are uncontroversial, these contemporary philosophers conclude that the set-theoretic paradoxes arise from the relevant, salient, non-logical principles about classes and, thus, take their solution to consist in the rejection of one or more of such principles. Accordingly, they never take the set-theoretic paradoxes to bring logic itself into question.[55]

By contrast, Russell explicitly takes the set-theoretic paradoxes to arise from logic -- or, rather, from its mischaracterisation. Unlike the contemporary philosophers alluded above, Russell early on considered the notion of class to be a logical notion in the sense that there exist logical particles in terms of which the notion may be expressed.[56] For instance, in *Principles*,

---

[54]The notion of class, as such, lacks the quality of so-called topic-neutrality characteristic of logical notions.

[55]It should be noted that some contemporary philosophers, most notably Quine, subscribe to some form of holism and, thus, claim that there is no fact of the matter about which principles, logical or non-logical, are responsible for the set-theoretic paradoxes.

[56]Although Russell early on treats the notion of class as a logical notion in this sense, he is not completely comfortable with the notion. As we shall see when we look at *Principles*, he perceives several difficulties with the notion many of which have nothing to do with the

Russell writes that the notion of class may be expressed in terms of the logical constant '$\in$' -- for example, if $\exists x(x \in y)$, then y is a class -- as well as in terms of the logical constant 'such that' and a variable that ranges over propositional functions.[57] With respect to the latter case, he writes:

> The values of *x* which render a propositional function true ϕx true are like the roots of an equation -- indeed the latter are a particular case of the former -- and we may consider all the values of *x* which are *such that* ϕx is true. In general, these values form a *class*, and in fact a class may be defined as all the terms satisfying some propositional function.
> [*Principles*, §23.]

In this respect, according to Russell, the salient principles that govern the notion of class -- more precisely, the salient principles that govern the logical particles in terms of which the notion of class may be expressed -- must be either axioms or theorems of the formal calculus in the way that, for instance, the salient principles that govern the truth-functional connectives are either axioms or theorems of the calculus. If this were not the case, the calculus would not be able to determine whether or not arguments which crucially employ the notion of class are logically valid. In *Principles*, Russell treats two such salient principles as axioms: these two are (equivalent to) the standard principles of comprehension and extensionality.[58]

In this light, it should be apparent why Russell takes the set-theoretic paradoxes to a arise from a mischaracterisation of the logical system. First, consider his class paradox. The

---

paradoxes. As a result, Russell later dispenses with the notion of class altogether.

[57]Russell considers the constants '$\in$' and 'such that' and the propositional variables to be logical particles, briefly, because these particles are required for the logical system to have both an adequate formal language and an adequate formal calculus in the senses described above in section 1. Cf. §§12 and 13 of *Principles*.

[58]See §§23 and 24.

formalisation of the argument to this paradox shows that only four interesting items are required. To repeat, they are the comprehension principle, the well-formedness of '$x \notin x$', universal instantiation, and existential instantiation. The comprehension principle is an axiom of the formal calculus. The well-formedness of '$x \notin x$' relies upon the primitive particles '$\neg$', '$\in$', and 'x', and the rules of formation. Universal and existential instantiation are two of the calculus's rules of inference. Thus, nothing of interest but items belonging to elements of the logical system are needed for the paradox. To solve Russell's class paradox, therefore, one must discover which elements are mischaracterised and how these should be characterised correctly. Secondly, consider the paradoxes of Burali-Forti and Cantor. The formalisations of the arguments to these paradoxes show that they also require these four items. To the extent that one or more of these items is mischaracterised and must be recharacterised in order to solve Russell's class paradox, such mischaracterisation may be taken to be responsible for these other paradoxes as well and the concomitant recharacterisation may be taken, likewise, to solve them. In this respect, the fact that they are concerned with transfinite ordinal and cardinal numbers is not as relevant to why they occur as generally had first been thought.

We now look at the semantic paradoxes: König's, Richard's, Berry's, Grelling's, and the Epimenides paradox. Recall that these paradoxes employ semantic notions like specifiability, applicability, and truth, and that their arguments appeal to special principles that govern these notions. For this very reason, contrary to Russell, Ramsey -- and others sympathetic to his analysis -- denies that the semantic paradoxes arise from a mischaracterisation of the logical system. As we saw earlier, according to Ramsey, these semantic notions are not logical notions but, rather, are theoretical notions that belong to some empirical theory of semantics. In this respect, according to Ramsey, the salient principles that govern the semantic notions are not principles of logic but, rather, are central principles of such a theory. As such, from the understanding that some of these salient principles are appealed to in an *essential* way by the arguments to the semantic paradoxes and that the other principles which are appealed to by these arguments are uncontroversial, Ramsey concludes that the semantic paradoxes arise, not from a mischaracterisation of the logical system, but from the relevant, salient, non-logical, semantic principles. As such, he takes their solution to consist, roughly speaking, in the rejection of one or more of such principles.

By contrast, Russell explicitly takes the semantic paradoxes to arise from a mischaracterisation of the logical system. There are three reasons for his doing so. The first reason is that, by his lights, the logical system itself contains substantive semantic principles and these play a crucial role in the arguments to the semantic paradoxes. Note that this first reason is not meant to be conclusive, but is only meant to jog the presupposition that logic must be innocent in this matter. The reason follows from Russell's conception of the logical system as discussed in section 1 above. Recall that, according to this conception, propositions and propositional functions are essentially *semantic* items in the sense that they are what various expressions of the formal language *mean*. Recall also that the logical system contains axioms, rules of inference, and rules of formation which crucially govern how such semantic items are to behave. It follows from this that Russell takes the arguments to the semantic paradoxes crucially to appeal to semantic principles of two kinds: principles that govern the semantic notions of proposition and propositional function and principles that govern the semantic notions of specifiability, applicability, and truth. Thus, the argument to the Epimenides paradox appeals to principles about the notions of proposition and truth; the argument to Grelling's paradox appeals to principles about the notions of propositional function and applicability; and the arguments to the paradoxes of König, Richard, and Berry appeal to principles about the notions of propositional function and specifiability.

The second reason that Russell has for taking the semantic paradoxes to arise from a mischaracterisation of the logical system is, roughly speaking, that Ramsey's own solutions to the semantic paradoxes are problematic. Before we look at why such solutions are indeed problematic, I should briefly review what Ramsey's specific solutions are.

In FM, Ramsey writes the following passage about König's, Richard's, and Berry's paradox.[59] The passage presents his diagnosis of why these paradoxes arise and his solutions to them.

> All of these result from the obvious ambiguity of 'naming' and 'defining'. The name or definition is in each case a functional symbol which is only a name or definition by meaning something. The sense in which it means must be made precise by fixing its order; the name or definition involving all such names or definitions will be of a higher order, and this removes the contradiction.[FM, p. 199.]

What Ramsey calls *naming* and *defining* I have been calling *specifying*, following Quine.[60] Roughly speaking, Ramsey here diagnoses that the paradoxes of König, Richard, and Berry arise from the presumption that there is a unitary semantic relation of specification and corresponding principles that govern this relation when, actually, there are an infinite number of such relations -- which may in a specific sense be seen to form a hierarchy -- as well as corresponding principles that govern these. By his lights, therefore, the arguments to these three paradoxes each turn on an equivocation between different specification relations and the solution to any of these must consist in making the relevant equivocation apparent.

This gloss on the above cited passage must be put more precisely. To begin with, consider what it is to say on a given occasion that x specifies y. Ramsey understandably takes it

---

[59]Although the following discussion focuses on these particular paradoxes, what is said applies *mutatis mutandis* to the other semantic paradoxes as well.

[60]See Quine, *Set Theory and its Logic*, p. 242.

to say that there is some description x constructed out of particles belonging to some fixed vocabulary E such that y uniquely satisfies x. In this respect, the expression 'x specifies y' is ambiguous and may be taken to mean on a given occasion one specification relation rather than another, depending on which fixed vocabulary E is presupposed on that occasion. Ramsey concludes from this consideration that, for any fixed vocabulary E_0, there exists an entire hierarchy of specification relations that is generated from E_0. This hierarchy is described as follows. At level (order) 0, there is the specification relation S_0 such that xS_0y iff x is a description constructed exclusively out of particles belonging to the vocabulary E_0 and y uniquely satisfies x. At level 1, there is the specification relation S_1 such that xS_1y iff x is an expression constructed exclusively out of particles belonging to $E_1 = E_0 \cup \{S_0\}$ and y uniquely satisfies x. Thus, $E_0 \subseteq E_1$ and $S_0 \subseteq S_1$. In this regard, an ω-ordered sequence of specification relations obtains such that, for all natural numbers i, j such that $i<j$, if $E_i \subseteq E_j$ and xS_iy, then xS_jy. Thus, $E_0 \subseteq E_1 \subseteq E_2 \subseteq \ldots$, and $S_0 \subseteq S_1 \subseteq S_2 \subseteq \ldots$.[61] In addition to this hierarchy of specification relations -- although Ramsey does not explicitly say so -- there must be principles that govern such relations (these will presumably belong to his alleged, empirical theory of semantics). Some of the principles will be disquotational in nature, having the form: $\forall y([F(y) \wedge \forall z(F(z) \rightarrow z=y)] \rightarrow cS_iy)$, where F is a some open formula, c is a Gödel number of F, and S_i is a specification relation appropriate to the vocabulary involved in F.

It may be helpful at this point to illustrate how Ramsey exploits such a hierarchy to block a given semantic paradox. Consider König's paradox. Recall its argument: Since there can be only $\aleph_0$ finite sequences of symbols, there can be only $\aleph_0$ ordinal numbers which are specifiable in terms of some fixed vocabulary E_0. However, there are certainly more than $\aleph_0$

---

[61]Here the relations S_i may be construed as sets of ordered-pairs.

ordinal numbers and, since the class of ordinal numbers is well ordered, there must be a least ordinal number β that is not specifiable. Hence, β is both not specifiable and specifiable. Now, mimicking this reasoning, Ramsey would say that given some E_0, there will be an entire hierarchy of specification relations S_i generated from E_0 such that, for any i, there will be a least ordinal number β that is not specifiable$_i$, where the notion of specifiability$_i$ is defined in the obvious way in terms of S_i and E_i. However, in virtue of the nature of the hierarchy, β will be specifiable$_{i+1}$, but since the notions of specifiability$_i$ and specifiability$_{i+1}$ are different, no contradiction actually results.

Let us now turn to consider why Ramsey's solutions to the semantic paradoxes are problematic. The reason is simple and is given by Goldfarb in his "Russell's Reasons for Ramification" (henceforth, RRR):

> Now Russell, I think, would ask why there is no relation that sums up (is the union of) the different [specification] relations that Ramsey postulates. (Such a union would reintroduce the paradox.)[RRR, p. 29.]

What Goldfarb in effect claims is that union operations render Ramsey's solutions ineffective by reintroducing contradiction. The argument for this claim may not be entirely apparent upon initial reflection and, so, in order to make good Russell's second reason, I shall try to spell it out. In this effort, I shall focus without loss of generality on König's paradox.

To begin with, Goldfarb's talk of the union of the specification relations presumes that there is a function that enumerates these relations. This presumption, however, is consonant with Ramsey's own discussion; thus, we may use the expression 'Spec' to designate this function.[62] In this regard, let $S = \cup Spec(i)$, where $i \in \omega$, so that xSy iff $\exists i(xSpec(i)y)$. Now, since the domain of S is countably infinite and the ordinal numbers are not, $\exists\beta\forall x\neg(xS\beta)$.

---

[62]Since the 'i' in 'S_i' above may be taken to be a schematic letter rather than a free variable, the use of 'S_i' above does not presume any function that enumerates the specification relations.

Since the ordinal numbers are well-ordered, there is a least such β -- call it β_1. Thus, $\{\forall x \neg(xS\beta_1) \wedge \forall\gamma[\forall x \neg(xS\gamma) \rightarrow \gamma \geq \beta_1]\}$ and for every ordinal number other than β_1, this statement is provably false. Hence, "$[\forall x \neg(xSy) \wedge \forall\gamma(\forall x \neg(xS\gamma) \rightarrow \gamma \geq y)]$" (ii) is provably true only of β_1. By disquotational principles, $cS\beta_1$, where 'c' is the Gödel number of the open formula (ii). So, $\exists x(xS\beta_1)$. But $\forall x \neg(xS\beta_1)$. Contradiction. Note that, so long as the required disquotational principles are available, this argument can be easily formalised.

The argument for Goldfarb's claim that union operations reintroduce contradiction might appear to make two moves that are open to Ramsey to challenge: the specification of S by an unrestricted union operation on the hierarchy of specification relations and the appeal to disquotational principles in order to infer $cS\beta_1$. With respect to the union operation, Ramsey could claim that, in the unrestricted form required to specify S, the operation is illegitimate and that, once properly restricted, it is unable to specify S. However, it is unclear how Ramsey could justify any such restriction. Since he does not take the semantic paradoxes to arise from a mischaracterisation of the logical system, he cannot appeal to type considerations to justify the restriction. The only other way of restricting the union operation is by restricting the axioms that govern it, but this method is conspicuously *ad hoc* and the result is inconsonant with the conception of the operation which the axioms are expected to embody. With respect to the appeal to the disquotational principles, Ramsey could claim that, although for every i, disquotational principles governing S_i -- where 'S_i' is a name and not a function expression -- must be available as part of an adequate semantic theory about specification, none of these principles need treat the function sign 'Spec' -- recall that 'Spec(i)' designates the same specification relation as 'S_i'. Thus, none of these principles need treat the term 'S' defined in terms of 'Spec'. Hence, no disquotational principle need be available to guarantee the required inference to $cS\beta_1$. However, to the extent that Ramsey must admit the relation S, which 'sums up' all the specification relations occurring in the hierarchy, the refusal to admit a semantic principle that governs S, *qua* semantic relation, would also appear conspicuously *ad hoc*. In this light, Ramsey's challenges, at least in the form given here, would be unsuccessful.

At this point, one may conclude that, to the extent that the above argument to the claim that union operations reintroduce contradiction is correct, Ramsey's solution to König's paradox is problematic. One may conclude, further, that, to the extent that Ramsey's solution to this paradox is typical of his solutions to all the other semantic paradoxes, his solutions to these are, likewise, problematic. Thus, to that extent, Russell has reason for taking the semantic paradoxes to a arise from a mischaracterisation of the logical system.

Before we look at Russell's third reason, I should address one question that one may now be prompted to ask. One might think that, even if Ramsey's solutions to the semantic paradoxes are problematic, surely the modern solutions are not and surely they do not suppose that these paradoxes arise from a mischaracterisation of the logical system. If this is so, does it not undermine Russell's second reason as it has here been explained? There are two points to make in response to this question. First, we have been looking at Russell's reasons for taking the semantic paradoxes to arise from a mischaracterisation of the logical system *in the light of* his conception of logic as discussed in section 1. Thus, insofar as Ramsey's conception of logic is comparable to Russell's in the way that the modern one is not, the consideration of the modern solutions to the semantic paradoxes does not in fact affect Russell's second reason in the way suggested. Secondly, if one examines closely the modern solutions to the semantic paradoxes, one finds that they in general divide into two separate kinds: the first kind is exemplified by the celebrated solution that Tarski put forward in his "Wahrheitsbegriff" and appeals to the objectlanguage/metalanguage distinction described in this paper;[63,64] the second kind is exemplified by Russell's own solution, his ramified theory of types, and appeals to some notion of ramification that is accomplished by restricting the quantifier rules -- note that, *qua* modern solution, the ramified theory of types need not rely upon Russell's particular conception of logic. Consider the first kind of modern solution. Curiously, although the objectlanguage/metalanguage distinction characteristic of this kind may be thought to be foreshadowed in Ramsey's hierarchy of specification relations, the way that it is standardly implemented by solutions of this kind is by means of a type-theoretic hierarchy as discussed in "Wahrheitsbegriff" and Gödel's "On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems I" (1931).[65] Such an implementation is required because, in order to define formally any of the semantic notions of specification, satisfaction, or truth for a particular language L_i, one must be able to quantify over (the equivalent of) binary relations between Gödel numbers of the formulae of L_i and (finite) sequences of the individuals, properties, and relations which are quantified over by the variables of L_i. Higher-order quantification, therefore, is needed. In this respect, modern solutions of the first kind actually appeal to resources that Russell would take to belong to logic and, accordingly, the consideration of them, rather than undermining Russell's second reason, in fact supports it. Now consider the second kind of modern solution. Church has recently defended this kind as a viable and attractive alternative to the first kind by pointing out the advantages that Russell's solution has over Tarski's "Wahrheitsbegriff" solution.[66] According to Church, one of these

[63]See "The Concept of Truth in Formalized Languages" in *Logic, Semantics, and Metamathematics*, (Indianapolis: Hackett, 1983).

[64]This distinction is to be found, if not explicitly, earlier in Hilbert's work.

[65]See "On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems I" (1931) in van Heijenoort, pp. 596-601. Note that the modern solutions to the semantic paradoxes also use 'coding' techniques that Gödel introduced in this paper.

advantages is that Russell's solution is more systematic than Tarski's in the sense that Tarski's solution requires assumptions that are overtly *ad hoc* in a way that those required by Russell's solution are not. Whether or not this is so, it is already clear that the consideration of Russell's solution and the kind of modern solution that it exemplifies, again rather than undermining his second reason, can only support it.

The third reason that Russell has for taking the semantic paradoxes to arise from a mischaracterisation of the logical system is, roughly speaking, that, although the arguments to the semantic paradoxes so far discussed all employ semantic notions like specifiability, applicability, and truth and appeal to special principles that govern such notions, it may be possible to generate other *semantic* paradoxes without having to rely explicitly on such notions and principles. That is, it may be possible to generate other semantic paradoxes from just the resources of the logical system itself.[67] Recall that, according to Russell, the logical system has variables that range over propositions and propositional functions. As such, the system by itself already has the resources to express, to a certain extent, semantic notions. For instance, if p_1 is a proposition and ψ is a propositional function that applies to p_1 and to nothing else, then the statement that p_1 is true can be expressed by the formula "$\exists p(\psi(p) \wedge p)$". Likewise, the statement that p_1 is false can be expressed by the formula "$\exists p(\psi(p) \wedge \neg p)$". Such expression is acceptable to Russell because he takes it that the variables that range over propositions can be used to *refer* to propositions as well as to *express* them.[68] In this respect, since the logical

[66]See Alonzo Church's "Comparison of Russell's Resolution of the Semantical Antinomies with that of Tarski" in *Journal of Symbolic Logic* 41 (1976) pp. 747-60. It should be noted that Church's defence is addressed to, among others, Quine who claimed that Russell's addition of his axiom of reducibility to the ramified theory of types revoked whatever security from contradiction the theory's ramification obtained. See Quine's *Set Theory and its Logic*, §35.

[67]By 'the logical system' here, I mean the system as it is characterised before Russell provides his solution to the paradoxes (the system implicit in *Principles*).

system by itself can, to a certain extent, express semantic notions such as truth and falsehood, perhaps one could generate semantic paradoxes from just the resources of the logical system, in a way that is analogous to the way that one generates them by explicitly employing the semantic notions and appealing to their special principles. If that is the case, such paradoxes would clearly have to be taken to arise from a mischaracterisation of the logical system. (Note that since Ramsey's proposed solutions to the semantic paradoxes do not focus at all on the characterisation of the logical system, they would be impotent in this particular situation.)

In his RRR, Goldfarb suggests one way that one could perhaps generate a version of the Epimenides paradox by appealing to just the logical system's resources: Consider the formula "$\forall p(\psi(p) \rightarrow \neg p)$" (iii).[69] It may be taken to express the proposition that any proposition that satisfies the propositional function ψ is false. Now, suppose that ψ is such that it is satisfied by and only by the proposition expressed by (iii). Let the proposition expressed by (iii) be q. Then, trivially, $q \rightarrow \forall p(\psi(p) \rightarrow \neg p)$. By universal instantiation, $\forall p(\psi(p) \rightarrow \neg p) \rightarrow (\psi(q) \rightarrow \neg q)$. Since q uniquely satisfies ψ, $(\psi q \rightarrow \neg q) \rightarrow \neg q$. Hence, by transitivity, $q \rightarrow \neg q$. However, since q is the proposition expressed by (iii), $\neg q \rightarrow \exists p(\psi(p) \wedge p)$. Since q uniquely satisfies ψ, $\exists p(\psi(p) \wedge p) \rightarrow (\psi(q) \wedge q)$. Hence, by transitivity again, $\neg q \rightarrow q$. Contradiction.

Goldfarb suggests that the crucial supposition that there is a ψ such that it is satisfied by q and only q is very weak. This suggestion, in fact, is confirmed by the following interesting circumstance: When Russell develops his 'substitutional theory' in 1905 and 1906, he introduces apparatus to the logical system that allows one to carry out Gödelian diagonalisation.

[68]Russell's view of propositional variables, in turn, relies on his further view that propositions are both, *qua* ordered complexes of terms, objects on a par with all other objects in the realm of being as well as the objects of judgement. This further view involves certain difficulties of interpretation, which we shall see in Chapter 2.

[69]Cf. p. 29.

By means of this diagonalisation, one *can* define (the equivalent of) such a ψ. In this respect, to the extent that it is legitimate to introduce such apparatus to the logical system, it is possible to generate this semantic paradox without explicit reliance on the semantic notions and their special principles. To that extent, these latter are not essential to the paradox and, therefore, solutions of Ramsey's kind are misdirected. Accordingly, Russell has reason to take the semantic paradoxes in general to arise from a mischaracterisation of the logical system.[70]

I now consider the mixed paradoxes: Russell's propositional function paradox and his proposition paradox. Before we look at why Russell takes these paradoxes to arise from a mischaracterisation of the logical system, I should first comment about why they do not *prima facie* fall under either of the two groups of paradoxes already discussed. Consider the propositional function paradox first. On the one hand, the argument to this paradox faultlessly mimics the argument to Russell's class paradox in the sense that the former argument can be easily obtained from the latter one, first, by replacing the latter argument's expressions for classes by expressions for monadic propositional functions and its expressions for membership by the notation for satisfaction and, then, by replacing the comprehension principle governing classes by one governing propositional functions. In addition, the propositional function paradox does not explicitly employ any of the semantic notions of specification, application, or truth. For these two reasons, one should feel reluctant to classify this paradox as a semantic paradox as opposed to a set-theoretic one. On the other hand, the argument to the propositional function paradox does not mention classes but instead mentions propositional functions. Recall that these are the *semantic* items meant by predicates and, more generally,

[70]Russell's own final solution, the vicious-circle principle and its attendant ramified theory of types, is such as to prevent the above argument from arriving at contradiction. According to this solution, q cannot fall within the scope of the variable p and, so, cannot satisfy the propositional function ψ.

open formulae. For this reason, one should feel reluctant to classify the paradox as a set-theoretic paradox as opposed to a semantic one. Next, consider Russell's proposition paradox. Recall that its argument mentions classes and appeals to a comprehension principle governing classes. To the extent that Burali-Forti's, Cantor's, and Russell's class paradox are classified as set-theoretic, in this regard, this paradox should be also. However, the argument to the proposition paradox also mentions propositions and the semantic notion of truth, and it appeals to an existence principle for propositions. To the extent that the Epimenides paradox is classified as semantic, in this regard, so should this paradox be. In this light, Russell's propositional function paradox and his proposition paradox strongly suggest that the two groups of paradoxes distinguished by Ramsey are neither mutually exclusive nor pair-wise exhaustive. They also further suggest that (from Russell's point of view) the distinction is not relevant to the paradoxes' solution.

We now turn to the question why Russell takes the propositional function paradox and proposition paradox to arise from a mischaracterisation of the logical system. From what we saw above concerning the set-theoretic and semantic paradoxes, the answer should already be clear. The formalisation of the argument to the propositional function paradox shows that only four interesting items are required: the comprehension principle for propositional functions, the well-formedness of x(x), universal instantiation, and existential instantiation. By Russell's light's, each of these items belongs to an element of the logical system and, thus, the paradox must arise from a mischaracterisation of at least one of these elements. Similarly, the formalisation of the argument to the proposition paradox shows that only the comprehension principle for classes, an existence principle for propositions, and trivial rules of inference are required. Since each of these items, by Russell's lights, belongs to an element of the logical system, this paradox must, likewise, arise from a mischaracterisation of at least one of these.

At this point, before we move on to consider Russell's second position on the nature of the paradoxes' solution, I should summarise what we have seen with respect to his first position. To repeat, the position is, roughly speaking, that all the modern paradoxes arise from logic. Russell's justification for this position may be put in terms of the three groups of paradoxes described in subsection (a). Briefly, he takes the set-theoretic paradoxes to arise from logic because he takes the notion of class to be a logical notion -- at least early on -- and because he takes the salient principles governing this notion to be logical principles. He has three reasons for taking the semantic paradoxes to arise from logic: logic contains substantive semantic principles; Ramsey's own solutions to the semantic paradoxes are problematic; and it may be possible to generate semantic paradoxes from just the resources of logic itself. Finally, Russell takes the mixed paradoxes to arise from logic for reasons similar to those concerning the set-theoretic and semantic paradoxes.

We now consider why Russell held the second position on the nature of the paradoxes' solution. To repeat, the position is that all the paradoxes arise from the same error -- for some appropriate notion of sameness of error -- and, so, their solution must be unitary. Russell has three interrelated reasons for holding the position. They are outlined in his "On Some Difficulties in the Theory of Transfinite Numbers and Order Types" (1905) (henceforth, "On Some Difficulties"), and are also discussed in later writing.[71]

The first reason for holding the second position is that, according to Russell, all of the arguments to the paradoxes have a particularly easily specifiable general form. He describes this form as follows:

---

[71]"On Some Difficulties in the Theory of Transfinite Numbers and Order Types" in EA, pp. 135-164. In this paper, Russell examines why the set-theoretic paradoxes arise and offers possible solutions to them. At this point, the semantic paradoxes had not yet entered the discussion.

(iv) Given a property ϕ and a function f, such that,if ϕ belongs to all the members of u, f u always exists, has the property ϕ, and is not a member of u; then the supposition that there is a class w of all terms having the property ϕ and that f w exists leads to the conclusion that f w both has and has not the property ϕ.["On Some Difficulties", p. 142.]

It is clear that if there is such a property ϕ, function f, and class w, contradiction quickly follows: Since every member of w satisfies ϕ, f w exists and satisfies ϕ. Conversely, since f w is not a member of w and w consists of all the items satisfying ϕ, f w does not satisfy ϕ. It is also clear that the arguments to Burali-Forti's, Cantor's, and Russell's class paradox have the general form (iv) that Russell describes. In the case of the argument to Burali-Forti's paradox, $\lambda x.\phi x$ is instantiated by λx(x is an ordinal number), and $\lambda u.fu$ is instantiated by λu(the ordinal number of u). In the case of the argument to Cantor's paradox, $\lambda x.\phi x$ is instantiated by $\lambda x(x=x)$, and $\lambda u.fu$ is instantiated by $\lambda u.\wp u$, the power-class function. In the case of the argument to Russell's class paradox, $\lambda x.\phi x$ is instantiated by $\lambda x(x \notin x)$, and $\lambda u.fu$ is instantiated by $\lambda u.u$, the identity function. However, it is not clear how any of the arguments to the semantic paradoxes instances the general form (iv). Although Russell does not explicitly consider these in "On Some Difficulties", when in later writings he does, he appears to take them to instance (iv) without explaining how.[72] By appealing to an interesting formalisation of semantics, perhaps it may be possible to show that the arguments to the semantic paradoxes, suitably represented, do indeed instance (iv).

The second reason for holding that all of the paradoxes arise from the same error is that there is a recipe for generating an endless number of paradoxes, all of which instance the general form (iv).[73,74] The recipe starts by specifying via transfinite recursion from a function f

[72]See "On Insolubilia", pp. 198-200; MPD, pp. 58-9.

[73]See "On Some Difficulties", pp. 142-3.

and a base point c a certain sequence that is isomorphic to On. Then, assuming that f satisfies certain conditions which many functions in fact do, the recipe deduces a contradiction. The recipe is sketched below:

(1) Pick a function f that is defined on all classes, and pick a class c.

(2) By definition via transfinite recursion from f and c, specify a well-ordered sequence as follows:

base: c is the first term.

successor: f c is the second term. For all other terms z of the sequence, the successor of z is f w, where w is the class of all terms in the sequence up to and including z.

limit: If u is an initial segment of the sequence ($\forall x \in u \forall y$(if y occurs in the sequence before x, then $y \in u$) with no greatest member, then the 'next' term of the sequence that occurs after all of the members of u is f u.

(3) Call the relation that orders the sequence specified in (2) *R*. By taking the ancestral of R, specify the property of being a member of the sequence.

(4) Assume that f satisfies the condition (*) that if u is a class of members of the sequence, f u is a member of the sequence that is not a member of u. Then, if w is the class of all members of the sequence, it follows that f w both is a member and is not a member of the sequence. Contradiction.

It should be clear that this recipe appeals to two substantive assumptions in order to generate the contradiction, both of which occur in (4): the assumption that f satisfies the condition (*) and the assumption of unrestricted comprehension needed to guarantee the existence of w. The first assumption need not be questioned, however, because there are various functions that one could take to be f which satisfy (*). Russell mentions two, along with corresponding base terms. One function, with its base term, generates the Burali-Forti contradiction: let $\lambda u.fu$ be λu(the ordinal of the class u) and let c be $\emptyset$, the empty set. Such an

---

[74]Note that this is not to say that the recipe generates all the paradoxes or even all the modern paradoxes discussed above. For instance, it does not generate Russell's class paradox, although -- as we shall see shortly -- it does generate a contradiction related to it.

f and c produce the sequence On of all ordinal numbers and, thus, f On∈On and f On∉On. The other function, with its base term, generates a contradiction related to Russell's class paradox: let $\lambda u.fu$ be $\lambda u.u$, the identity function, and let c be any well-founded class. Such an f and c produce a sequence S of terms, none of which is a member of itself, such that f S∈S and f S∉S. It is noteworthy that if c is ∅, the sequence S consists of the von Neumann ordinals; as such, Russell anticipates this important sequence long before von Neumann describes it (1923).[75]

Russell takes the recipe to be able to generate an infinite number of contradictions, although he never describes an effective procedure for obtaining the required number of f's and c's. Of course, there are trivial procedures for doing so: for instance, pick c to be successive members of the von Neumann sequence of ordinals. Such trivial procedures, however, only yield trivial variation between the contradictions that are generated by the recipe. One may perhaps object that they do not yield enough variation for Russell to take the existence of the recipe to be a good reason for holding the second position. In light of the limitative results of Gödel and Church, one may perhaps object further that it is doubtful whether effective procedures that are in this sense non-trivial exist. Even if these objections are correct, one may still say that the recipe has at least this merit: it may naturally be viewed as describing prototypically the kinds of important processes referred to by the third reason for holding the second position -- a reason that does not suffer this apparent difficulty.

This third reason for holding that the paradoxes all arise from the same error is the following:

> ... the contradictions result from the fact that, according to current logical assumptions, there are what we may call *self-reproductive* processes and classes. That is, there are some properties such that, given any class of terms

---

[75]von Neumann, *Zur Einführung der transfiniten Zahlen*. Acta Sci. Math. (Szeged) *1*, 199-208.

> all having such a property, we can always define a new term also having the property in question. Hence we can never collect *all* the terms having the said property into a whole; because, whenever we hope we have them all, the collection which we have immediately proceeds to generate a new term also having the said property.["On Some Difficulties", p.144.]

According to Russell, therefore, some mischaracterisation of logic gives rise to *self-reproductive* processes of the kind that he describes, which in turn give rise to the paradoxes. These processes are self-reproductive in the sense that they are iterative and, thus, can be thought of as realised in a sequence of 'stages.' For any one of these processes, at any given stage, "given [a] class of terms all having [the] property [associated with the process], we can always define a new term also having the property in question." The definition of this new term takes us to the next stage.

There are two points to note about these self-reproductive processes. First, in light of this consideration of stages, the description of them cited above is similar to the description of the recipe, although the former emphasises an aspect of *circularity* that the latter does not. In this respect, the two sequences specified in the above discussion of the recipe may be viewed as paradigms of such processes. Secondly, although Russell in "On Some Difficulties" intends to appeal to these processes specifically in order to characterise how the set-theoretic paradoxes arise, they clearly may be appealed to to characterise how all of the paradoxes arise. Indeed, Russell does exactly this later in "On 'Insolubilia'", ML, and *Principia* itself. In these works, however, the appeal to self-reproductive processes is spelled out more precisely as his vicious-circle analysis. In this respect, Russell's third reason for holding that the paradoxes all arise from the same error -- *viz.*, that their arguments all appeal to self-reproductive processes of a certain sort -- is one that deserves considerable respect and attention. Since Russell's initial investigations into the matter, it indeed has received this.

Before moving on the next chapter, I should summarise what we have just seen with respect to his second position on the nature of the paradoxes' solution. To repeat, the position is, roughly speaking, that all the modern paradoxes arise from the same error. Russell has three reasons for holding this position. The first reason is that, according to Russell, the paradoxes all instance an easily specifiable general form. We saw, however, that it is not entirely clear whether or not the semantic paradoxes do in fact instance this form. His second reason is that there is a recipe for generating an infinite number of paradoxes. We saw here that it is doubtful whether the paradoxes which this recipe generates vary enough in character for the existence of the recipe to count as a good reason. Finally, the third reason is that all the arguments to the paradoxes appeal to certain 'self-reproductive processes.' The appeal to such processes later becomes spelled out as Russell's vicious-circle analysis. It is noteworthy that this analysis may be seen to lead to his eventual solution to the paradoxes, *viz.*, his ramified theory of types. I examine the analysis in Chapter 3.

Chapter 2

The Logical Theory of *The Principles of Mathematics*

As a result of his exposure to the work of G. Peano and G.E Moore, Russell produced his *Principles of Mathematics, volume 1*. In this volume, he put forward the beginnings of a logical theory whose precise details he intended to explain in a companion volume 2. Before Russell could start to work on this companion volume, however, he discovered his class paradox and several logicians discovered other paradoxes soon afterward. Accordingly, he postponed the work on the companion volume and devoted himself instead to finding the solution to the paradoxes. After much effort, in 1906 he came upon what he took to be their solution, *viz.*, his celebrated *vicious-circle principle*. Shortly afterward, Russell began work on the intended companion volume and, as is well-known, this work culminated in the publication of *Principia Mathematica*.[76]

Not surprisingly, the logical theory put forward in *Principia* to a certain extent resembles the logical theory implicit in *Principles*. Both theories have the same items as their subject matter -- that is, individuals, propositional functions, and propositions -- and, roughly speaking, share most of their basic principles. Indeed, the logical theory of *Principia* -- call it *LT(PM)* -- may be understood as the result of modifying the logical theory implicit in *Principles* -- call it *LT(POM)* -- in order to abide by the vicious-circle principle. In this light, in order to understand LT(PM) one should first understand LT(POM) as well as the vicious-circle principle. In fact, when Russell presents LT(PM) in *Principia*, he appears to assume that his audience is already familiar with LT(POM) for he elaborates very little on the nature of several

---

[76]See the Preface to *Principia*, p. v.

of LT(PM)'s important elements -- for example, its individuals, propositional functions, propositions, and logical constants.

Toward the end of understanding LT(PM), I briefly examine LT(POM) in this chapter (I examine Russell's vicious-circle principle in the next chapter). In this effort, I confine my attention both in breadth and in depth. That is, I only examine those aspects of LT(POM) that are relevant to understanding LT(PM) and, even when examining such aspects, I do so only to the extent that it is relevant in this respect. Since Russell's discussion of LT(POM) in *Principles* is extremely rich, there is much to LT(POM) to which my effort cannot be thought to do justice.

1. The Ontology of LT(POM)

In Part I of *Principles*, entitled "The Indefinables of Mathematics," Russell puts forward an ontological view that describes the ontology he takes LT(POM) to have, his realm of being. Peter Hylton in his RIAP aptly calls this ontological view *Platonic Atomism*. The realm of being it describes is atomistic in two respects. First, every 'term' -- Russell's word meaning object in general -- in the realm of being is either absolutely simple -- that is, it is not composed of anything else -- or it is obtained by an operation of composition ultimately starting from other terms that are simple. Secondly and more significantly, every term in the realm of being, roughly speaking, is what it is -- has the nature that it does -- independently of the circumstances of every other term in the realm of being. More precisely, every simple term is what it is independently of the circumstances of every other term and every compound term is what it is independently of the circumstances of every term that is not a constituent of it. Russell made this claim by saying that all relations are external. As is well-known, he did so as part of his rejection of British Idealism which claimed that all relations are internal -- in other words, every term is what it is in virtue of every relation it bears to every other term. The realm of being Russell's ontological view describes is platonistic in the sense that certain of its terms are abstract as opposed to concrete. Numbers, classes, properties, and relations are among such abstract terms. Russell advocated that the realm of being countenance abstract terms for the same reason that Frege did: they are required to give an account of mathematical discourse.

Like any logical theory, LT(POM) contains several logical categories and each term in the realm of being belongs to one or more of such categories according to the logical role(s) it may play. Among others, there are the categories of thing, concept, propositional function, denoting concept, class, proposition, variable, and logical constant. As the categories of thing, concept, propositional function, proposition, variable, and logical constant play constitutive roles in LT(PM), I focus on them in the following.

(a) Propositions[77]

Propositions play two important roles in LT(POM). First, they are 'complexes' of terms and, as such, are themselves terms in the realm of being on a par with all other terms in the realm of being. As complexes, they may be thought of as ordered in the sense that only certain orderings of words of certain kinds constitute clauses that mean propositions and different orderings of the same words constitute different clauses that may mean different propositions. Secondly, propositions are the objects or vehicles of judgement. As such, they are what we judge, understand, suppose, know, assert, and so on. Crucially, Russell claims that, save for a few notable exceptions[78], what a proposition is about is *in* the proposition itself and, so, when we

[77]Although I shall not discuss the matter, Russell's account of propositions descends directly from G.E. Moore's account.

judge, suppose, or know a particular proposition, we are judging, supposing, or knowing about the terms that belong to the complex that is that proposition.

One may note the stark contrast between this view of Russell's and Frege's more familiar view. According to the latter view, what we judge, suppose, or know are thoughts and when we judge, suppose, or know a particular thought, we are judging, supposing, or knowing about the referents of the senses that compose that thought. Such referents do not in any way -- on most readings -- figure in the composition of the thought.

This stark contrast was actually a subject taken up in correspondence between Frege and Russell. Thus, in a response to Frege's comment that:

> ... Mont Blanc with all its snowfields is not itself a component part of the thought that Mont Blanc is more than 4000 meters high.[79]

Russell famously writes:

> I believe that in spite of all its snowfields Mont Blanc itself is a component part of what is actually asserted in the proposition 'Mont Blanc is more than 4000 meters high'. ... we assert the object of the thought (an objective proposition, one might say) in which Mont Blanc is itself a component part.[80]

The contrast may be highlighted by considering the predicative part of the proposition in question, *viz.*, the property of being more than 4000 meters high. By Russell's lights, if this property is a term in the realm of being, the proposition in question is the complex of Mont Blanc and this property. That the property itself is a component of the proposition gives rise to several problems to which Frege's view on the face of it does not. Among others, there is the

---

[78]See the discussion of Russell's theory of denoting concepts below.

[79]Letter to Russell dated 13 November 1904. *Gottlob Frege: Philosophical and Mathematical Correspondence*, ed. B. McGuiness (Chicago: The University of Chicago Press, 1980), p. 163.

[80]Letter to Frege, dated 12 December 1904. *op. cit.*, p. 169.

problem of explaining in what the unity of the proposition consists. Russell states the problem in a rather sharp way in *Principles*:

> The twofold nature of the verb, as actual verb and as verbal noun, may be expressed, if all verbs are held to be relations, as the difference between a relation in itself and a relation actually relating. Consider, for example, the proposition "A differs from B." The constituents of this proposition, if we analyze it, appear to be only A, Difference, and B. Yet these constituents, thus placed side by side, do not reconstitute the proposition. The difference which occurs in the proposition actually relates A and B, whereas the difference after analysis is a notion which has no connection with A and B. It may be said that we ought, in the analysis, to mention the relations which difference has to A and B, relations which are expressed by *is* and *from* when we say "A is different from B." These relations consist in the fact that *A* is referent and *B* relatum with respect to difference. But "*A*, referent, difference, relatum, *B*" is still merely a list of terms, not a proposition. A proposition, in fact, is essentially a unity, and when analysis has destroyed the unity, no enumeration of constituents will restore the proposition. The verb, when used as a verb, embodies the unity of the proposition, and is thus distinguishable from the verb considered as a term, though I do not know how to give a clear account of the precise nature of the distinction.[*Principles*, §54.]

As we shall see in the next chapter, Russell will change his view slightly. In response to the vicious-circle principle, he will conclude that a term that plays a predicative role in a given proposition cannot figure in that proposition in the same way as a term that plays a nonpredicative role in it.

Those who are familiar with Russell's later 1918 theory of *Logical Atomism* may now be prompted to ask how LT(POM)'s propositions relate to facts. Recall that the 1918 theory countenances facts as items distinct from propositions[81] and it says that a proposition is true if and only if there exists a fact with which it may correspond. Truth is thus definable in terms of such a correspondence and, accordingly, the 1918 theory cannot countenance 'false' or merely possible facts. Recall also that the 1918 theory says that the world is simply the totality of all

---

[81]According to the 1918 theory, a proposition is an indicative sentence, not what such a sentence means.

the facts that there are. LT(POM) differs rather sharply from this later theory. To begin with, LT(POM) does not countenance facts as items distinct from propositions. Rather, it has as a logical constant the property of being true and, as such, this property is an *indefinable* term in the realm of being.[82] Facts may be defined in terms of this property as follows: If the property happens to apply to a given proposition, then and only then that proposition is true and may be taken to be a(n actual) fact. In this respect, to the extent that the world is simply a totality of facts, by LT(POM)'s lights, the totality in question must be all the true propositions and none of the false ones.

It is noteworthy that Russell's view has one consequence about propositions that, for different reasons, Frege, the early Wittgenstein, and more contemporary philosophers of language all claim, *viz.*, that a proposition, *qua* meaning of a sentence, has the nature that it does irrespectively of whatever else may be the case -- in particular, irrespectively of its truth value. This consequence follows from the circumstance that a proposition is a term in the realm of being and that every term in the realm of being has the nature that it does irrespectively of the circumstances of every other term.

So far, we have not looked at the requirements that the terms that make up a proposition must satisfy in order that they may do so. I address this matter in what follows.

(b) Things and Concepts.

Here I consider what Russell says about the internal structure of what might be called *atomic propositions*. These are propositions that are not composed of simpler propositions; that is, they are propositions none of whose proper parts is a proposition.

---

[82]See *Principles*, §12, "On the Nature of Truth and Falsehood" in *Philosophical Essays*. Russell's view about the role of truth in LT(POM) descends from what Moore says about it.

In §48 of *Principles*, Russell writes that any atomic proposition may be understood to be composed of terms of two kinds, one term that plays a predicative role in the proposition in question and one or more terms that play a nonpredicative role in the proposition. In general, when such an atomic proposition is expressed by a natural language sentence, the term playing the predicative role in the proposition will be correlated to the sentence's verb or adjective(s) and the term(s) playing the nonpredicative role will be correlated to the sentence's noun(s) and/or noun phrase(s).

Significantly, Russell writes that any term that plays a predicative role in an atomic proposition may play a nonpredicative role in some other atomic proposition. Any such term Russell calls a *concept*. He further writes that there are terms in the realm of being that play and can only play a nonpredicative role in atomic propositions. Any such term he calls a *thing* (from now on I shall write 'Thing' as opposed to 'thing' to distinguish Russell's technical usage from ordinary usage). Russell's justification for this claim seems to be similar to Aristotle's justification for the claim that substances exist.

In an effort to illustrate the above claims, Russell in §48 considers the proposition that Socrates is human. Clearly, in this proposition Socrates is a term that plays a nonpredicative role whereas the property of being human is a term that plays a predicative one. According to Russell, Socrates is such that he could never play anything but a nonpredicative role in a proposition and, therefore, he must be a Thing. The property of being human, on the other hand, is a concept, for there are atomic propositions such as the one considered here in which it plays a predicative role and there are atomic propositions in which it plays a nonpredicative role as well -- Russell offers the proposition that humanity belongs to Socrates as an instance of the latter kind.

One may note here yet another point of contention between Russell and Frege. In contrast to Russell, Frege famously claims in his "On Concept and Object" that whatever plays a predicative (or unsaturated) role may not play a nonpredicative (or saturated) role as well. Ironically, in their arguments toward their respective conclusions on the matter, Russell and Frege both make appeal to the conditional that if it were not the case that a predicative item could play a nonpredicative role, then one could not even say of the predicative item in question that it plays a predicative role. However, whereas Russell *contraposes* after such an appeal, Frege simply *detaches*.

It is noteworthy that Russell writes very little about the nature of Things and concepts that is at all specific and, of the terms in the realm of being with which we may be familiar, he never identifies the ones that he takes to be Things and the ones that he takes to be concepts. This paucity of detail actually serves his purpose since the categories of Thing and concept are logical categories of LT(POM) and, *qua* logical theory, LT(POM) should be indifferent to the specific features of the world.[83]

(c) Variables

LT(POM)'s variables play an integral role in the construction of what might be called *compound propositions*. These are propositions that are not atomic. They also play an integral role in the construction of propositional functions.

Russell's account of LT(POM)'s variables is very complicated and thus difficult to explain. Keeping our stated end in mind, I only offer the following brief remarks. First, as we saw in chapter 1, §1, the variables are actual terms in the realm of being on a par with all other terms. They are not symbols but terms meant by symbols. Secondly, they are all completely

[83]However, since LT(POM) may be construed as a theory describing the most general features of the realm of being, this last claim may be difficult to spell out in any clear way.

unrestricted and, so, they all have the same universal range of variation. In this respect, there is only one kind of variable and, indeed, Russell talks about *the* variable:

> By making our *x* always an unrestricted variable, we can speak of *the* variable, which is conceptually identical in Logic, Arithmetic, Geometry, and all other formal subjects.[*Principles*, §88]

However, there must be infinitely many tokens of this one kind -- in the sense that the variables x and y occurring in Rxy are different tokens -- since propositions and propositional functions may contain any number of variables in them. Thirdly, although the range of a variable may be infinite, the variable itself must have only finite complexity. This follows from Russell's theory of denoting concepts explained in Chapter V of *Principles*. Roughly speaking, according to this theory, every term that figures in a proposition must have only finite complexity since otherwise our 'finite' minds would not be able to comprehend it. However, there are propositions that clearly talk about terms having infinite complexity such as the class of natural numbers and, so, there must be terms having only finite complexity that figure in these propositions and represent or 'denote' such other terms -- note here that this circumstance constitutes an exception to the general claim that a proposition contains what it is about. For instance, Russell takes the term *all numbers* to be one of these terms having only finite complexity:

> With regard to infinite classes, say the class of numbers, it is to be observed that the concept *all numbers*, though not itself infinitely complex, yet denotes an infinitely complex object. This is the inmost secret of our power to deal with infinity. An infinitely complex concept, though there be such, can certainly not be manipulated by the human intelligence; but infinite collections, owing to the notion of denoting, can be manipulated without introducing any concepts of infinite complexity.[*Principles*, §72]

Not surprisingly, Russell also takes the variable to be another of the terms having only finite complexity that figures in propositions and represents or denotes an infinite collection, namely, its infinite value range.[84]

(d) Logical Constants.

Like LT(POM)'s variables, LT(POM)'s logical constants play an integral role in the construction of compound propositions and in the construction of propositional functions. Roughly speaking, LT(POM)'s logical constants include among others the universal quantifier[85], material implication, membership, class abstraction, and truth.[86] Unfortunately, Russell says very little about these other than that they are *indefinable*. Indeed, because they are indefinable, little can be said about them:

> The logical constants themselves are to be defined only by enumeration, for they are so fundamental that all the properties by which the class of them might be defined presuppose some of the terms of the class. But practically, the method of discovering the logical constants is the analysis of symbolic logic ... [*Principles*, §10]

The circularity that Russell alludes to here is obviously the circularity that Quine talks about in "Truth by Convention"; the circularity such that for any logical constant, the specification of its semantics will employ that logical constant or others that are likewise circularly defined.[87]

---

[84]Strictly speaking, the terms in question having only finite complexity, which Russell calls *denoting concepts*, are the terms meant by locutions of the following form: an F, any F, every F, all F, some F, and the F. In general, every denoting concept stands in an indefinable relation of *denoting* with some other term in the realm of being. Whenever any of these concepts occurs in a proposition, the proposition will not be about the concept in question but about the term that it denotes. The variable is an integral part of any denoting concept signalled by 'any F'.

[85]Actually, in *Principles*, Russell writes very little that is explicit about the universal and existential quantifiers. Rather, he writes extensively about what he calls *formal implication*, which is exhibited in propositions having the form: $\forall x[\phi(x) \rightarrow \psi(x)]$. Obviously, universal and existential quantification can be defined in terms of formal implication.

[86]See *Principles*, §12.

[87]Cf. *Principles*, §16.

In spite of the remark cited above, one can make a few comments about the logical constants. First, like the variables of LT(POM), these are terms on a par with all other terms; they are not symbols. Unfortunately, many have interpreted Russell's talk of logical constants -- as well as his talk of variables and terms -- as talk about symbols and, as a result, have accused him of making use/mention errors in several places where he actually does not. What is worse, many have wholly misconstrued Russell's logical theory as a result. Secondly, the logical constants of material implication and the universal quantifier are functions. In particular, material implication is a function from pairs of propositions to propositions, from pairs of propositional functions to propositional functions, or from pairs of propositional functions and propositions -- where only one member of each pair is a propositional function -- to propositional functions; the universal quantifier is a function from propositional functions to either propositions or propositional functions, depending on the number of distinct free variables occurring in the argument to this function. Moreover, in each case the function involved in some sense figures in the values of the function. Thus, where the propositions p and q are arguments to the function of material implication, the value of this function, $p \rightarrow q$, contains p, the function, and q. LT(POM)'s conception of material implication and the universal quantifier, accordingly, differs markedly from more contemporary conceptions of these logical constants. However, LT(POM)'s conception resembles the more contemporary ones to the extent that it takes compound propositions whose main operators are either material implication or the universal quantifier to be evaluated as true or false in more or less the same way in which these others do. Thus, by LT(POM)'s lights, the proposition $p \rightarrow q$ is true just in case p is false or q is true and the proposition $\forall x.\phi(x)$ is true just in case $\phi(x)$ is true for every value of x. Thirdly, the logical constant of material implication is *also* an important logical relation. Thus, Russell writes:

> But it is plain that where we validly infer one proposition from another, we do so in virtue of a relation which holds between the two propositions whether we perceive it or not: ... The relation in virtue of which it is possible for us validly to infer is what I call material implication.[*Principles*, §37]

By Russell's light's, therefore, a proposition p stands in the relation of material implication to a proposition q if and only if q is a logical consequence of p. Fourthly, although it may not be apparent, all the other usual logical constants may be defined in terms of material implication, the universal quantifier, and truth. The reason is that truth, *qua* property of propositions, may be employed in much the same way in which some logical theories employ the *false*. As such, negation may be defined thus: $\neg p \ .{=}._{df} \ \forall r[(r \text{ is a proposition}) \rightarrow (p \rightarrow T(r))]$.[88] Conjunction, disjunction, equivalence, and existential quantification may be defined similarly.

(e) Propositional Functions

Here I consider the terms in the realm of being that Russell calls *propositional functions*. From the discussion above, it should already be clear that such terms figure within compound propositions rather in the way in which open formulae figure in containing sentences. Indeed, such terms are the semantic analogues of open formulae.

Russell's account of propositional functions is not easy to comprehend. Indeed, some of its claims are not obviously mutually consistent. Here I only offer a selection of the account's most salient claims: First, the notion of propositional function is a generalisation of the notion of concept. One may roughly define the notion as follows: x is a propositional function iff x is a concept or x is a term built up from concepts by means of the logical operators (the connectives and quantifiers).[89] In this respect, whereas a concept is to be

[88]Russell often drops the predication of truth and simply writes the variable ranging over propositions: $\forall r[(r \text{ is a proposition}) \rightarrow (p \rightarrow r)]$.

understood as a simple term in the realm of being which may play a predicative role -- for instance, the concept of being red as Russell later considers it -- a propositional function is to be understood as this as well as any arbitrarily complex term which may play a predicative role -- for instance, $\lambda x \lambda y. x<y$ as defined in terms of the ancestral of the successor relation. Secondly, propositional functions are *functions*. That is, they are functions from Things, propositional functions, and propositions to propositions and, in spite of these unusual arguments and values, they are functions in the same sense in which mathematical functions are. Indeed, Russell later writes that the functionality of any mathematical function derives from its being, roughly speaking, constructed out of some propositional function.[90] Thirdly, a propositional function is the result of taking a proposition, *qua* ordered complex of terms, and replacing one or more of such terms with one or more variables. Thus, Russell writes:

> $\phi(x)$, the propositional function, is what is denoted by *the* proposition of the form ϕ in which x [the variable] occurs. [*Principles*, §86]

and

> Given any proposition (not a propositional function), let a be one of its terms, and let us call the proposition $\phi(a)$. Then in virtue of the primitive idea of a propositional function, if x be a [variable], we can consider the proposition $\phi(x)$, which arises from the substitution of x in place of a.[*Principles*, §93]

One consequence of this claim is that insofar as something cannot be both a proposition and propositional function, a proposition cannot contain a(n unbound) variable as one of its terms.[91] A variation on this third claim is that a propositional function is the result of taking a proposition and simply extracting at least one of its terms. In §482 of *Principles*, Russell

---

[89]This is somewhat inaccurate. I offer a precise recursive definition in the next section.

[90]See *Principia*, *30, p. 232.

[91]The significance of this consequence is mitigated by the fact that Russell claims that propositional functions are terms that can be asserted *simpliciter*.

colourfully calls such a result the *rump* of the proposition. Fourthly, a propositional function is essentially an ambiguity.[92] This claim coheres with the third claim insofar as a variable is considered as an ambiguity and propositional functions contain variables. However, it is clear that in this instance Russell is committing a genuine use/mention error. The expression '$\phi(x)$' is ambiguous in that it means either the proposition $\phi(a)$ or the proposition $\phi(b)$ or the proposition $\phi(c)$, depending on whether the sign 'x' means either a, b, or c. Nevertheless, insofar as the expression '$\lambda x.\phi(x)$' means a particular term in the realm of being which is a function having among its values $\phi(a)$, $\phi(b)$, and $\phi(c)$, '$\lambda x.\phi(x)$' is not ambiguous and it makes little sense to say of what it means, $\lambda x.\phi(x)$, that it is ambiguous. Fifthly, propositional functions are intensional terms in the sense that coextensive propositional functions need not be identical. This claim follows from the claim that propositional functions are functions from terms to propositions. Thus, the propositional functions ϕ and ψ may be such that $\forall x(\phi(x) \leftrightarrow \psi(x))$, yet ϕ be not identical to ψ.

(f) Classes

As I suggested above, LT(POM) countenances classes as terms in the realm of being. More precisely, LT(POM) guarantees their existence twice over. First, it contains an unrestricted comprehension axiom that says that every propositional function $\lambda x.F(x)$ determines a class $\{x:F(x)\}$. Secondly, it contains denoting concepts and according to the theory of these concepts, if F is a propositional function, then *all F's* is a denoting concept that denotes *the* class $\{x:F(x)\}$. Any proposition containing *all F's*, rather than being about this concept, will be about such a class.

---

[92]See *Principia*, pp. 38-39.

Russell makes many interesting claims about LT(POM)' classes in *Principles*, but since LT(PM) does not countenance classes at all, we need not consider them.

2. The Logical System of LT(POM)

In this section, I examine LT(POM)'s rules of formation, axioms, and rules of inference. Two points should be noted from the start. The first point is that whereas Russell writes extensively in *Principles* about the nature of the various logical categories of terms belonging to the ontology of LT(POM), he writes little that is explicit about LT(POM)'s rules of formation, axioms, and rules of inference. Thus, unlike the above description of the nature of LT(POM)'s ontology, much of what I say here is at best a reconstruction of what may be taken as implicit in *Principles*. The second point is that when Russell does write somewhat explicitly about these elements of LT(POM), he rarely focuses his attention on syntactic considerations. Rather, as always he focuses it on the realm of being. Thus, instead of writing about the rules of formation for formulae and sentences, he writes about the rules of formation for propositional functions and propositions. Russell is to a certain extent justified in doing so because -- as we saw in Chapter 1, §1 -- he takes the formal language in question to mimic perfectly the realm of being that it is about. He takes it to do so in such a way that, for instance, whatever claims he makes about propositions and propositional functions should be easily translated to analogous claims about sentences and formulae.[93] To a certain extent, I shall describe the elements of LT(POM)'s logical system with Russell's particular focus.[94]

---

[93]Note that this circumstance coheres well with Russell's claim that, in general, a proposition contains what it is about.

[94]One may note a further contrast between Russell and Frege in the light of these two points.

(a) Rules of Formation for Propositions and Propositional Functions.

Since I have already discussed to a certain extent the nature of atomic propositions, I begin by looking at the formation of these. Then, after considering atomic propositional functions, I look at compound propositions and propositional functions. Finally, I examine two complications.

(i) Atomic Propositions:

As we saw above, there are two structural requirements that an atomic proposition of LT(POM) must satisfy. The first requirement is that an atomic proposition of LT(POM) must consist of the predication of some n-ary concept to a sequence of n terms. That such predication may be polyadic was of course not an insignificant detail at the turn of the century. The second requirement is that an atomic proposition of LT(POM) must respect the distinction between Things and concepts. Recall that according to this distinction, a concept may play either a predicative or nonpredicative role (or both) in an atomic proposition, whereas a Thing may only play a nonpredicative role in such a proposition. Thus, let c and d be Things, E a singulary concept, and F a binary concept. Then the following structures are atomic propositions by LT(POM)'s lights (braces indicate predication): E(c), E(d), F(c,d), E(F), F(E,c), E(E), F(E,F); and the following structures are ill-formed: c(E), c(d), E(c,d), F(d). Note that the concepts E and F occurring in the above atomic propositions of LT(POM) are less constrained in the way in which they do than are the predicates occurring in the atomic formulae of first-order predicate logic. Accordingly, LT(POM) is more type-free than this more contemporary logical theory.

(ii) Atomic Propositional Functions:

Atomic propositional functions may be specified by reference to the third claim made above about propositional functions. In particular, an atomic propositional function may be understood as the result of taking an atomic proposition and replacing one or more of its terms with one or more variables. Thus, let x, y, and z be variables. Then the following structures are atomic propositional functions of LT(POM): x(y), x(z), x(y,z), x(x), y(y,y), E(x), x(d), F(x,y), F(E,z). Clearly, the variables occurring in these atomic propositional functions are even less constrained in the way in which they do than are the Things and concepts occurring in the atomic propositions illustrated just above. The reason is that since all the variables range over every term in the realm of being, no distinction analogous to the distinction between Things and concepts applies to them, nor does any distinction analogous to the distinction between m-ary concepts and n-ary concepts, where $m < n$. This circumstance only serves to highlight the degree of type-freedom enjoyed by LT(POM).

(iii) Compound Propositions and Propositional Functions:

Like Frege's logical theory and unlike earlier logical theories, LT(POM) employs iteration in the characterisation of compounds. Thus, compound propositions and propositional functions are defined inductively. The inductive rules are given as follows (where appeal is made to the standard notions of freedom and bondage): Let ϕ and ψ be propositions or propositional functions. Then $\phi \neg \psi$ is a propositional function if one of ϕ or ψ is; otherwise, it is a proposition. Let $\pi(x)$ be a propositional function with the variable x's occurring free -- this condition is required because, officially, Russell does not countenance vacuous quantification. Then $(x)\pi(x)$ is a proposition if all of its variables are bound; otherwise, it is a propositional function. I should note here that, following Peano, Russell calls free variables *real* and bound ones *apparent*.

(iv) Complications:

There are two complications that the above rules of formation do not address. The first complication concerns propositional functions, *viz.*, that no resources have been offered by which the variables occurring in a propositional function may be distinguished as between those which act as arguments and those which act as parameters. The possibility of such a distinction is required in order that the comprehension axioms concerning propositional functions be able to guarantee the existence of those propositional functions whose existence is expected to be so guaranteed. The resources for making the distinction between arguments and parameters will be discussed in detail in Chapter 5 when we look at LT(PM).

The second complication concerns propositions. Recall that since LT(POM)'s propositions are terms in the realm of being on a par with all other terms and since LT(POM)'s variables range over every term that there is, the variables range over propositions. As a result, one should expect there to be propositions of, say, the following form in LT(POM): $(p)(q)[\{Proposition(p) \,\&\, Proposition(q)\} \rightarrow R(p,q)]$, where R is some propositional function from pairs of propositions to propositions. Indeed, the above rules of formation guarantee this circumstance. However, in addition to propositions of this form, there are propositions having the following peculiar forms in LT(POM):

$(p)[Proposition(p) \rightarrow p]$,

$(p)[Proposition(p) \rightarrow (p \vee \neg p)]$,

$(p)(q)[\{Proposition(p) \wedge Proposition(q)\} \rightarrow (p \rightarrow (q \rightarrow p))]$.

Russell takes the first one to express the statement that whatever p may be, if it is a proposition, then it is the case. He takes the second one to express the statement that whatever p may be, if it is a proposition, then either it or its negation is the case. He takes the third one to express the statement that whatever p and q may be, if they are propositions, then if p is the case, then q is the case only if p is. Propositions exhibiting forms of this sort are obviously difficult to interpret coherently. I shall examine this matter below.

(b) Axioms and Rules of Inference

Russell explicitly puts forward in *Principles* many of the axioms and rules of inference belonging to LT(POM). These axioms and rules resemble those belonging to a standard higher-order logical theory. Of the axioms and rules that may be taken to occur only implicitly in *Principles*, only a few have features that are nonstandard. I briefly summarise the relevant considerations below.

Propositional Axioms:

In §§18-9 of *Principles*, Russell explicitly puts forward a collection of truth-functional tautologies which include the following familiar ones: $(p \rightarrow q) \rightarrow (p \rightarrow q)$; simplification: $pq \rightarrow p$; syllogism: $[(p \rightarrow q).(q \rightarrow r)] \rightarrow (p \rightarrow r)$; importation: $(p \rightarrow (q \rightarrow r)) \rightarrow (pq \rightarrow r)$; exportation: $(pq \rightarrow r) \rightarrow [p \rightarrow (q \rightarrow r)]$; composition: $[(p \rightarrow q).(p \rightarrow r)] \rightarrow (p \rightarrow qr)$; reduction: $((p \rightarrow q) \rightarrow p) \rightarrow p$; $(p \vee q) \leftrightarrow [(p \rightarrow q) \rightarrow q]$. After putting these forward, he makes the following interesting remark about the completeness of the propositional fragment of LT(POM):

> From this point we can prove [with the rules of *modus ponens* and substitution] the laws of contradiction and excluded middle and double negation, and establish all the formal properties of logical multiplication and addition -- the associative, commutative and distributive laws. Thus the logic of propositions is now complete.[*Principles*, §19]

Quantification Axioms:

Although Russell does not explicitly put forward any quantification axioms in *Principles*, the axioms of universal instantiation satisfying the schema '$(x)\phi(x) \rightarrow \phi(y)$' might be taken to be implicit in some of what he says.[95] It is important to note that, however such axioms might be formulated in LT(POM), they must be restricted in order to respect the distinction between

---

[95]See *Principles*, §§25, 40-45.

Thing and concept. That is, consider a proposition having the form '(F) --- Ft ---', where the variable F stands in predicative position in the complex Ft. Since a Thing may never stand in predicative position in any complex, no Thing may instance F even though F's range is completely unrestricted. Thus, appropriate restrictions are required in order to reflect this circumstance within the formal system.

Identity Axioms:

In §§24 and 26 of *Principles*, Russell explicitly defines identity in the way in which it is defined in standard higher-order logic. As a result, LT(POM) requires no specific axioms to govern this notion.

Comprehension Axioms:

In §§23 and 24 of *Principles*, Russell rather explicitly puts forward unrestricted comprehension axioms.[96] These axioms concern only the existence of classes but they are very similar to the ones that LT(POM) requires in order to insure the existence of the propositional functions that Russell takes to be in its ontology. To this extent, comprehension axioms concerning the existence of propositional functions might be taken to be an implicit part of LT(POM).

Rules of Inference:

In §18 (4) of *Principles*, Russell writes:

> A true hypothesis in an implication may be dropped, and the consequent asserted.

This is his statement of the rule of *modus ponens*. He does not appear to say anything explicit

---

[96]It is noteworthy that one of these axioms has the form of Frege's Basic Law V.

about the rule of universal generalisation. It may be taken to be an implicit part of LT(POM).

3. Difficulties Concerning LT(POM)

As one might have expected, LT(POM) involves several difficulties in addition to its being inconsistent. I examine these as follows.

(a) The first difficulty concerns certain quantified propositions. Recall that Russell takes it that there are propositions of the following form in LT(POM):

$(p)[\text{Proposition}(p) \rightarrow p]$,

$(p)[\text{Proposition}(p) \rightarrow (p \vee \neg p)]$,

$(p)(q)[\{\text{Proposition}(p) \wedge \text{Proposition}(q)\} \rightarrow (p \rightarrow (q \rightarrow p))]$.

One may ask how Russell justifies countenancing propositions having forms of this sort. By our lights, insofar as the variables p and q are construed as ranging over terms in the realm of being -- that is, objectually -- the contexts in which they figure must be predicative if the sense of such contexts is to be explainable.[97] Since contexts such as 'p', '$p \vee \neg p$' and '$p \rightarrow (q \rightarrow p)$' are not predicative, the sense of propositions having forms of the above sort is in this respect brought into question.

That Russell countenances such propositions suggests that, in so doing, he is conflating the two very different roles that LT(POM)'s account of proposition claims that propositions may play. Recall that, according to the account, a proposition is a term in the realm of being on a par with all other terms and it is an object of judgement. In virtue of playing the first role, a

[97]One may argue that quantified propositions are what are expressed by ordinary language statements of the kind that is exemplified by such locutions as 'Whatever thing you choose, it is self-identical' and 'Whatever person you choose, he is mortal'. As such, these propositions' quantified variables must figure within predications occurring within them.

proposition may be quantified over, referred to, predicated of, or taken as argument to a truth-functional connective. In virtue of playing the second role, a proposition may be expressed, asserted, supposed, believed, known, or merely understood. In this respect, perhaps Russell mistakenly countenances propositions of the kind in question as a result of taking propositions as in general able to play these two roles 'simultaneously'. Thus, in virtue of playing the first role, a given proposition may be quantified over and, in virtue of playing the second role, it may figure within a truth-functional compound or stand simply by itself.

It is noteworthy that there seems to be a certain analogy between, on the one hand, the mention and expression of propositions and, on the other hand, the mention and use of sentences. To this extent, the following statements about the mention and use of sentences might be relevant. One may *use* a sentence to make a claim. One may *mention* a sentence and predicate something of it. One may universally quantify over a collection of sentences and predicate something of each. However, normally, one may neither use a sentence while it is being mentioned and predicated of nor may one use it while it is being quantified over and likewise predicated of.[98]

(b) The second difficulty concerns LT(POM)'s connectives *qua* formal objects. As one may have already gathered, Russell treats these ambiguously. Consider, for instance, the conjunction sign $\wedge$. On the one hand, Russell treats it as a sentential connective such that if S_1 and S_2 are sentences, then $\ulcorner S_1 \wedge S_2 \urcorner$ is the sentence that is their conjunction -- as such, $\ulcorner S_1 \wedge S_2 \urcorner$ is true if and only if both S_1 and S_2 are true.[99] On the other hand, Russell treats $\wedge$ as a functor that

---

[98]Liar and truth-teller sentences may count as exceptions to this claim but appeal to these does not help to explain the sense of the propositions mentioned above.

[99]Note that here sentences are to be distinguished from names.

means the 'conjunction function' C such that if P_1 and P_2 are propositions, $C(P_1,P_2)$ is the propositional compound that is their conjunction. As a functor, $\wedge$ grammatically relates to names, not sentences. The two senses of $\wedge$ in question are of course related. If S_1 expresses P_1 (under appropriate circumstances) and S_2 expresses P_2 (under appropriate circumstances), then $\ulcorner S_1 \wedge S_2 \urcorner$ will express $C(P_1,P_2)$.

The difficulty at issue here with respect to the conjunction sign is that, for some LT(POM) formulae containing one or more occurrences of this sign, it is not entirely clear which if any of such occurrences to treat as a sentential connective as opposed to a functor. Consider, for instance, '$(p)[\text{Proposition}(p) \rightarrow \{(p \wedge p) \rightarrow p\}]$'. Since $\wedge$ is flanked by variables, it may be read as a functor that 'operates' on them yielding a singular term. However, since '$p \wedge p$' is an antecedent of a conditional, $\wedge$ may equally well be read as a sentential connective.[100]

Similar remarks may be made about LT(POM)'s other connectives. The material implication sign $\rightarrow$, however, distinguishes itself in that Russell treats it in a further third way, *viz.*, as a dyadic predicate expressing the relation of logical consequence. To this extent, the difficulty specifically with respect to $\rightarrow$ might be thought to be even greater than it is with respect to any of the other LT(POM) connectives. However, some have arrived at ways of reading occurrences of this sign so as to make sense of formulae containing them. Consider, for instance,

$$(p)(q)[\{\text{Proposition}(p) \wedge \text{Proposition}(q)\} \rightarrow (p \rightarrow (q \rightarrow p))].$$

Richard Cartwright has suggested that quantified formulae of this kind -- that is, formulae whose principal logical constant after outside quantifiers are stripped away is $\rightarrow$ -- may be made sense of by reading their occurrence of $\rightarrow$ that is the principal logical constant after outside

[100] Of course, neither reading makes complete sense of the formula in question in the light of the first difficulty.

quantifiers are stripped away as the dyadic relation mentioned above and the other occurrences of ¬ (as well as the other LT(POM) connectives) as functors. Note that this consideration appears to support the suggestion raised in the discussion of the first difficulty that Russell takes propositions as in general able to play their two characteristic roles 'simultaneously'.

It is interesting that Quine accuses Russell of making a use/mention error in treating the material implication sign as able to express the relation of logical consequence. Briefly, according to Quine,[101] *implication* -- i.e., logical consequence -- is a relation that holds between sentences (and their schemata) such that, if 'S_1' and 'S_2' are names mentioning sentences, S_1 implies S_2 if and only if, roughly speaking, any interpretation that makes S_1 true also makes S_2 true.[102] Furthermore, according to Quine, a *conditional* is a compound sentence having the obvious form such that it is true if and only if its antecedent is false or its consequent is true. Moreover, whenever one utters a conditional, one does not mention its antecedent or consequent; rather, one uses them. In this respect, implication is to be clearly distinguished from the conditional. (They are of course intimately related to each other. In fact, one test for implication between sentences is that the corresponding conditional be truth-functionally valid.) It is not surprising, in this light, that Quine diagnoses Russell's treating the material implication sign as able to express the relation of logical consequence to be the simple confusion of use with mention.

Since Russell takes logical consequence as a relation that is borne by propositions rather than sentences or their schemata, Quine's diagnosis is somewhat inaccurate. It would be

---

[101]See *Mathematical Logic*, §5, pp. 27-33, and *Methods of Logic*, pp. 49-51.

[102]Quine's explanation of implication is actually more complicated than this as it takes implication to be principally a relation between schemata rather than sentences. It allows implication to be construed as a relation between sentences only derivatively.

better to talk about the confusion of *expression* with mention rather than about the confusion of *use* with mention. However, to the extent that there is a significant analogy -- as I suggested above -- between these two pairs --- use/mention and expression/mention -- Quine's diagnosis is insightful. In any case, the difficulty remains of how to read various occurrences of LT(POM)'s connectives contained in LT(POM) formulae.

(c) The third difficulty is somewhat technical but is closely related to the first and second difficulties. Consider the propositions focused upon in the discussion of the first difficulty:

$(p)[\text{Proposition}(p) \rightarrow p]$,

$(p)[\text{Proposition}(p) \rightarrow (p \vee \neg p)]$,

$(p)(q)[\{\text{Proposition}(p) \wedge \text{Proposition}(q)\} \rightarrow (p \rightarrow (q \rightarrow p))]$.

Recall that propositions having forms of this sort are not included among those propositions that are covered by the inductive rules of formation. It is thus natural to ask how such rules could be modified so as to cover them.

The simplest modification is that which results in the base clauses' countenancing the variables as atoms. However, since the variables range over all terms in the realm of being, propositional and nonpropositional alike, these would be very peculiar atoms. In addition, unless the recursive clauses were significantly modified in turn, they would countenance constructions of the following bizarre kind as propositional functions: $p \rightarrow \phi(c)$, $p \rightarrow q$, and $\neg q.(x)\phi(x)$. Such constructions would become propositions, by Russell's lights, only after something having a form like '$(p)[\text{Proposition}(p) \rightarrow \ldots]$' was placed over them.

Perhaps we could avoid these unsavoury consequences by resorting to a different modification of the rules of formation. However, it should be clear that any modification that managed to avoid such consequences would be awkwardly complex and rather *ad hoc*. To this extent, a satisfactory answer to the original question does not appear forthcoming. I should point out that this technical difficulty to some measure does not arise in LT(PM). Since this later theory is typed, it has variables whose ranges are restricted to propositions. These restricted variables are countenanced as atoms by the theory's rules of formation.

(d) The fourth difficulty is similar to the preceding one. Consider the proposition $(\phi).\phi(c)$. On the one hand, since LT(POM)'s variables range over all the terms that there are, ϕ ranges over Things as well as concepts. On the other hand, according to the distinction between Things and concepts, a Thing cannot stand in predicative position within an atomic proposition. Thus, where t is a Thing, the complex t(c) is ill-formed and, as such, it does not make sense to say of c that it is t (recall that for this reason, LT(POM)'s axioms of universal instantiation had to be suitably restricted). Clearly, the two circumstances in question here appear to be at odds with one another. On an objectual reading of quantification, the proposition $(\phi).\phi(c)$ is the proposition that whatever ϕ may be, ϕ applies to c. If ϕ may be t, then, on the face of it, this proposition 'implies' that t applies to c. But then it implies something that cannot be said.

Russell himself explicitly acknowledges this difficulty in *Principles*:

> If *xRy* implies *x'Ry'*, whatever *R* may be, so long as *R* is a relation, then *x* and *x'*, *y* and *y'* are respectively identical. But this principle introduces a logical difficulty from which we have been hitherto exempt, namely a variable with a restricted field; for unless *R* is a relation, *xRy* is not a proposition at all, true or false, and thus *R*, it would seem, cannot take *all* values, but only such as are relations.[*Principles*, §28]

Because LT(PM) is typed and, so, the ranges of its variables are restricted, this difficulty does not arise in this later theory.

(e) The fifth difficulty concerns a feature of propositional functions that I have neglected to mention so far. According to Russell, like propositions, propositional functions are dual-natured: they are terms in the realm of being on a par with all other terms and they are objects of judgement. Indeed, Russell says that propositional functions may be *asserted*. He writes:

> Let $\phi(x)$ be a propositional function whose argument is x; then we may assert $\phi(x)$ without assigning a value to x. This is done, for example, when the law of identity is asserted in the form "*A* is *A*."[103]

That propositional functions are objects of judgement and may, thus, be asserted raises two questions: why does Russell make this claim and how is one to understand the assertion of propositional functions?

Russell offers (at least) one reason for making this claim, *viz.*, that, to the extent that mathematical proofs involve tracts of reasoning that make essential use of open formulae, mathematical proofs make essential use of asserted propositional functions. Thus, he writes in ML:

> The distinction between asserting ϕx and asserting $(x).\phi x$ was, I believe, first emphasized by Frege (*1893*, p. 31). His reason for introducing the distinction explicitly was the same which caused it to be present in the practice of mathematicians, namely, that deduction can only be effected with *real* variables, not with apparent variables. In the case of Euclid's proofs, this is evident: we need (say) some one triangle *ABC* to reason about, though it does not matter what triangle it is. The triangle *ABC* is a *real* variable; and although it is *any* triangle, it remains the *same* triangle throughout the argument. But in the general enunciation the triangle is an apparent variable. If we adhere to the apparent variable, we cannot perform any deductions, and this is why in all proofs real variables have to be used. Suppose, to take the simplest case, that we know "ϕx is always true", that is, "$(x).\phi x$", and we know "ϕx always implies ψx", that is, "$(x).\{\phi x$ implies $\psi x\}$". How shall we infer "ψx is always true", that is, $(x).\psi x$"? We know it is always true that, if ϕx is true and if ϕx implies ψx, then ψx is true. But we have no premisses to the effect that ϕx is true and ϕx implies ψx; what we have is: ϕx is *always* true, and ϕx *always* implies ψx. In order to make our inference, we must go from "ϕx is always true" to ϕx, and from "ϕx always implies ψx" to "ϕx implies ψx", where the x, while remaining any possible argument, is to be the same in both. Then, from "ϕx" and "ϕx implies ψx", we infer "ψx"; thus

[103] *Principia*, p. 92.

> ψx is true for any possible argument and therefore is always true. Thus, in order to infer "*(x). ψx*" from "*(x) ϕx*" and "*(x).{ϕx implies ψx}*", we have to pass from the apparent to the real variable and then back again to the apparent variable. This process is required in all mathematical reasoning which proceeds from the assertion of all values of one or more propositional functions to the assertion of all values of some other propositional function, as, for example, from "all isosceles triangles have equal angles at the base" to "all triangles having equal angles at the base are isosceles". In particular, this process is required in proving *Barbara* and the other moods of the syllogism. In a word, *all deduction operates with real variables* ... [104]

It is important to note that even if it *were* correct that mathematical proofs could only be formalised by making use of open formulae,[105] one would not be compelled to make appeal to the yet-to-be-explained notion of asserted propositional function. By speaking from a metatheoretic point of view, one could easily explain the validity of such free variable reasoning -- characteristic of, for instance, EI subproofs -- by appealing to the fact that the logical rules of inference are such that, if a given model M and assignment S to the free variables satisfy the formulae figuring on the first n lines of a proof, then M and S will satisfy the first n+1 lines of the proof. Needless to say, Russell does not readily ascend to such a metatheoretic point of view.

Next, consider the second question how is one to understand the assertion of propositional functions. According to Russell, if $\lambda x(F(x)\rightarrow G(x))$ is a propositional function, then one is to understand its assertion as one understands the assertion of whatever is meant by "Any F is G".[106] More generally, if $\lambda x.H(x)$ is a propositional function, then one is to understand its assertion as one understands the assertion of whatever is meant by "Anything is

---

[104]ML, pp. 157-8.

[105]The formal system that Quine presents in *Mathematical Logic* does not need open formulae.

[106]In the following discussion, I abuse the distinction between a schema and a sentence whose schematisation the schema is. However, no confusion should result.

H". In response to this answer, one may ask in turn how is one to understand *any*. Russell has an interesting answer to this question:

> *Any* [*F*] denotes only one F, but it is wholly irrelevant which it denotes, and what is said will be equally true whichever it may be. Moreover, *any* [F] denotes a variable [F], that is, whatever particular [F] we may fasten upon, it is certain that *any* [F] does not denote that one; and yet of that one any proposition is true which is true of any [F].[*Principles*, §60]

Thus, where 'F' is to be instanced by a count noun, any F is some 'undetermined' or 'arbitrary' F, but no 'particular' one. Russell speaks here as if he were either expressing a contradiction of the form '$\exists x[F(x) \,\&\, \forall y(F(y) \rightarrow \neg x{=}y)]$', or committing himself to the existence of some mysterious object 'any F' which is not an F but may be taken to be one for argument's sake.

Since, for any propositional function $\lambda x(F(x) \rightarrow G(x))$, one is to understand its assertion as one understands the assertion of whatever is meant by "Any F is G", Russell explains how one is to understand a real variable x in terms of how one understands *any*. Considering numerical variables, he writes:

> Originally, no doubt, the variable was conceived dynamically, as something which changed with the lapse of time, or, as is said, as something which successively assumed all values of a certain class. This view cannot be too soon dismissed. If a theorem is proved concerning *n*, it must not be supposed that *n* is a kind of arithmetical Proteus, which is 1 on Sundays and 2 on Mondays, and so on. Nor must it be supposed that *n* simultaneously assumes all its values. If *n* stands for any integer, we cannot say that *n* is 1, nor yet that it is 2, nor yet that it is any other particular number. In fact, *n* just denotes any number, and this is something quite distinct from each and all of the numbers. It is not true that 1 is any number, though it is true that whatever holds of any number holds of 1. The variable, in short, requires the indefinable notion of *any* ...[*Principles*, §87]

In this respect, if $\lambda x.H(x)$ is a propositional function, then to assert that H(x) is to ascribe $\lambda x.H(x)$ to *any* x or to an *arbitrary* x.

It is noteworthy that Russell later offers a slightly different gloss on this matter. Rather than saying that to assert that H(x) is to ascribe $\lambda x.H(x)$ to any x, he says that to assert that H(x) is to assert some 'undetermined' value of the propositional function $\lambda x.H(x)$, but not any 'definite' value -- that is, it is to assert *any* value of $\lambda x.H(x)$. Thus Russell writes in *Principia*:[107]

> Any Value "ϕx" of the function ϕ^x can be asserted. Such an assertion of an ambiguous member of the values of ϕ^x is symbolised by "$\vdash . \phi x$."[108]
>
> When we assert something containing a real variable, as in *e.g.* "$\vdash . x=x$," we are asserting *any* value of a propositional function.[109]
>
> When we assert something containing a real variable, we cannot strictly be said to be asserting a *proposition*, for we only obtain a definite proposition by assigning a value to a variable, and then our assertion only applies to one definite case, so that it has not at all the same force as before. When what we assert contains a real variable, we are asserting a wholly undetermined one of all the propositions that result from giving various values to the variable. It will be convenient to speak of such assertions as *asserting a propositional function*. The ordinary formulae of mathematics contain such assertions; for example "$\sin^2 x + \cos^2 x = 1$" does not assert this or that particular case of the formula, nor does it assert that the formula holds for *all* possible values of x, though it is equivalent to this latter assertion; it simply asserts that the formula holds, leaving x wholly undetermined;[110]

In this respect, the difference between this later gloss and the earlier one is a matter of which notion *any* is taken to apply to.[111]

---

[107]"^x" represents Russell's circumflexed *x*.

[108]*Principia*, p. 17.

[109]*ibid.* p. 18.

[110]*ibid.* p. 18.

[111]In ML Russell appears to offer a combination of the two glosses. His explanation of the distinction between *any* and *all* begins as follows:

> Given a statement containing a variable x, say "$x=x$", we affirm that this holds in all instances, or we may affirm any one of the instances without deciding as to which instance we are affirming. The distinction is roughly the same as that between the general and the particular enunciation in Euclid. The general enunciation tells us something about (say) all triangles, while the particular enunciation takes one triangle and asserts the same thing of this one triangle. But the triangle taken is *any* triangle, not some one special

In any case, both the earlier and later glosses on how one is to understand the assertion of a propositional function are difficult to comprehend. To the extent that Russell takes such assertions to figure in LT(PM), this difficulty besets this later logical theory as much as it does LT(POM).

The difficulties discussed here of course pale beside the problem that LT(POM) gives rise to the paradoxes. I turn to consider Russell's solution to this problem in the next chapter. I should note now that in addition to resolving the problem, the solution to a certain extent, removes the third and fourth difficulties discussed above. The other difficulties, however, are inherited unaffected by LT(PM).

---

triangle; and thus, although, throughout the proof, only one triangle is dealt with, yet the proof retains its generality. If we say: "Let *ABC* be a triangle, then the sides *AB* and *AC* are together greater than the side *BC*", we are saying something about *one* triangle, not about *all* triangles; but the one triangle concerned is absolutely ambiguous, and our statement consequently is also absolutely ambiguous. We do not affirm any one definite proposition, but an undetermined one of all the propositions resulting from supposing *ABC* to be this or that triangle. This notion of ambiguous assertion is very important, and it is vital not to confound an ambiguous assertion with the definite assertion that the same thing holds in *all* cases.

The distinction between (1) asserting any value of a propositional function and (2) asserting that the function is always true is present throughout mathematics, as it is in Euclid's distinction of general and particular enunciations. In any chain of mathematical reasoning, the objects whose properties are being investigated are the arguments to *any* value of some propositional function.[ML, p. 156]

Chapter 3

The Vicious-Circle Principle

I claimed at the beginning of the last chapter that the logical theory of *Principia*, LT(PM), may be understood as the result of modifying the logical theory implicit in *Principles*, LT(POM), in order to abide by Russell's vicious-circle principle. Therefore, in order to understand LT(PM) one should first understand LT(POM) as well as the vicious-circle principle. We briefly looked at LT(POM) in the last chapter. I now turn to examine the vicious-circle principle. This principle was Russell's considered solution to the paradoxes. As is well known, he entertained several other possible solutions before he arrived at this one. To provide a little context, I very briefly survey the more salient of these early solutions before examining the vicious-circle principle in detail.

1. The Early Solutions

The survey covers four of the early solutions. The first solution, rather surprisingly, appears in an appendix to *Principles* under the rubric the *Doctrine of Types*; it is a version of simple type theory and, accordingly, it is a logical theory itself, different in spirit as well as in detail from LT(POM). *Qua* simple type theory, it avoids giving rise to the set-theoretic paradoxes by failing to countenance structures like '$x \in x$' as well-formed. Although it resembles standard versions of simple type theory such as that presented by Tarski in his "Wahrheitsbegriff" -- at least to the extent that it contains a hierarchy of types consisting of a bottom type of individuals (LT(POM)'s Things) and then a type of classes of individuals, and then a type of classes of these, and so on -- the Doctrine of Types differs from these versions in several respects. Most notably, unlike these others, it distinguishes between the range of significance of, say, a monadic propositional function and the type of a term -- in Russell's sense -- that may fall under it. By its lights, such a range is a superset of such a type. The Doctrine of Types differs further in that, by its lights, all the ranges of significance form a type, the numbers form a type, and the propositions form a type, where each of these types lies outside the hierarchy mentioned above. Because the propositions form a type, the Doctrine of Types is inconsistent,[112] and for this reason alone, Russell rejected it as a viable solution.[113]

Russell offers the three other solutions in his 1905 "On Some Difficulties in the Theory of Transfinite Numbers and Order Types". These solutions all concern LT(POM)'s comprehension axioms. Recall that such axioms are appealed to in order to derive Cantor's paradox, Burali-Forti's paradox, Russell's class and propositional function paradox, and his proposition paradox. For this reason in 1905 Russell suspects that they are responsible for the paradoxes. More precisely, since he in some sense takes such axioms as a required part of any correct formalisation of logic, he suspects that they must be mischaracterised in some way and that such a mischaracterisation is responsible for the paradoxes. If this is the case, then in what does such a mischaracterisation consist? Consider how LT(POM) characterises the comprehension axioms. Since it characterises them as unrestricted, they all have the form

$$\exists f \forall v_1 \ldots \forall v_n [f(v_1, \ldots, v_n) \leftrightarrow \text{----}]$$

where '----' represents any formula not containing the variable f free but possibly containing other variables free among which of course may be any or all of $v_1, \ldots, v_n$.[114] Call such a

---

[112]See *Principles*, §500, or the discussion of Russell's proposition paradox in Chapter 1, §1.

[113]Others have claimed that Russell rejected the Doctrine of Types also because it does not treat of the semantic paradoxes and because its ranges of significance appear rather *ad hoc*. See C. Chihara and I. Copi.

[114]Since LT(POM) contains something like second-order quantification, for any given n, it may be thought to have a single comprehension axiom having the form '$\forall G \exists f \forall v_1 \ldots \forall v_n [f(v_1, \ldots$

formula an *s-formula*. In this regard, in 1905 Russell specifically suspects that the mischaracterisation of the comprehension axioms consists in LT(POM)'s countenancing all possible formulae -- satisfying the constraint of not containing the variable f free -- as s-formulae when only a very restricted collection of such formulae may be so countenanced. If his suspicion is correct, then one must determine which of the possible formulae are to be countenanced as s-formulae in order to arrive at the proper characterisation of the comprehension axioms. Toward this end, Russell puts forward in "On Some Difficulties" three possible theories that make such a determination: these are the three solutions with which we are at present concerned.[115]

The first of the three solutions is the *zigzag theory*. Briefly, according to this theory, a formula is to be countenanced as an s-formula if and only if such a formula is *fairly simple*. This particular determination is motivated by the circumstance that appeal to comprehension axioms whose s-formulae are rather complicated or recondite such as '$\neg x \in x$' engenders paradox whereas appeal to those axioms whose s-formulae are in some sense simple never does. Of the three theories put forward in "On Some Difficulties", the zigzag theory is the least worked out. Russell failed to arrive at anything like an explanation of the conditions for a formula's being 'fairly simple'. Indeed, no one since has succeeded in arriving at such an

---

$,v_n) \leftrightarrow G]$' as opposed to an infinite number of such axioms, each of which satisfies the appropriate schema. However, this point is rather insignificant since LT(POM) in any case requires a different comprehension axiom for each n. There is no escaping an infinite axiomatisation. As we shall see, the case is much worse for LT(PM).

[115]I should note that rather than carrying out his discussion of these matters in "On Some Difficulties" at the metalinguistic level, Russell as always does so at the object level. Thus, rather than asking which formulae are to be countenanced as s-formulae, he asks which propositional functions determine classes or whatever else are taken to be comprehended by the axioms in question. He calls the propositional functions that do determine these items *predicative*. This is a word that acquires at least two other usages later on.

explanation. Some suggest that Quine is the one who has come the closest to doing so by means of his set theory *New Foundations*.[116]

The second of the three solutions is the *theory of the limitation of size*. According to this theory, a formula is to be countenanced as an s-formula if and only if the item such a formula specifies is not *too large*. This particular determination is motivated by the circumstance that the set-theoretic and mixed paradoxes all make reference to very large classes or 'very large' propositional functions such as the class of all classes, the class of all ordinals, the class of all non-self-membered classes, and the propositional function that applies to all and only non-self-satisfying propositional functions. Unfortunately, Russell did not succeed in arriving at a clear explanation of the conditions for a formula's not specifying an item that is too large. Interestingly, some have understood Zermelo-Fraenkel set theory as providing such an explanation.[117] By their lights, its axioms of Aussonderung, pairing, power-set, union, and replacement determine collectively which formulae are to be countenanced as s-formulae in such a way that, given that certain sets are taken to exist, these axioms affirm the existence of other sets whose size is, roughly speaking, not very much larger than that of these given sets.

The third solution is the *substitutional theory*. According to this theory, *no* formula is to be countenanced as an s-formula. Clearly, this theory is the most revisionary of the three. Unlike LT(POM), it does not countenance propositional functions or classes. In their stead, it countenances a primitive operation of substitution S which is such that given any terms c and d and a proposition p containing c as constituent, S(p,d,c) is the proposition that results from

---

[116]For instance, see Gödel, "Russell's Mathematical Logic", p. 125.

[117]Azriel Levy is one. However, Michael Hallett in *Cantorian Set Theory and Limitation of Size* argues against such an understanding.

substituting d for c in p. The motivation behind this theory is simply that all of the paradoxes make reference to propositional functions or classes. After expending much effort to work out the theory, Russell eventually abandoned it for two reasons: the theory is technically very cumbersome and it leads to contradictions of its own.

2. Vicious Circularity

In response to Russell's "On Some Difficulties", H. Poincaré put forward his own account of which formulae are to be countenanced as s-formulae in his 1906 "Les mathématiques et la logique". According to this account, a formula is to be countenanced as an s-formula if and only if it is not what he designates *viciously circular*. Poincaré put forward this account because he took there to be some sort of pernicious circularity on a par with definitional circularity implicitly involved in the arguments to the paradoxes. The circularity in question might be taken as the circularity pointed up by Russell in his discussion of *self-reproductive processes*.[118] In this light, it is not surprising that later, in his 1906 "On 'Insolubilia'", Russell acceded to Poincaré's account.

At this point, even if one has some understanding of the circularity that both Poincaré and Russell point up, one may be prompted to ask for an explanation of the conditions for a formula's being viciously circular. Russell to a certain extent offers an answer to this question when he states his celebrated vicious-circle principle. I employ the qualifier "to a certain extent" because such an explanation should, strictly speaking, speak at the metalinguistic-level whereas Russell as always states the principle at the object-level. The principle may, however, be construed metalinguistically.

---

[118]See the discussion about self-reproductive processes at the end of §2, Chapter 1. I discuss the nature of the circularity involved near the end of this chapter.

As one may have expected, Russell provides more than one rendering of the vicious-circle principle: at least eight appear in the corpus. Here are four of them.

> Whatever involves an apparent variable must not be among the values of that variable.[119]

> Whatever involves *all* of a collection must not be one of the collection.[120]

> If, provided a certain collection had a total, it would have members only definable in terms of that total, then the said collection has no total.[121]

> More generally, given any set of objects such that, if we suppose the set to have a total, it will contain members which presuppose this total, then such a set cannot have a total.[122]

Some have remarked that these renderings do not express a single principle. In "Russell's Mathematical Logic", Gödel claims that 'corresponding to the phrases "definable only in terms of", "involving", and "presupposing", we have really three different principles ...'.[123] Others have remarked that none of the renderings are particularly clear.

In order to arrive at a clear understanding of Russell's answer to the question about the conditions for a formula's being viciously circular, one must examine his vicious-circle principle in detail. I do so by examining two of the above renderings: the rendering that employs "presuppose" and the rendering that employs "definable". For convenience, I dub these *VCP1* and *VCP2*. The examination will show that, although there is a sense in which Gödel's claim is correct, it is not very significant.

---

[119] "On 'Insolubilia'", p. 198.

[120] ML, p. 155.

[121] *ibid.*

[122] *Principia*, p. 37.

[123] Kurt Gödel, *Collected Works*, Volume II, p. 127.

Two points should be noted before I start. First, as Quine has argued,[124] definition is best understood as what occurs when a new notation is introduced as short for an old one. Thus, to the extent that VCP2 really concerns comprehension axioms, its employment of "definable" is misleading. That is, for any given comprehension axiom c, c's s-formula is not a definiens and, so, c does not define anything. Rather, it affirms the existence of some item i which satisfies a certain formula -- i.e., the formula following c's outer existential quantifier. Since it is usually the case that there can be only one such item, c may also be understood to *specify* the item i. In this respect, I shall talk about specification rather than definition in the following.

Secondly, although it may not yet be clear in what a formula's being viciously circular consists, it should be rather clear what such circularity is not. In particular, since definition is not really at issue, the circularity in question is not the circularity of 'smuggling in the definiendum into the definiens'.[125] Likewise, since the arguments to the paradoxes do not, on the face of it, commit *petitio principii*, it is not the circularity of smuggling a conclusion among premisses. I should remark that because of these circumstances, Quine concluded that, to the extent that Russell and Poincaré thought that they had observed some circularity responsible for the paradoxes, they were confused. Quine of course did not attend to the details of LT(POM) and, as a result, his assessment is not altogether fair -- as we shall see.

(a) VCP1:

Recall Russell's expression of VCP1:

---

[124] *Set Theory and its Logic*, p. 242.

[125] *ibid.*

> More generally, given any set of objects such that, if we suppose the set to have a total, it will contain members which presuppose this total, then such a set cannot have a total.

In order to become clear about what Russell intends VCP1 to say, one should consider some of the expressions that it employs. To begin with, consider 'set'. Russell clearly intends this expression to have a meaning that is more general in its application than that which it currently has (or may be taken to have), *viz.*, the concept of set as explained by Zermelo-Fraenkel set theory. In addition to intending 'set' to mean such a concept, Russell intends it to mean the concepts collection, extension of a predicate, and proper class. To this extent, perhaps the expression 'totality' might suit his intentions more aptly.[126] Next, consider 'member'. Russell intends this expression to have a meaning that is correlatively more general in its application than that which it may currently be understood to have, *viz.*, the concept of member of ZF set. In addition to intending 'member' to mean such a concept, Russell intends it to mean the concepts part, component, constituent, subset, subclass, and so on. Next, consider the expression 'to have a total'. Clearly, by Russell's lights, to say of an item like a collection that it has a total is to say of it that it is a term in the realm of being. In other words, it is to say that the item may play the logical role of subject and, as such, the quantifier and substitution rules apply to it in the usual way. Finally, consider 'presuppose'. Russell intends this expression to mean a very important metaphysical relation -- that of *presupposition*. Although he does not write much that is explicit in order to explain the nature of this relation, his usage of the expression makes it evident that presupposition is a relation of ontological dependence in the sense, for instance, that the definiteness of wholes may be taken ontologically to depend on the definiteness of their parts and the definiteness of sets may be taken ontologically to depend on

---

[126]Here and just below I use the expressions 'totality' and 'item' as Russell uses 'class as many' -- that is, I treat them as a Boolos plural.

the definiteness of their elements. The relation, as such, is irreflexive and at least as strong as supervenience -- for if c ontologically depends on d, c would not exist if d did not.

Having these glosses, one may rephrase VCP1 as follows: Given any totality T such that, if we suppose T to be a term, it will contain terms that ontologically depend on it, then T cannot be a term. Better: No totality that is a term may contain a term which ontologically depends on it. Given the notion of well-foundedness from set theory,[127] one may rephrase VCP1 still more succinctly: All totalities that are terms are well-founded. Thus understood, VCP1 implies that no wholes may be proper parts of themselves and that no sets may be members of themselves or members of ... of members of themselves.

Admittedly, VCP1 so understood enjoys a certain intuitive appeal. Many in the history of philosophy have taken it as correct. Leibniz implicitly appealed to it in one of his arguments for the existence of so-called corporeal substances.[128] Kant implicitly appealed to it in his argument to the thesis of the second antinomy.[129] Russell and Poincaré apart, more recently Gödel has claimed that VCP1 is plausible.[130] Notwithstanding the intuitive appeal that it enjoys, consistent set theories have been developed -- more precisely, consistent relative to ZF's consistency -- that deny VCP1 by asserting the existence of non-well-founded sets. Moreover, since the axiom of regularity is independent of ZFC, ZFC itself has nonstandard models containing non-well-founded sets. To this extent, one might perhaps deny VCP1 without

[127]Roughly speaking, x is well-founded if and only if each of the branches of its membership tree is finite in length. Thus, each begins with either the null-set or an urelement.

[128]See Leibniz's letter of April 1687 to Arnault, C.I. Gerhardt, ed., *G.W. Leibniz: Die Philosophischen Schriften*, Vol. II, p. 96.

[129]*The Critique of Pure Reason*, A434/B462 to A444/B472.

[130]"Russell's Mathematical Logic", p. 127.

contradicting oneself. Needless to say, these points would not sway the conviction of anyone who was impressed by its intuitive appeal.

Importantly, Russell takes VCP1 thus understood as incompatible with many of the type-free features of LT(POM). Indeed, he argues from VCP1 and certain LT(POM) claims about propositions to the conclusion that no propositional function may apply to itself *with sense*. The argument goes roughly as follows:[131] Let $\lambda x.\phi x$ be a monadic propositional function. By Russell's lights, "a function is not a well-defined function unless all its values are *already* well-defined."[132] Suppose that the proposition ϕc is a value of $\lambda x.\phi x$. Since the well-definedness of $\lambda x.\phi x$ depends on the *prior* well-definedness of ϕc, $\lambda x.\phi x$ presupposes ϕc. Since c is a constituent of ϕc, ϕc presupposes c. By the transitivity of presupposition, $\lambda x.\phi x$ presupposes c. By its irreflexivity, $\lambda x.\phi x$ and c are distinct. At this point, Russell concludes that $\lambda x.\phi x$ may not apply to itself with sense. In other words, he concludes not that the complex $\phi(\lambda x.\phi x)$ is false but rather that it is ill-formed and, thus, that the judgement that $\lambda x.\phi x$ is ϕ cannot be expressed. One may ask of course how he justifies this particular conclusion when it appears, on the face of it, that he should rather conclude that the complex $\phi(\lambda x.\phi x)$ is false.[133] This question is difficult since its answer would probably appeal to special features of presupposition as Russell conceives this relation and to details of the background logic within which his argument is supposedly carried out.[134] It is noteworthy that Zermelo-Fraenkel set theory with the axiom of regularity respects VCP1 in the sense that all the sets that

---

[131]See *Principia*, p. 39.

[132]*Principia*, p. 39, my italics.

[133]Hylton in his RIAP seems to think that the conclusion follows from the argument. See RIAP, pp. 300-1.

[134]In so arguing, perhaps he is already using a type-theoretic logic.

it talks about are well-founded.[135] This theory, however, does not realise in any way this last conclusion that Russell takes to follow from VCP1.

Not surprisingly, several interesting claims also incompatible with the type-free features of LT(POM) follow in turn from Russell's conclusion. I mention three. The first is that, contrary to the position of LT(POM) that for any propositional function $\lambda x.\phi x$ and for any term c in the realm of being, $\lambda x.\phi x$ may be predicated with sense of c and such a predication produces either the proposition or the propositional function ϕc, $\lambda x.\phi x$ may only be predicated with sense of those terms belonging to some proper subclass of all the terms in the realm of being. As is well-known, Russell calls such a proper class a *type*. The second claim is that, contrary to the position of LT(POM) that for any propositional function $\lambda x.\phi x$, the variable x that figures in $\lambda x.\phi x$ is completely unrestricted, such an x must be restricted to range only over the proper subclass or type of terms of which $\lambda x.\phi x$ may be predicated with sense. This claim, strictly speaking, only follows from the conjunction of Russell's conclusion with the additional assumption that a variable figuring in a propositional function must be restricted to range over only those terms of which the function may be predicated with sense. Although the assumption is reasonable, one should recall that LT(POM) violates it with respect to the fourth difficulty discussed in Chapter 2. The third claim is that, contrary to the LT(POM) position that there is only one kind of variable, to the extent that there are propositional functions having different types of terms of which they may be predicated with sense, there will be variables of different kinds ranging over these different types.

Given that any propositional function $\lambda x.\phi x$ has associated with it some type -- that is, the type of terms of which it may be predicated with sense -- any term c in the realm of being

[135] I am only considering the theory's standard model here.

may be understood to fall under one or more types, *viz.*, the types associated with the propositional functions which may be predicated of c with sense. In order to avoid confusion, for a given propositional function $\lambda x.\phi x$, call the type of terms of which it may be predicated with sense its *argument-type*, and, for a given term c, call the type or types that c may fall under simply c's *type(s)*.[136]

At this point, one may be prompted to ask how the types are all configured. By Russell's lights, VCP1 at least partly determines the answer to this question. Not only does he take VCP1 to require that the argument-type t of a given propositional function f be restricted so as not to contain, for instance, f itself, but he takes VCP1 to require that t not contain any other propositional function g whose own argument-type contains f. More generally, Russell takes VCP1 to require that the argument-types of propositional functions be configured so that no sequence of the following kind occur: f_1 applies to f_2 with sense, f_2 applies to f_3 with sense, ... , f_{i-1} applies to f_i with sense, f_i applies to f_1 with sense. In other words, the types must be configured so as to realise a well-founded ordering. There are of course many kinds of well-founded ordering and, so, one may ask which particular ordering is so realised. Unfortunately, even by Russell's lights, VCP1 does not have anything to say to this question. Also, since the circumstance that the types realise a well-founded ordering does not determine whether or not the types are cumulative, one may ask whether or not they are. That is, one may ask whether or not the types are such that, for any two types t_1 and t_2 where t_1 is greater than (according to the ordering) t_2, some of the items falling under t_2 may also fall under t_1. Here, again, VCP1 even by Russell's lights does not have anything to say.

It is noteworthy that many have taken the following configuration of types as the most

[136]For the sake of simplicity, I have been focusing on monadic propositional functions. However, analogous claims apply to polyadic propositional functions.

plausible well-founded ordering (some have even taken it as the only possible one):[137] First, consider the collection of all terms in the realm of being that Russell calls *Things*. These are the terms that do not apply to anything. Take this collection as a type and call it *level-0*. Next, consider the collection of monadic propositional functions that apply to all and only level-0 items with sense. Take this collection as a type and call it *level-1*. Next, consider the collection of monadic propositional functions that apply to all and only level-1 items with sense. Take this collection as a type and call it *level-2*. Clearly, this procedure may be iterated indefinitely. Moreover, it may be extended without difficulty to cover polyadic propositional functions and propositions. The result of such an iteration and extension is a well-founded ordering in which there figures every type of every term in the realm of being. This ordering resembles the simple hierarchy of types presented by Tarski in his "Wahrheitsbegriff" and by Gödel in his 1931 paper at least to the extent that its types are noncumulative and, since there is a least type -- i.e., level-0 -- as opposed to several minimal types, the types of the monadic propositional functions are *well-ordered*.

There are two points to make about the configuration of types described above. On the one hand, the description is only a sketch. In order to provide a more precise characterisation of the configuration, one would have to specify how the configuration relates to the logical system that concerns it. For instance, suppose that for some fixed y, '$\lambda x.\phi xy$' means some monadic propositional function f in the configuration. Although the above description indicates roughly how the type of f relates to its argument-type, it says nothing about how the type of the parameter y relates to these two types. Needless to say, in order to provide a detailed characterisation of the configuration one would have to specify other important features of it as

---

[137]Indeed, Peter Hylton in his RIAP incorrectly takes VCP1 to necessitate it. See RIAP, pp. 301-2.

well. On the other hand, since VCP1, even by Russell's lights, already underdetermines the particular kind of well-founded ordering that the types realise as well as whether or not the types are cumulative, it *a fortiori* underdetermines such relatively fine features of the configuration.

Now that we have some understanding of VCP1, we should return to the question posed at the beginning of this section about an explanation of the conditions for a formula's being viciously circular (and, hence, failing to be an s-formula). In the light of what we have seen, we may expect the explanation to go roughly as follows: a formula is viciously circular if and only if it attempts to express a predication that, according to VCP1 and the claims that Russell takes to follow from it, cannot be expressed. In this respect, formulae of the following kind whose variables are not restricted to types that form a well-founded ordering are viciously circularity: $x(x)$, $x(y) \wedge y(x)$, $x(y) \wedge y(z) \wedge z(x)$, $x \in x$.

It is noteworthy that, in this light, the comprehension axioms that are appealed to in the arguments to the set-theoretic and mixed paradoxes all involve s-formulae that are viciously circular. To the extent that no viciously circular formula is to be countenanced as a legitimate s-formula, the deduction of these paradoxes, at least by means of the arguments described in Chapter 1, §2, is obstructed. To this extent, one may say that violation of VCP1 gives rise to such paradoxes, where such violation may be taken to consist in employing viciously circular formulae as s-formulae.[138]

(b) VCP2:

[138]One should conclude further that insofar as a viciously circular formula attempts to express a predication that it cannot express, such a formula -- far from being considered as a legitimate s-formula -- cannot even be considered as well-formed.

Recall Russell's statement of VCP2:

> If, provided a certain collection had a total, it would have members only definable in terms of that total, then the said collection has no total.

In order to become clear about what Russell intends VCP2 to say, one should consider some of the expressions that it employs. To begin with, consider its expressions 'collection', 'member', and 'had a total'. Clearly, these may be understood in terms of the glosses at which we arrived when we looked at VCP1. Next, consider 'definable'. As we saw above, it is best to read this expression as meaning *specifiable*. Finally, consider 'in terms of'. Russell's application of VCP2 makes it clear that he intends this to mean something '. . . *by quantifying over*'. In this respect, one may rephrase VCP2 as follows: No totality T may contain a term that is specifiable only by quantifying over T.

VCP2 thus understood recommends rather clearly its own explanation for a formula's being viciously circular (and, hence, failing to be an s-formula), *viz.*, that a formula F is viciously circular when and only when, if F were to be countenanced as an s-formula, it would contain at least one quantified variable whose range contained the term whose existence would be affirmed by the comprehension axiom having F as its s-formula. To illustrate, consider the following familiar comprehension axiom from standard second-order logic that affirms the existence of the property of being a natural number ('0' denotes zero and 's' denotes the successor function):

$$\exists N \forall z[N(z) \leftrightarrow \forall G\{(G(0) \wedge \forall u(G(u) \rightarrow G(s(u)))) \rightarrow G(z)\}]$$

The axiom's s-formula contains the quantified variable G that ranges over the term $\lambda x.Nx$ whose unique existence the axiom affirms. As such, by VCP2's lights, the axiom's s-formula is viciously circular and, so, is not to be countenanced as a legitimate s-formula.[139,140]

Not surprisingly, VCP2 thus understood is inconsonant with many of the type-free features of LT(POM). It is inconsonant in at least three respects. First, to the extent that there are s-formulae containing quantified variables, VCP2 requires that there be some restricted variables since, by its lights, only restricted variables may occur bound in s-formulae. Again, call the range of any of such restricted variables a *type*. Secondly, if for every s-formula containing a bound variable, there is another s-formula containing a bound variable that ranges over the item the first s-formula specifies, then, since the second bound variable must have a type different from that of the first, there must be an infinite number of restricted variables whose types are different one from the other. Note that this claim does not imply that every variable must be restricted -- a situation different from that associated with VCP1. Thirdly, recall that, according to LT(POM), a propositional function is the result of taking a proposition and replacing one or more of its constituent terms with one or more variables. In this respect, insofar as there are an infinite number of restricted variables whose types are different one from the other, there will be an infinite number of propositional functions whose *argument-types* are different one from the other. Each of these functions will apply with sense to only some of all the terms in the realm of being.

Given that VCP2 requires types of its own, one may be prompted to ask how such types are all configured. VCP2 partly determines the answer to this question. For not only

---

[139]One should note that the range of the variable G in the above s-formula is just the totality T of which VCP2 forbids existence.

[140]One may recognise that contemporary logicians call an s-formula that is viciously circular in this sense (as well as its containing comprehension axiom) *impredicative*. They call an s-formula that is not so viciously circular (as well as its containing comprehension axiom) *predicative*. Note that this is a sense of 'predicative' that differs from Russell's usage. (Russell uses 'predicative' in yet another way in *Principia* as we shall see in Chapter 5.)

does VCP2 require that there be no specification of a term t_1 via some formula F containing quantified variables $v_1, \ldots, v_i$, some of whose types contain t_1, but it requires that there be no specification of a term t_1 via some formula F containing quantified variables $v_1, \ldots, v_i$, some of whose types contain a term t_2, which is in turn specified via a formula F' containing quantified variables $w_1, \ldots, w_j$, some of whose types contain t_1. More generally, VCP2 requires that its types be configured so that no sequence of the following kind occur: t_1 is specified 'in terms of' the quantified variables $v_1, \ldots, v_i$, some of whose types contain t_2, which is specified 'in terms of' the quantified variables $w_1, \ldots, w_j$, some of whose types contain $t_3, \ldots t_n$, which is specified 'in terms of' the quantified variables $z_1, \ldots, z_k$, some of whose types contain t_1. In other words, VCP2 requires that its types be configured so as to realise a well-founded ordering. There are of course many kinds of well-founded ordering and, so, one may ask which particular ordering is so realised. Unfortunately, VCP2 does not have anything to say to this question. Also, since the circumstance that the types realise a well-founded ordering does not determine whether or not the types are cumulative, one may ask whether or not they are. Here, again, VCP2 does not have anything to say.

Interestingly, many have taken the following configuration of types as the most plausible well-founded ordering (some have even taken it as the only possible one): First, consider the collection of all terms in the realm of being that Russell calls *Things*. Take this collection as a type and call it *order-0*. Next, consider the collection of propositional functions that apply to all and only order-0 items with sense and whose existence need not be affirmed by comprehension axioms whose s-formulae contain quantified variables ranging over items not falling under order-0. The propositional functions of this collection must exist if any do. The collection is closed under the operation of specifying a propositional function by means of a comprehension axiom whose s-formula possibly contains bound variables ranging over order-0, free parameter variables ranging over the collection, but no bound variables ranging over it. Note that since such an s-formula contains no bound variables ranging over the collection, VCP2 does not require that a propositional function specified by means of it fall under a type different from that of any of its possible free parameter variables. Take the collection in question as a type and call it *order-1*. Next, consider the collection of propositional functions that apply either to all and only order-1 items with sense or to all and only order-0 items with sense or to both, and whose existence need not be affirmed by comprehension axioms whose s-formulae contain bound variables ranging over items not falling under order-0 or under order-1. This collection is closed under the operation of specifying a propositional function by means of a comprehension axiom whose s-formula possibly contains bound variables ranging over order-0 or order-1, free parameter variables ranging over the collection, but no bound variables ranging over it. Note that since such an s-formula contains no bound variables ranging over the collection, VCP2 does not require that a propositional function specified by means of it fall under a type different from that of any of its possible free parameter variables. Take the collection in question as a type and call it *order-2*. Clearly, this procedure may be iterated indefinitely. Moreover, it may be extended without difficulty to cover propositions. In this respect, a proposition or a propositional function of order-i, roughly speaking, may be specified by means of a comprehension axiom whose s-formula possibly contains bound variables of orders i and orders less than i but contains no bound variables of higher orders. The result of such an iteration and extension is a well-founded ordering in which there figures every type of every term in the realm of being. The types are noncumulative and, since there is a least type -- i.e., order-0 -- as opposed to several minimal types, the types of the propositional functions are *well-ordered*.

There are two points to make about the configuration of types described above. On the one hand, the description is only a sketch. In order to provide a more precise characterisation of the configuration, one would have to specify, among other things, how the configuration relates to the logical system that concerns it. On the other hand, since VCP2 already underdetermines the particular kind of well-founded ordering that the types realise as well as whether or not the types are cumulative, it *a fortiori* underdetermines such relatively fine details of the configuration.

Before I move on to discuss difficulties, I should note that the comprehension axioms that are appealed to in the arguments to the semantic paradoxes all involve s-formulae that are viciously circular in the sense of the explanation recommended by VCP2. For instance, the argument to the liar paradox requires the existence of a proposition that says of itself that it is false. This is affirmed by the following comprehension axiom:

$$\exists q(q \leftrightarrow \exists p(\psi(p) \wedge \neg p))$$

where $\lambda x.\psi x$ is some propositional function true of q and only q -- recall the discussion of the liar paradox in Chapter 1, §2. Here, the quantified variable 'p' ranges over the proposition q and, so, the s-formula is viciously circular. In this respect, to the extent that no viciously circular formula is to be countenanced as a legitimate s-formula, the deduction of the semantic paradoxes, at least by means of the arguments described in Chapter 1, §2, is obstructed. To this extent, one may say that violation of VCP2 gives rise to such paradoxes, where such violation may be taken to consist in employing viciously circular formulae as s-formulae.

(c) Problems:

I now consider two problems concerning VCP1 and VCP2. To provide a little context, I should summarise some of the above discussion. Around 1905-6 Russell concluded, roughly speaking, that the modern paradoxes arise from a mischaracterisation of the comprehension axioms and that this mischaracterisation consists in its countenancing all possible formulae as s-formulae when only a very restricted collection of such formulae may be so countenanced. Following Poincaré, Russell took this collection to consist of those formulae that are not *viciously circular* and stated his vicious-circle principle in an effort to provide an explanation of the conditions for a formula's being viciously circular. He stated the principle in several renderings and, so, in order to arrive at a clear understanding of the explanation that he intended the principle to provide, I have examined two of these renderings, VCP1 and VCP2.

The first problem: In the light of the examination of VCP1 and VCP2, we have seen that they offer different explanations of the conditions for a formula's being viciously circular. Roughly speaking, by VCP1's lights, a formula is viciously circular if and only if it is a formula of the following kind whose variables are not restricted to types that form a well-founded ordering: $x(x)$, $x(y) \wedge y(x)$, $x(y) \wedge y(z) \wedge z(x)$, $x \in x$. By VCP2's lights, a formula F is viciously circular when and only when, if F were to be countenanced as an s-formula, it would contain at least one quantified variable whose range contained the term whose existence would be affirmed by the comprehension axiom having F as its s-formula. Thus, strictly speaking, Gödel's claim that Russell's various renderings of the vicious-circle principle express different principles is correct. Perhaps, one may construe VCP1 and VCP2 as offering explanations that are mutually compatible in the sense that each explanation only states a possible condition for a formula's being viciously circular and, so, a possible condition for a formula's failing to be an s-formula. To this extent, one would have to construe VCP1 and VCP2 each as indicating a source of the modern paradoxes. This construal, however, appears incompatible with one of Russell's positions on the nature of the solution to the paradoxes, *viz.*, that all the paradoxes arise from the same error and, so, their solution must be unitary. Although such an incompatibility may be resolved by countenancing a rather disjunctive notion of error, we shall see a more satisfactory resolution below.

The second problem: Whereas many are impressed by VCP1's intuitive appeal -- as we saw above -- many find it difficult to accept VCP2. Indeed, Gödel and Quine argue against it. Thus, Gödel writes:

> If, however, it is a question of objects that exist independently of our constructions, there is nothing in the least absurd in the existence of totalities containing members which can be described (i.e. uniquely characterized) only by reference to this totality.[141]

Speaking of classes, Quine writes:

> For we are not to view classes literally as created through being specified -- hence as dated one by one, and as increasing in number with the passage of time. ... The doctrine of classes is rather that they are there from the start. .. It is reasonable to single out a desired class by citing any trait of it, even though we chance thereby to quantify over it along with everything else in the universe.[142]

By their lights, if propositions and propositional functions are construed *realistically* -- that is, if they are construed as existing in the realm of being independently of our activities -- then they may be specified by means of s-formulae containing quantified variables whose ranges contain them. In other words, by their lights, VCP2 is incompatible with a realistic construal of propositions and propositional functions.

To illustrate the point, Quine considers the description 'the most typical Yale man'. The logically explicit expression of this description involves quantified variables that range over all Yale scores including those of the person it specifies and, as such, it violates VCP2. However, to the extent that the relevant domain of quantification may be construed realistically, such a description is clearly innocuous.

---

[141] "Russell's Mathematical Logic", pp. 127-8

[142] *Set Theory and its Logic*, p. 243.

Because Russell put forward VCP2, Gödel and Quine conclude that Russell construed propositions and propositional functions *constructivistically*. That is, they conclude that he construed them as items that are in some sense 'constructed' by us. In this respect, perhaps they are brought into being when we affirm their existence by means of comprehension axioms. If so, then of course we may not use s-formulae containing quantified variables whose ranges contain such items whose existence we are affirming.

Quine makes it clear that any attempt to give a coherent explanation of a constructivist construal of propositions and propositional functions would likely appeal to temporal notions and, as such, it would doubtfully succeed. For consider the difficulties encountered if, in order to affirm the existence of a propositional function by means of a comprehension axiom, one had to employ an s-formula containing quantified variables whose ranges contained only terms whose existence had *already* been affirmed.

Of course, the consideration of such difficulties is irrelevant since it is clear that Russell did not construe propositions and propositional functions constructivistically. By his lights, propositions and propositional functions are terms in the realm of being which are such that each has the nature that it does independently of the circumstances of every other term in the realm of being. To this extent, Gödel and Quine draw a false conclusion. However, their discussion still presents us with the following problem: to explain how VCP2 may be compatible with a realistic construal of propositions and propositional functions.

Interestingly, the two problems described above may be resolved by attending to the details of LT(POM). Consider the second problem first. Gödel and Quine may be taken to argue successfully that, on -- what one may call for lack of a better nomenclature -- the standard account of comprehension,[143] VCP2 *is* incompatible with a realistic construal of a given domain of quantification. Although LT(POM)'s account of comprehension approaches very closely the standard account, it differs from it in two salient respects. By appealing to these two respects, one may explain how VCP2 may be compatible with a realistic construal of propositions and propositional functions. Goldfarb points up the two respects in his RRR.

The first respect in which LT(POM)'s account of comprehension differs from the standard account is that according to LT(POM)'s account, on an abstract level, for any proposition or propositional function specified by means of an s-formula, the form of such a specified term is the same as the form of the s-formula at least in the very *weak* sense that, if t is a segmentable item that figures in the latter, then there will be a term in the realm of being that is the meaning of t which figures in the former. By contrast, according to the standard account, for any item specified by means of an s-formula, there is no interesting sense in which the form of the item, to the extent that it may be taken to have a form, is the same as the form of the s-formula. Goldfarb may be read as pointing up this contrast in the following:

> ...the comprehension axioms for propositions and propositional functions that are implicit in the system involve not so much the specification of these entities as the presentation of them. One is not characterizing a proposition or propositional function: one is giving it.[144]

The contrast may be highlighted by considering Zermelo-Fraenkel set theory. No one would claim that, for any set specified by means of an s-formula, the form or structure of the set is the same as that of the s-formula. At best, one would claim that the s-formula picks out all of the

---

[143]Here, for the purposes of exposition, I am supposing there is such a thing as a standard account of comprehension or, more precisely, a standard account of what is involved in the affirmation of comprehension axioms. What such an account would be should become clear from the discussion of that with which it will be contrasted: the LT(POM) account of comprehension.

[144]RRR, p. 32.

set's members and that these alone determine its form or structure.

Although Goldfarb does not explicitly say so, the claim that, by LT(POM)'s lights, a proposition/propositional function and its s-formula in a weak sense have the same form follows from two features of LT(POM). First, roughly speaking, although LT(POM)'s implicit comprehension axioms for propositions and propositional functions may be taken as quantified biconditionals having the respective forms:

$$\exists p(p \leftrightarrow \_\_\_)$$

$$\exists f \forall v_1 \ldots \forall v_n(f(v_1,\ldots,v_n) \leftrightarrow ---),$$

where '___' represents any formula and '---' represents any formula containing no free occurrence of f, these axioms may alternatively be taken as quantified identities having the respective forms:

$$\exists p(p = \lambda.\_\_\_)^{145}$$

$$\exists f(f = \lambda v_1 \ldots \lambda v_n.---).$$

In this respect, 'p' and the formula represented by '___' are singular terms meaning the same term in the realm of being. The same holds for f and the abstraction expression obtained by prefixing the formula represented by '---' with the appropriate abstraction operators. That the comprehension axioms may be taken in this alternative way follows from the ontological commitments implicit in LT(POM)'s operation of propositional function abstraction.[146]

[145]The lone lambda operator acts as a nominalising device here.

[146]Although I have not discussed this aspect of LT(POM), we shall see it when we look at LT(PM)'s propositional function abstraction in Chapter 5. Strictly speaking, for any formula F and for any variables $u_1, \ldots, u_i$ occurring free in F, there will be the abstraction expression represented by $\lambda u_1 \ldots \lambda u_i.F$. Then, from the identity having the form '$\lambda u_1 \ldots \lambda u_i.F = \lambda u_1 \ldots \lambda u_i.F$', one may existentially generalise to obtain the statement having the form: $\exists f(f = \lambda u_1 \ldots \lambda u_i.F)$. In the case of LT(PM), the type of the variable f must be the type of the abstraction expression.

Secondly, recall that by Russell's lights,[147] the formal language of the logical theory perspicuously mirrors the realm of being at least in the weak sense that, for any formula meaning some proposition, if t is a segmentable item figuring in the formula, then there will be a term that is the meaning of t figuring in the proposition and, further, that for any propositional function abstraction expression E meaning some propositional function, if t is a segmentable item figuring in the formula from whose abstraction E results, then there will be a term that is the meaning of t figuring in the propositional function. It should be clear that this second feature of LT(POM) in conjunction with the first implies the claim in question, *viz.*, that a proposition/propositional function and its respective s-formula in a weak sense have the same form.

The second respect in which LT(POM)'s account of comprehension differs from the standard account is that, according to LT(POM)'s account, a variable *qua* term in the realm of being presupposes in Russell's sense the range over which it varies as well as the terms belonging to this range. Goldfarb argues toward this claim in RRR. Indeed, he suggests the even stronger claim that, according to LT(POM)'s account, a variable is identical to its range. We saw, however, in Chapter 2 that this is impossible.

Before we look at Goldfarb's argument, we should see what follows from the claim at hand. Briefly, suppose the negation of VCP2. Then there is a comprehension axiom, say, of the following form whose s-formula contains a quantified variable g whose range contains the term f whose existence the axiom affirms:

$$\exists f \forall v_1 \ldots \forall v_n \{ f(v_1, \ldots, v_n) \leftrightarrow (g)(\text{---}g\text{---}) \}$$

By the claim pertaining to the first respect -- call it *claim 1* -- the propositional function f has

---

[147]See Chapter 1, §1.

the same form as the s-formula at least to the extent that the variable g must figure in f. Hence, f presupposes g. By the claim pertaining to the second respect -- call it *claim 2* -- g presupposes the terms belonging to its range and, so, g presupposes f. By the transitivity of presupposition, f presupposes itself. Since presupposition is irreflexive, this consequence clearly repugns VCP1. Thus, insofar as VCP1, claim 1, and claim 2 are held, the original supposition must be rejected. In other words, VCP1 together with claims 1 and 2 implies VCP2.

Now, consider Goldfarb's argument to claim 2:

> I wish only to point to a consequence of having variables that lack complete generality. Once such variables are used, the question of the nature of the variable (as an entity) becomes far more urgent. Different variables can have different ranges; it then appears that our understanding of a proposition or a propositional function that contains quantified variables will depend quite heavily on an understanding of what those ranges are. The variable must carry with it some definite information; it must in some way represent its range of variation. Therefore, I would speculate, Russell takes a variable to presuppose the full extent of its range.[148]

Goldfarb's statement of the argument is terse. Perhaps, it may be expanded as follows. Recall that Russell has a compositional account of understanding according to which in order to understand a complex such as a proposition, one must understand its parts as well as how they are combined. Recall also that Russell takes it that one cannot understand in a 'direct way' a term having infinite complexity but, rather, one can only understand such a term by means of understanding another term having finite complexity that represents the former term in some way.[149] In this light, consider a quantified propositional function, say, $\lambda y(x).\phi xy$ in which the variable x's range is infinite. By the compositional account of understanding, someone j's

---

[148]RRR, p. 37.

[149]See *Principles*, §72; here of course Russell appeals to the LT(POM) relation of denoting.

understanding of $\lambda y(x).\phi xy$ will depend on her understanding of its quantified variable x. This understanding in turn will depend on her understanding of the range over which x varies. Note that since VCP1 requires that the range of any variable be restricted, this latter dependence is non-trivial. Since the range over which x varies has infinite complexity, j's understanding of this range must come by way of the circumstances that she understands x and that x represents the range over which it varies. To the extent that x does represent such a range, perhaps x may be taken to presuppose it. This last step is of course rather quick and would have to be argued for. I forego any such undertaking here.

Interestingly, Goldfarb's argument may also be expanded into a *de-epistomologised* form: By Russell's lights, the identity of the propositional function $\lambda y(x).\phi xy$ supervenes on the identity of the variable x. The identity of x supervenes on the range over which it varies. The identity of this range supervenes on the identity of the terms that belong to it. Thus, the identity of $\lambda y(x).\phi xy$ supervenes on the identity of the terms belonging to the range of x. To the extent that supervenience approximates presupposition, Goldfarb's conclusion -- that is, claim 2 -- follows.

Let us prescind from considering the difficulties of Goldfarb's argument and return to the second problem posed above: to explain how VCP2 may be compatible with Russell's realistic construal of propositions and propositional functions. As we saw above, VCP1 together with claims 1 and 2 implies VCP2. Thus, insofar as VCP1 and claims 1 and 2 may be taken as salient features of Russell's particular realistic construal, not only do we see how VCP2 *is* compatible with it, but we also see that such a construal indeed requires VCP2.

Next, turn to the first problem posed above: to the extent that VCP1 and VCP2 may be construed each as indicating a source of the paradoxes, to explain how such a construal may be compatible with Russell's position that all the paradoxes arise from the same error. At this point, it should be clear how the explanation should go. By Russell's lights, the one error in question is violation of VCP1. When we looked at VCP1, we saw how such violation occurs in the arguments to the set-theoretic and mixed paradoxes. When we looked at VCP2, we saw how its violation occurs in the arguments to the semantic paradoxes. And just above, we saw how violation of VCP2 crucially involves violation of VCP1.

Recall that I said at the beginning of this section that, roughly speaking, because Russell took there to be some sort of circularity responsible for the paradoxes, he put forward his vicious-circle principle. However, because Quine did not see any such circularity, he concluded that Russell was confused about the matter. In the light of the above discussion, one should come to a conclusion different from Quine's. That is, one should conclude that, to the extent that the vicious-circle principle may in essence be identified with VCP1, there is a circularity to which Russell attended, *viz.* the circularity involved in non-well-founded structures.

Before I conclude this chapter, I should offer a brief review. In 1900, Russell put forward a characterisation of logic, *viz.*, the logical theory LT(POM) implicit in *Principles*. In 1901, however, he derived his class paradox within LT(POM) and other logicians derived several other paradoxes soon afterward. In the face of these paradoxes, Russell adopted two salient positions on the nature of their solution. First, the paradoxes arise from some mischaracterisation of logic and, so, their solution must consist in LT(POM)'s reform. Secondly, all of the paradoxes arise from the same error and, so, their solution must be unitary. Thus, the error must be a single mischaracterisation of LT(POM). In 1905-6 Russell took the mischaracterisation in question to consist in LT(POM)'s construal of the comprehension axioms. More precisely, he took it to consist in the construal's countenancing all possible formulae as s-formulae when only those that abide by the vicious-circle principle may be so countenanced.

Much of our attention recently has been focused exclusively on how to interpret this particular principle. At this point, I should address a question that one may have been prompted to ask for some time, *viz.*, what justifies Russell's identification of the mischaracterisation responsible for the paradoxes. To respond, I cite three reasons. First, as we saw when we looked at VCP1 and VCP2, if only those formulae that abide by these renderings of the vicious-circle principle are countenanced as s-formulae, none of the paradoxes arise, at least by means of their familiar arguments. Secondly, Russell's identification of the mischaracterisation clearly satisfies his two positions on the nature of the solution to the paradoxes. Thirdly, the vicious-circle principle, insofar as it may in essence be taken as VCP1, is intuitively appealing, if not compelling.

To the extent that these reasons justify Russell's particular identification of the mischaracterisation of logic, he may take logic's correct characterisation as provided by the result of modifying LT(POM) in order to abide by the vicious-circle principle. In 1908 Russell adumbrated such a result, LT(PM), in his "Mathematical Logic as Based on the Theory of Types" and in 1910 he explained it in detail in *Principia Mathematica*. Although the modification in question has been described as merely an adjustment of which formulae are to be countenanced as s-formulae, strictly speaking it of course is much more extensive. As we saw when we looked at VCP1 and VCP2, significant structural changes to LT(POM) must be effected in order that the result abide by these renderings of the vicious-circle principle. These changes must be such that logical categories -- that is, types, levels, and orders -- play a far more serious role in the result than they do in LT(POM).

In the next two chapters, I look at this result, LT(PM), in detail. In Chapter 4, I confine my attention to *Principia Mathematica*'s *Theory of Deduction*, the propositional fragment of LT(PM). In Chapter 5, I consider the rest of LT(PM), the *Theory of Apparent Variables*, LT(PM)'s quantification theory.

Chapter 4

The Theory of Deduction

In this chapter and the next I explain the logical theory of *Principia Mathematica* LT(PM).[150] More specifically, in this chapter I explain the propositional fragment of LT(PM), which Russell dubs *the theory of deduction*, and in the next chapter I explain the quantificational fragment of LT(PM), which he dubs *the theory of apparent variables*. Russell presents the propositional fragment of LT(PM) in *Principia*'s Chapters *1 through *5 and its quantificational fragment in Chapters *9 through *21. For convenience, call the former fragment $LT_d(PM)$ and the latter fragment $LT_{av}(PM)$.

Recall that in Chapters 2 and 3 I said that LT(PM) may be understood as a result of modifying the logical theory implicit in *Principles*, LT(POM), in order to abide by the vicious-circle principle. As such, LT(PM)'s fragments LT_d(PM) and LT_{av}(PM) may be understood as *theories of types* at least in the sense that logical categories play a serious role within them. For expository purposes, I shall for the most part prescind from considering LT_d(PM)'s type-theoretic features while explaining LT_d(PM) in this chapter. These features are a subset of the type-theoretic features of LT_{av}(PM) and I explain these latter features in detail in Chapter 5. Although my prescinding should cause no confusion, one should remain cognizant that every term in LT_d(PM)'s ontology is of particular logical categories such as types and so are many elements of LT_d(PM)'s vocabulary.

I should remark now that Russell's presentation

[150]Greek letters are used throughout this chapter as metalinguistic variables. In particular, "α", "α'", "β", and "β'" are generally used to range over variables while "ϕ", "ϕ'", "ψ", and "ψ'" are generally used to range over formulae.

in *Principia* of LT_d(PM) at times falls short of contemporary standards of rigour and clarity. For instance, it does not explicitly distinguish the sundry roles played by its axioms, rules of inference, and even rules of formation, usually simply calling them all "principles". It also does not explicitly distinguish semantic considerations from syntactic ones.

In contrast to my explanation of LT(POM) in Chapter 2, my explanation of LT_d(PM) will largely concern its formal elements. However, their semantic analogues will always be clear.

1. The Ontology of LT_d(PM):

Apart from the modifications required in order to abide by the vicious-circle principle, the ontology of LT_d(PM) is very similar to that of the propositional fragment of LT(POM). It contains propositions, variables ranging over propositions, negation functions, and disjunction functions. The differences between the two ontologies here in question may be roughly summarised as follows -- VCP1 and VCP2 require most of the differences: (i) every term in LT_d(PM)'s ontology is of certain logical categories such as types; (ii) the range of any LT_d(PM) variable is some type; (iii) whereas the material implication function and the property of being true are primitive logical constants of LT(POM), negation functions and disjunction functions are the primitive logical constants of LT_d(PM); these are like the material implication function in that they are terms in the ontology in question on a par with all other terms and in that they are genuine functions; they differ from LT(POM)'s material implication function in that their scopes of application are types.

Before moving on to the next section, I should remark that Russell actually asserts in Chapter II of the Introduction to the first edition of *Principia* that propositions do not exist:

> Owing to the plurality of the objects of a single judgment, it follows

> that 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. That is to say, the phrase which expresses a proposition is what we call an "incomplete" symbol; it does not have meaning in itself, but requires some supplementation in order to acquire a complete meaning. This fact is somewhat concealed by the circumstance that judgment in itself supplies a sufficient supplement, and that judgment in itself makes no *verbal* addition to the proposition. Thus "the proposition 'Socrates is human'" uses "Socrates is human" in a way which requires supplement of some kind before it acquires a complete meaning; but when I judge "Socrates is human," the meaning is completed by the act of judging, and we no longer have an incomplete symbol. The fact that propositions are "incomplete symbols" is important philosophically, and is relevant at certain points in symbolic logic.[151]

This quotation expresses part of Russell's *multiple-relation theory of judgment*, a theory that he begins to articulate at the time of *Principia*'s publication. Clearly, insofar as LT(PM) has quantified variables ranging over propositions,[152] it is difficult to see how Russell can consistently make the claims quoted above while holding that LT(PM) is a correct characterisation of logic. Accordingly, I shall charitably discount such claims for the purposes of this exposition.

2. <u>Vocabulary:</u>

As every term in LT_d(PM)'s ontology is of certain logical categories, so are many of the elements of LT_d(PM)'s vocabulary. More precisely, if a term i in LT_d(PM)'s ontology is of logical categories $t_1, ..., t_n$, and an element e of LT_d(PM)'s vocabulary either refers to or ranges over i, then e will be of logical categories $t_1, ..., t_n$ as well.[153]

---

[151] *Principia*, p. 44.

[152] See Chapter *14.

[153] Of course, i and e cannot be of $t_1, ..., t_n$ in the same sense.

Variables:

For every LT_d(PM) type t, LT_d(PM) has an infinite alphabet of variables.[154] Each of these is to be thought of as superscripted by some visible index of t. As is well-known, Russell in fact never visibly superscripted any of the LT(PM) variables presented in *Principia*. Concerning these variables, he expected his readers to follow his convention of *typical ambiguity*, according to which one is to supply such variables with whatever indices may be required.

Concerning the semantics of LT_d(PM)'s variables, if t is an LT_d(PM) type, the variables of type t · ·e taken to range over propositions of that type.

Logical Constants:

LT_d(PM) possesses as primitive truth-functional connectives negations signs and disjunction signs. More precisely, for every type of LT_d(PM) formula, LT_d(PM) possesses a negation sign that applies to LT_d(PM) formulae of that type -- the type of a formula is explained in Chapter 5, §6. The case is similar for the disjunction signs. Accordingly, these truth-functional connectives are typed in the sense that they apply to formulae of only certain types.

Grouping Devices:

LT_d(PM) possesses as grouping devices dots, parentheses, and braces.

3. Rules of Formation:

Russell puts forward in *Principia* propositions *1.7 and *1.71 in order to specify which sequences of primitive particles are to count as LT_d(PM) formulae:

---

[154] LT_d(PM) types will be explained in Chapter 5.

*1.7 If p is an elementary proposition, $\neg p$ is an elementary proposition. (Pp)[155]

*1.71 If p and q are elementary propositions, $p \vee q$ is an elementary proposition. (Pp)[156]

Although *1.7 and *1.71 are apparently correct parts of an inductive specification, they suffer three difficulties: they are simply classified as 'primitive propositions' (Pp) and, as such, their particular role is not made discernible from the roles played by LT_d(PM)'s axioms and rules of inference; they do not distinguish between syntactic and semantic considerations; and they do not make explicit what the priorities of grouping are that would guarantee the unique readability of compounds. Concerning the second difficulty, it may be thought that Russell is to be excused since he thinks that the logical constants of negation and disjunction are actually in the propositions that they help to compose. Accordingly, to that extent he can ambiguously refer by the same means with impunity to the syntactic sequence of signs used to express a given proposition p and the proposition p itself. However, Russell cannot be so easily excused because it is doubtful that he would take the syntactic devices -- the dot and the parenthesis/brace -- that are required for the unique readability of compounds -- or their semantic analogues, whatever these would be -- to be in the propositions for whose expression

[155]Note about *Principia*'s Reference Numbers: In order to facilitate reference to propositions and definitions -- such reference is required in the annotations of demonstrations -- every proposition and definition in *Principia* is introduced with a decimal number which is immediately preceded by an asterisk (The asterisk is needed because Russell uses 'plain' numerals to label propositions occurring in his deductions). The integral part of a proposition's (definition's) reference number is also used to indicate the chapter in which such a proposition (definition) is introduced. Thus, if "*n.d" refers to a given proposition (definition), "*n" will refer to its chapter. Russell usually, but not always, follows the convention of introducing a definition at the beginning of its chapter with a reference number whose decimal part is less than 0.1. So, for instance, he introduces the definition for the subset predicate at the beginning of Chapter *22 with the reference numeral "*22.01."

[156]*Principia*, p. 97.

they are so required. In that case, he cannot ambiguously refer to propositions and their expressions with impunity. This point brings us to the third difficulty. Although Russell makes no mention in *1.7 and *1.71 of any syntactic devices required to guarantee unique readability, in the execution of *Principia* he employs Peano's dot notation in concert with parentheses and braces. A brief description of the dot notation is given in Chapter I of the Introduction to the first edition of *Principia*,[157] but how it is to be used must be gathered from his practice.[158]

Bracketing consideration of the circumstance that the logical constants are typed in the sense described in §2 above, below I present $LT_d(PM)$'s rules of formation, which *1.7 and *1.71 attempt to articulate. Unlike *1.7 and *1.71, however, the rules thus presented are explicitly concerned with $LT_d(PM)$'s formulae. They also make explicit what the priorities of grouping are by specifying how the dot and parenthesis/brace notations are to be employed. In order to make their specification inductively, the rules inductively define a function dn (dot number) from expressions to natural numbers.

Atomic: Any variable α is an atomic formula, and $dn(\alpha)=-1$;

Negation: If α is a variable, $\ulcorner\neg\alpha\urcorner$ is a formula and $dn(\ulcorner\neg\alpha\urcorner)=0$;

If ϕ is a formula which is not a variable, $\ulcorner\neg(\phi)\urcorner$ is a formula and $dn(\ulcorner\neg(\phi)\urcorner)=0$;

Disjunction: If α and β are variables, $\ulcorner\alpha\vee\beta\urcorner$ is a formula and $dn(\ulcorner\alpha\vee\beta\urcorner)=0$;

If α is a variable and ϕ is a formula which is not, $\ulcorner\alpha\vee A\phi\urcorner$ is a formula where

[157] P. 9.

[158] Although Russell makes no attempt to give inductive rules governing the dot notation, it is interesting that rather simple ones governing both dot and parenthesis/brace notations can be stated. These are presented just below.

> A is a configuration of $dn(\phi)+1$ dots and $dn(\ulcorner \alpha \vee A \phi \urcorner)=dn(\phi)+1$;[159]
>
> If ϕ is a formula which is not a variable and α is a variable, $\ulcorner \phi A \vee \alpha \urcorner$ is a formula where A is a configuration of $dn(\phi)+1$ dots and $dn(\ulcorner \phi A \vee \alpha \urcorner)=dn(\phi)+1$;
>
> If ϕ and ψ are both formulae which are not variables, $\ulcorner \phi A \vee B \psi \urcorner$ is a formula where A is a configuration of $dn(\phi)+1$ dots and B is a configuration of $dn(\psi)+1$ dots and $dn(\ulcorner \phi A \vee B \psi \urcorner)=max[dn(\phi),dn(\psi)]+1$.

Braces may be used in place of parentheses.

This definition might appear overly complicated. In particular, if one is only interested in the mechanics of the dot and parenthesis/brace notation, Russell's own description in *Principia*'s Introduction might seem more appropriate. It is certainly simpler and more intuitive:

> Dots immediately preceded or followed by "V" ... serve to bracket off a proposition. ... The general principle is that a larger number of dots indicates an outside bracket, a smaller number indicates an inside bracket. ... The scope of the bracket indicated by any collection of dots extends backwards or forwards beyond any *smaller* number of dots ... until we reach either the end of the asserted proposition or a *greater* number of dots or an *equal* number ... [D]ots only work away from the adjacent sign of disjunction ... [pp. 9-10]
>
> In a proposition containing several signs of [disjunction], the one with the greatest number of dots before or after it is the *principal* one ... [p. 11]

Although Russell's description is heuristically useful, it is unfortunately incomplete. Specifically, it does not say how a negation sign relates to the dot and a parenthesis/brace in a

---

[159]The dots are of course configured as follows:

1 dot	.
2 dots	:
3 dots	:.
4 dots	::
5 dots	::.
etc.	

formula.

4. Definitions:

LT_d(PM) possesses definitions concerning the material implication, conjunction, and equivalence signs. Before we look at these, I ought first to review what Russell explicitly says about definition in general in *Principia*.[160,161]

In Chapter I of the Introduction to the first edition of *Principia*, Russell writes:

> A definition is a declaration that a certain newly-introduced symbol or combination of symbols is to mean the same as a certain other combination of symbols of which the meaning is already known.[162]

Evidently, this comment is very similar to a comment about definition that Frege makes in his *Begriffsschrift*, where he says that a definition stipulates that its *definiendum* is "to have the same content" as its *definiens*.[163]

In the next paragraph, Russell makes the following three familiar-sounding remarks about definition: a definition is concerned with its *definiendum* and *definiens qua* formal objects, not with what they symbolise; a definition is always theoretically unnecessary -- that is, its *definiens* may be used in any place where its *definiendum* is used and, so, this latter is only a mere "typographical convenience"; as typographical conveniences, however, the *definienda* are extremely useful, since without appeal to them, formulae would become too long to be readable.

---

[160]Here, we shall see that there is much in common between what Russell and Frege have to say about the matter.

[161]I shall focus on explicit definition here. Russell later introduces contextual definitions.

[162]*Principia*, p. 11.

[163]*From Frege to Gödel*, p. 55.

These few remarks of Russell's about definition raise the question whether one should regard each new definition that is put forward in the course of *Principia* as metatheoretically specifying an extension of the current language by the addition of the *definiendum* as a primitive particle and as likewise specifying a conservative extension of the current theory by the addition of the definition as an axiom.[164] Or, should one regard each new definition that is put forward in the course of *Principia* as metatheoretically specifying a new notation, the *definiendum*, which is to be considered as an abbreviation for and to be paraphrased away by an old notation, the *definiens*, whenever one is concerned with the official language and theory of $LT_d(PM)$. Regarded in this light, a definition does not extend the current language or theory in any way; it merely indicates how certain symbols are to be provisionally used as short for the expressions belonging to the official language and theory.

Unfortunately, Russell's above-cited remarks about definition do not answer this question in any decisive way. Some remarks are compatible with both conceptions of definition, some only support the former while others only support the latter. His remarks that each definition makes for shorter formulae and that it is metatheoretic in nature are compatible with either of the two conceptions of definition. His remark that each definition declares that its *definiendum* is to mean the same as its *definiens* speaks against the latter conception for, on that conception, the *definiendum* does not mean anything at all.[165] Later we shall see that one of $LT_d(PM)$'s implicit yet primitive rules of inference allows the occurrence of any *definiendum*

---

[164]Kunen explains such conservative "1-step" extensions in his *Set Theory: An Introduction to Independence Proofs*, Ch. 1, §13.

[165]By comparison, it is noteworthy that after making the similar-sounding remark about identical content in *Begriffsschrift*, Frege says something that clearly supports the former conception: a definition "is immediately transformed into a [judgement]". That is, it is to be treated as an axiom after its introduction.

within a formula to be substituted by its *definiens* and vice versa. The admission of such a rule coheres much better with the former conception of definition than with the latter, since a system's primitive rules of inference are generally understood to apply to pieces of its official language and those only. By contrast, what Russell says in *Principia* about contextual definition and incomplete symbols would seem on the face of it specifically to support the latter conception.[166] Since nothing very significant depends upon which conception of definition one ascribes to $LT_d(PM)$, I shall treat any $LT_d(PM)$ definition according to the former conception. That is, I shall treat it as specifying a conservative extension of $LT_d(PM)$.

Russell's method of stating an explicit definition in *Principia* is similar to the method that Frege employs in *Begriffsschrift*. After the corresponding asterisk numeral, the definition's *definiendum* is written, followed by the identity sign "=", followed by its *definiens*, finally followed by the symbols "Df".[167] Dots are used between these four items if they are required to indicate the intended grouping. The general form of a definition is therefore:

$$*n.m \quad x_1x_2...x_i.=.y_1y_2...y_j \qquad Df.$$

The definitions concerning the material implication signs are represented in *Principia* in typically ambiguous fashion by proposition *1.01:

$$*1.01. \quad p\supset q.=.\neg p\vee q \qquad Df$$[168]

Here, an occurrence of 'p' or 'q' is an occurrence of an $LT_d(PM)$ variable, not an occurrence of a compound formula. The definitions concerning the conjunction signs and the equivalence signs are likewise represented in *Principia* in typically ambiguous fashion by propositions *3.03

[166]See Chapter III of the Introduction to the first edition and Chapter *14.

[167]The identity sign without the "Df" is used in *Principia* later on to mean identity.

[168]*Principia*, p. 94.

and *4.01:

*3.01. $p.q. =. \neg(\neg p \vee \neg q)$ Df[169]

*4.01. $p \equiv q. =. p \supset q. q \supset p$ Df[170]

5. Axioms:

Russell explicitly presents all of LT_d(PM)'s axioms in Chapter *1. Before we look at these, we ought first to see what Russell says about axioms in general in *Principia*. In the Introduction to the first edition of *Principia*, Russell characterises the axioms of a formal system as those formulae which must be assumed without proof and remarks that such formulae are needed since all inference proceeds from some formulae that have been previously 'asserted'.[171] Two points may immediately be made about this first statement. The first is that its appeal to the notions of assumption and assertion sounds rather pragmatic. Clearly, however, a purely syntactic characterisation of axiomhood may be given. The second point is that some inference -- or, rather, some formalisations of inference -- can proceed from no previous formula at all, as in, say, some kinds of natural deduction and sequent calculus formal system. None of these systems requires any particular formulae to be selected as axioms. Of course, Gentzen's discovery of such systems some twenty years after *Principia*'s publication was a significant accomplishment in the development of logic.[172]

In the Introduction Russell further remarks that, like the situation with respect to a

[169]*Principia*, p. 111.

[170]*Principia*, p. 117.

[171]*Principia*, p. 12.

[172]See Gerhard Gentzen's "Investigations into Logical Deduction" (1935) in *The Collected Papers of Gerhard Gentzen*.

formal system's primitive particles, it is to some extent optional which formulae should be selected as the axioms, although the better the axiomatisation the fewer and simpler are the axioms. He goes on to say something curious, *viz.*, that the specific formulae that he has selected as LT_d(PM)'s axioms are not particularly obvious. Obviousness, of course, is not such a reliable criterion of truth in these matters, since principles which had been thought to be unquestionably obvious led to contradiction. But if these axioms are not particularly obvious and obviousness is not such a reliable guide to truth anyway, and if the axioms cannot be proved within LT_d(PM) in any epistemologically interesting way -- the notion of deduction itself appeals to them -- the question naturally arises how their acceptance as axioms is to be justified. Russell answers this question with the following piece of reasoning:

> The proof of a logical system is its adequacy and its coherence. That is: (1) the system must embrace among its deductions all those propositions which we believe to be true and capable of deduction from logical premisses alone, though possibly they may require some slight limitation in the form of an increased stringency of enunciation; and (2) the system must lead to no contractions ...[173]

Interestingly, Russell uses reasoning of this kind to justify his notorious axioms of reducibility.[174] The gist of the reasoning is that the axioms of logic in general and of LT_d(PM) in particular are to be justified in much the same way in which high-level principles of natural science are justified -- that is, by something like the abductive method. Russell first discusses this method of justifying principles of logic -- which he calls the *regressive method* -- in 1907[175] and, in important respects, his discussion anticipates Quine's later doctrines of *holism* and

---

[173]*Principia*, pp. 12-3.

[174]However, Russell does not do so with respect to formulae of *Principia* that express the content of the axioms of infinity and choice. As is well known, he never accepts these formulae as axioms but treats them whenever he is required to do so as mere hypotheses.

[175]See "The Regressive Method of Discovering the Premises of Mathematics"(1907) in EA.

naturalism.[176]

The five formulae that Russell selects as LT_d(PM)'s axioms are by now well-known. I list them as follows along with their names and asterisk numerals.

*1.2	$p \vee p . \rightarrow . p$	Principle of Tautology (Taut.)
*1.3	$q . \rightarrow . p \vee q$	Principle of Addition (Add.)
*1.4	$p \vee q . \rightarrow . q \vee p$	Principle of Permutation (Perm.)
*1.5	$p \vee (q \vee r) . \rightarrow . q \vee (p \vee r)$	Associative Principle (Assoc.)
*1.6	$q \rightarrow r . \rightarrow : p \vee q . \rightarrow . p \vee r$	Principle of Summation (Sum.)

Only a few remarks need be made about the axioms. First, Russell is not unmistakeably clear when he states these five expressions about whether they are axioms or axiom schemata. His later remarks, however, tend to favour construing these expressions as single axioms. For instance, in *Introduction to Mathematical Philosophy* (IMP), he notes that the legitimacy of substitutions of the kind required when *1.2 to *1.6 are treated as axioms must be insured by means of a rule of inference.[177] Secondly, Russell does not clearly distinguish these axioms as fundamentally different with respect to the role that they play within LT_d(PM) from the rules of inference and formation that he puts forward; he simply labels all of these 'primitive propositions'. Thirdly, several reductions of this axiomatisation of propositional logic have been put forward since *Principia*'s publication. Paul Bernays proposed the simplest of these reductions which consists of Russell's axioms save the fourth one, whose non-independence Bernays discovered;[178] J.G.P. Nicod proposed another reduction

[176]See, for instance, "Two Dogmas of Empiricism", "On What There Is", *Word and Object*, and "Epistemology Naturalized" among others.

[177]IMP, p. 151.

[178]*Mathematische Zeitschrift*, vol. 25 (1926), pp. 305-320.

which consists of replacing Russell's axioms by the following four similar-looking ones:[179]

$$p \vee p . \rightarrow . p$$
$$p . \rightarrow . p \vee q$$
$$p \vee (q \vee r) . \rightarrow . q \vee (p \vee r)$$
$$q \rightarrow r . \rightarrow : p \vee q . \rightarrow . p \vee r$$

Likewise, Goetlind and Rasiowa proposed yet another reduction in which Russell's axioms are replaced by the following three similar-looking ones:[180]

$$p \vee p . \rightarrow p$$
$$p \rightarrow . p \vee q$$
$$q \rightarrow r . \rightarrow : p \vee q . \rightarrow . r \vee p$$

It should be noted for the sake of completeness that much terser axiomatisations of propositional logic have been put forward which, for instance, make use of only one sentential connective, one axiom, and two rules of inference.[181,182] However, not only are these axiomatisations dissimilar from LT_d(PM)'s but their primitive connectives are hard to use and their axioms and rules of inference are complicated. Accordingly, interest in them consists principally in the fact that they permit rather simple proofs of metatheorems about them.

5. <u>Rules of Inference:</u>

LT_d(PM) possesses three rules of inference. In Chapter *1 Russell states one such rule, *modus ponens* (MP). The other two are interchange of definitional equivalents and substitution.

---

[179]*Proceedings of the Cambridge Philosophical Society*, vol. 19 (1917-20), pp. 32-41.

[180]*Norsk Matemisk Tidsskrift*, vol. 29 (1947), pp. 1-4; *ibid.*, vol. 31 (1949), pp. 1-3.

[181]LT_d(PM) has 3 primitive rules of inference.

[182]Nicod introduces the following example in the paper cited above.
Connective: Nand;
Axiom: $[p|(q|r)]|\{[t|(t|t)]|[(s|q)|((p|s)|(p|s))]\}$
Rules of inference: substitution, from p and $p|(q|r)$ to r.

Unfortunately, these are never explicitly stated; that they are rules of LT_d(PM) is only made apparent in the Summary of Chapter *2, where Russell explains the method by which a deduction is to be carried out, and in the presentation of the deductions themselves.[183] All three rules of inference are described as follows.

(a) Modus Ponens:

At this point, we have seen all of the explicitly enunciated LT_d(PM) primitive propositions save the one that specifies MP -- every one of these others has either been a rule of formation or an axiom. Ironically, as if it were in some respect more fundamental than these primitive propositions, Russell puts forward the primitive proposition that specifies MP pages before any of these others. It is proposition *1.1:

*1.1 Anything implied by a true elementary proposition is true.[184]

That *1.1 is a specification of MP may not be apparent at first blush. This difficulty arises partly from its having a rather semantical flavour.

Several remarks may be made about *1.1. First, unlike *1.2 through *1.6 (the axioms), *1.1 is not expressed in the object language of LT_d(PM). In this regard, Russell says that although one might initially think that such expression is possible, it clearly is not. He suggests that MP may be initially thought to be specified by the following formula of LT_d(PM): $p \rightarrow : p \rightarrow q . \rightarrow q$. But then he responds:

> This is a true proposition, but it holds equally when p is not true and when p does not imply q. It does not, like the principle we are concerned with, enable

---

[183]In IMP, p. 151, Russell acknowledges the failure to introduce explicitly in *Principia* LT_d(PM)'s rule of substitution.

[184]*Principia*, p. 94.

us to assert q simply, without any hypothesis.[185]

Here, Russell is alluding in a somewhat obscure way to the important distinction between axioms and rules of inference, which Lewis Carroll had vividly illustrated.[186] Russell goes on to make the following curious remark:

> We cannot express the principle symbolically, partly because any symbolism in which *p* is variable only gives the *hypothesis* that *p* is true, not the *fact* that it is true.[187]

Of course, one *can* express the rule of inference symbolically; I shall do so shortly. What *is* correct is rather that one cannot express it in the object language of LT_d(PM).[188] The question then arises why Russell concludes that MP cannot be specified symbolically. One plausible consideration has to do with the metalevel/objectlevel distinction. It appears that Russell does not clearly recognise the metatheoretic character of specifications concerning LT_d(PM). Thus, he does not clearly acknowledge the metatheoretic character either of the specification of LT_d(PM)'s rules of inference or that of its axioms. Both of these, however, talk about LT_d(PM) formulae, not about the subject matter of which such formulae treat. Accordingly, to specify a rule of inference of LT_d(PM) is to specify which formulae may be appended to any given sequence of formulae that constitutes an LT_d(PM) deduction. And to specify the axioms of LT_d(PM) is just to say which LT_d(PM) formulae are to be so-called. The fact that one way of doing so is by displaying such formulae perhaps misleadingly encourages one to think that the axioms of a formal system may be specified in that system's object language. It is thus

---

[185]*ibid.*

[186]See "What the Tortoise said to Achilles" in *Mind*, N.S. Vol.iv, p. 278, 1895.

[187]*ibid.*

[188]I bracket considerations of coding the metalanguage into the object language here.

plausible to suppose that because Russell does not clearly recognise the metatheoretic character of such specifications, he does not try to formulate them symbolically in anything but the object language. And whereas he might appear to succeed with respect to the axioms, he manifestly does not respect to MP. The result is his conclusion that MP is not symbolically specifiable at all.[189]

Secondly, as noted above, Russell's specification of MP is peculiarly semantical. *1.1 appeals to propositions, implication, and truth rather than to the respective syntactic counterparts formulae, the horseshoe, and deduction. What makes this especially curious is that Frege, by contrast, in his *Grundgesetze der Arithmetik*[1893] -- a work with which Russell was familiar -- had already specified the rules inference for his formal system in a purely syntactic way -- something that was considered to be a great advance in the development of logic. Moreover, Frege distinctly recommended this method of specification in order to obtain deductions that enjoy the quality of so-called 'gaplessness'.[190] As I mentioned Chapter 1, §1, Russell in carrying out the work of *Principia* also strives for gapless deduction. Thus, in these respects, his semantical specification of MP is particularly incongruous. The incongruity may be partially explained away by the following two plausible but somewhat speculative

---

[189]The fact that Russell does not clearly distinguish $LT_d(PM)$'s axioms from its rules of inference -- calling them all 'primitive propositions' -- also makes it plausible to suppose that he also mistakenly presumes that since many of the primitive propositions -- *viz.*, the axioms -- *appear* to be specifiable in the language of $LT_d(PM)$, MP should be also. After all, by his lights, *1.1 is not noticeably different from these others. MP is not, however, specifiable in the language of $LT_d(PM)$ and Russell in fact acknowledges this situation, as I cited above. As a result of this insuperability, perhaps Russell simply concludes that the rule cannot be specified 'symbolically' at all. This suggestion is not as odd as it might at first appear. Russell makes a similar conclusion with regard to describing the theory of types, something on a par with specifying a formal system's rule of inference. He acknowledges that, in describing the theory, he must violate its own type restrictions. Such description, therefore, cannot be grammatically stated within the theory and, so, cannot be stated 'symbolically' at all.

[190]See *Foundations of Arithmetic*, p. 103; *Begriffsschrift*, pp. 5-6.

considerations. To begin, it is plausible that the semantical flavour of *1.1 may simply be the result of Russell's often-noted failure to attend carefully in all cases to the distinction between syntax and semantics. In other words, *1.1 may perhaps best be construed as a poorly worded expression of what Russell actually means to say by it: a purely syntactic specification of MP:

For any formulae ϕ and ψ, from ϕ and $\ulcorner\phi\rightarrow\psi\urcorner$, infer ψ.

Next, it is plausible to suppose that Russell thinks that to the extent that MP is clearly specifiable at all, it should be specifiable in the object language of LT_d(PM). As I said in Chapter 1,§1, he views logic in general as providing an all-encompassing framework within which anything with sense can be said and within which all reasoning can be carried out and strictly speaking *only* within such a framework. Since LT_d(PM) is a subpart of logic and since MP cannot be specified within its object language -- as Russell himself notes -- he has likely relinquished any effort to specify the rule in a clear, syntactic manner. The result is *1.1.

(b) Interchange of Definitional Equivalents:

This rule is stated as follows:

> For any two formulae ϕ and ϕ' such that with respect to some LT_d(PM) definition ϕ is a definiens and ϕ' a definiendum, if the two formulae ψ and ψ' are such that ψ contains one or more occurrences of ϕ (ϕ') and ψ' is like ψ except for containing ϕ' (ϕ) at some places where ψ contains ϕ (ϕ'), infer ψ' from ψ.

As I said above, this rule is implicit in *Principia*. One clear application of it occurs in the demonstration of *2.01 (the proposition that $p\rightarrow\neg p.\rightarrow.\neg p$):[191]

From: $\vdash: \neg p \vee \neg p.\rightarrow.\neg p$

To: $\vdash: p\rightarrow\neg p.\rightarrow.\neg p$

[191]*Principia*, p. 100.

Another clear application occurs in the demonstration of *3.26 (the proposition that p.q.¬.p):[192]

From: ⊢: ¬(¬p∨¬q).¬.p

To: ⊢: p.q.¬.p

(c) Substitution:

Although LT_d(PM)'s rules of substitution are implicit, they are very important. They are stated as follows:

First Rule of Substitution:

> If α and α' are variables of the same type and the formulae ϕ and ϕ' are such that ϕ contains one or more occurrences of α and ϕ' is like ϕ except for containing an occurrence of α' in every place where ϕ contains an occurrence of α, from ϕ, infer ϕ'.

A clear application of this rule occurs in the demonstration of *2.15 (the proposition that ¬p¬q.¬.¬q¬p):[193]

From: ⊢: ¬¬(¬p)

To: ⊢: q¬¬(¬q)

Second Rule of Substitution:

> If the formulae ψ and ψ' are such that ψ contains one or more occurrences of variable α and ψ' is like ψ except for containing formula ϕ in every place where ψ contains an occurrence of α, then from ψ, infer ψ'.[194]

---

[192]*Principia*, p. 112.

[193]*Principia*, p. 102.

[194]I am here bracketing type-theoretic considerations. A more comprehensive statement of this rule of inference that does not bracket them is given in Chapter 5 under the rubric *Second*

A clear application of this rule also occurs in the demonstration of *2.15:

From: $\vdash :. q \to r . \to : p \to q . \to . p \to r$

To: $\vdash :. \neg p \to \neg(\neg q) . \to . \neg q \to \neg(\neg p) : \to :. \neg p \to q . \to . \neg p \to \neg(\neg q) : \to : \neg p \to q . \to . \neg q \to \neg(\neg p)$

$[\neg p \to q / p, \neg p \to \neg(\neg q) / q, \neg q \to \neg(\neg p) / r]$

Rule of Substitution for Propositional Variables.

Chapter 5

The Theory of Apparent Variables

In this chapter I explain $LT_{av}(PM)$ -- that is, the fragment of LT(PM) that Russell calls the *Theory of Apparent Variables*.[195] Russell presents $LT_{av}(PM)$ in Chapters *9 through 21 of *Principia*. In so doing he offers two formulations of the theory; he offers one in Chapter *9 as an alternative and he offers the other in Chapters *10 through *21 as the formulation to be employed throughout the rest of *Principia*. Accordingly, my explanation of $LT_{av}(PM)$ will principally concern the latter formulation and, thus, Chapters *10 through *21. To the extent, however, that certain features of this formulation are only characterised in Chapter *9, the explanation will concern this chapter as well.

Recall that in Chapter 4 I said that LT(PM) may be understood as a result of modifying the logical theory implicit in *Principles*, LT(POM), in order to abide by the vicious-circle principle and that, as such, its subtheories $LT_d(PM)$ and $LT_{av}(PM)$ may be understood as *theories of types*. Although I was generally able to prescind from considering $LT_d(PM)$'s type-theoretic features while explaining $LT_d(PM)$ in Chapter 4, to explain $LT_{av}(PM)$ in this chapter I must consider its type-theoretic features from the start.[196]

As is well-known, $LT_{av}(PM)$ has two notions of logical category -- the notion of *type* and the notion of *order*. Concerning the former, several have asked which particular notion it is. On the one hand, in *Principia* Russell explicitly defines a certain relation -- the relation of

---

[195]Greek letters are used throughout this chapter as metalinguistic variables. In particular, "α", "α'", "β", and "β'" are generally used to range over $LT_{av}(PM)$'s variables while "ϕ", "ϕ'", "ψ", and "ψ'" are generally used to range over $LT_{av}(PM)$'s formulae.

[196]Note that these type-theoretic features include those of $LT_d(PM)$.

being of the same type -- which requires a very fine-grained notion of type. On the other hand, this definition apart, Russell's presentation of $LT_{av}(PM)$ in *Principia* is rather inexplicit (he makes few references to types) and, as a result, $LT_{av}(PM)$ may be taken as having any of an entire spectrum of notions of type ranging from the most coarse-grained to the most fine-grained. In this chapter, for purposes of exposition, I first explain $LT_{av}(PM)$ as having a rather coarse-grained notion of type and then, at the end, I explain $LT_{av}(PM)$ as having a rather fine-grained notion of type.

$LT_{av}(PM)$'s notion of order plays the roles played by the notions of *order* and *level* discussed in Chapter 3 -- recall that these two notions were introduced in order to abide by VCP1 and VCP2. As such, $LT_{av}(PM)$'s notion of order in a clear sense obstructs the deduction within $LT_{av}(PM)$ of the set-theoretic, mixed, and semantic paradoxes and, thus, contributes to their unitary solution. Although many have thought otherwise, $LT_{av}(PM)$'s notion of order is more coarse-grained than its notion of type in the sense that, in general, every order may be taken as a superset of several types. Because many have thought otherwise,[197] $LT_{av}(PM)$ has been called a *ramified theory of types* as opposed to a *simple theory of types*. This appellation, however, more suitably fits such theories as ramified analysis whose notion of order is more fine-grained than its notion of type. In any case, as we shall see, $LT_{av}(PM)$ differs only slightly from a simple-type-theoretic version of itself, the difference mainly consisting in subtleties in the rules of substitution.

In understanding LT(PM) -- and, hence, $LT_{av}(PM)$ -- as a result of modifying LT(POM) in order to abide by the vicious-circle principle, one should remain cognizant of the fact that it

---

[197]See I. Copi, *The Theory of Logical Types*, Chap. 3, especially pp. 82-91; S. Kleene, *Introduction to Metamathematics*, p. 44; W. and M. Kneale, *The Development of Logic*, pp. 660ff; and F. Ramsey, "The Foundations of Mathematics", pp. 175-6.

could be only one among many such results -- recall the discussion in Chapter 3 about VCP1 and VCP2's underdetermining the structure of the logical categories there concerned. One should note, however, that LT_{av}(PM) is a particularly attractive one of such results insofar as it may be taken to provide a unitary solution to the paradoxes. This last characterisation of course challenges Ramsey and Quine's criticism that, far from being unitary, it is a rather *ad hoc* combination of two separable theories, one superimposed upon the other, in which the superimposing theory -- the theory of orders -- specifically obstructs the deduction of the semantic paradoxes and the theory superimposed upon -- the theory of types proper -- specifically obstructs the deduction of the set-theoretic paradoxes. My contrary characterisation will be made good in what follows.

Needless to say, Russell's presentation in *Principia* of LT_{av}(PM) is similar to that of LT_d(PM) in that it suffers the same flaws. For instance, it does not distinguish syntactic and semantic considerations nor does it distinguish the roles played by the axioms, rules of inference, and rules of formation. As Chapter 4 focused enough upon these flaws, not much attention will be given them here.

As with LT_d(PM), the explanation of LT_{av}(PM) will largely concern its formal elements. However, their semantic analogues will always remain apparent.

1. The Ontology of LT_{av}(PM):

Apart from the modifications required in order to abide by the vicious-circle principle, the ontology of LT_{av}(PM) is very similar to that of LT(POM). In addition to containing everything that the ontology of LT_d(PM) contains -- that is, propositions, variables ranging over propositions, negation functions, and disjunction functions -- the ontology of LT_{av}(PM) contains Things, concepts, propositional functions, variables ranging over these terms, and universal quantifier functions. The differences between the ontologies of $LT_{av}(PM)$ and LT(POM) may be roughly summarised as follows -- VCP1 and VCP2 require most of the differences: (i) trivially, $LT_{av}(PM)$ calls its Things *individuals* and its concepts *matrices*; (ii) every term in $LT_{av}(PM)$'s ontology is of a particular type and order; (iii) the range of any $LT_{av}(PM)$ variable is some type; (iv) for n>0, the scope of application of any $LT_{av}(PM)$ n-adic propositional function is in the relevant sense some n types; (v) the scope of application of any $LT_{av}(PM)$ negation function is some type; the scope of application of any universal quantifier function is some type; and the scope of application of any disjunction function is in the relevant sense some two types;[198] (vi) $LT_{av}(PM)$'s ontology contains no classes. In Chapter *20 of *Principia* Russell very cleverly shows how class terms may be introduced as "incomplete symbols" by contextual definition. As a result, he is able to simulate class theory within $LT_{av}(PM)$ and, accordingly, $LT_{av}(PM)$ has been called a *no-class theory of classes*.

Before moving on to the next section, I should remark that some have claimed that there are aspects of Russell's 1905-6 substitutional theory which are apparent in $LT_{av}(PM)$.[199] For instance, certain of such aspects might be taken as apparent in Russell's definition of being of the same type. However correct their claim may be, it is important to understand that there is no interesting sense in which the ontology of $LT_{av}(PM)$ is similar to that of the substitutional theory. Recall that although the ontology of the substitutional theory contains propositions, it does not contain propositional functions. In their stead, the theory countenances a primitive operation of substitution S such that given any terms c and d and a proposition p containing c as constituent, S(p,d,c) is the proposition that results from substituting d for c in p.

---

[198]See the Summary of Chapter *9 in *Principia*.

[199]See, for instance, Goldfarb's RRR, pp. 36-8.

2. <u>Vocabulary:</u>

As every term in $LT_{av}(PM)$'s ontology is of a particular type and order, so are many of the elements of $LT_{av}(PM)$'s vocabulary. More precisely, if an term i in $LT_{av}(PM)$'s ontology is of a particular type t and order o and an element e of $LT_{av}(PM)$'s vocabulary either refers to or ranges over i, then e will be of type t and order o as well.[200] Accordingly, I begin this section by explaining $LT_{av}(PM)$'s types and orders.[201] I then consider $LT_{av}(PM)$'s variables and logical constants *qua* formal objects.

Orders:

$LT_{av}(PM)$'s orders are natural numbers including zero.

Types:

$LT_{av}(PM)$'s types are ordered pairs of *pretypes* and orders. These are defined recursively as follows:

Base Step:

> The empty sequence $\langle\rangle$ is a pretype. For any order n, the ordered pair of the empty sequence and n, $\langle\langle\rangle, n\rangle$, is a type.

Inductive Step:

> If $\langle p_1,o_1\rangle, \langle p_2,o_2\rangle, \ldots, \langle p_i,o_i\rangle$ are types, $i>0$, then the sequence s of these types is a pretype. If m is the least number greater than the orders of $o_1, o_2, \ldots, o_i$, then, for any $n \geq m$, n is an *order appropriate to* the pretype s and the ordered pair of s and n, $\langle s, n\rangle$, is a type.

---

[200] Of course, i and e cannot be of t and o in the same sense.

[201] Although my characterisation of $LT_{av}(PM)$'s coarse-grained types and orders as given below differs from that given by A. Church, "Comparison of Russell's Resolution of the Semantical Antimonies with that of Tarski", I am indebted to him for his characterisation.

One should note that although the types defined above might appear unfamiliar, considered without their order components they are very similar to the types characteristic of simple type theories.

Variables:

For every type t, LT_{av}(PM) has an infinite alphabet of variables. Each of these is to be thought of as superscripted by some visible index of t.

Concerning the semantics of LT_{av}(PM)'s variables, if a type $t = <<>, 0>$, the variables of type t are taken to range over the individuals and, as such, are called *individual variables*. If a type $t = <<t_1,t_2,...,t_i>, o>$ for some types t_1, t_2, ..., t_i and appropriate order o, the variables of type t are taken to range over propositional functions of that type where the arguments of such functions are in the relevant sense of types t_1, t_2, ..., t_i. As such, the variables of type t are called *propositional function variables*. If a type $t = <<>, n>$, $n>0$, the variables of type t are taken to range over propositions of that type and, as such, are called *propositional variables*.

Russell calls those propositional function variables -- as well as the functions over which such variables range -- whose orders are the least possible appropriate to their pretypes *predicative*. Note that this is a third sense of the word. As we shall see, all of LT_{av}(PM)'s variables are predicative in the second sense, *viz.*, that of being of logical categories that respect VCP2.

Logical Constants:

In addition to the primitive truth functional connectives of LT_d(PM) -- that is, the negation signs and disjunction signs -- LT_{av}(PM) possesses, for every one of its variables α, a universal quantifier (sign) $\ulcorner(\alpha)\urcorner$. Like the truth functional connectives, every universal quantifier is typed in the sense that it is taken to apply to formulae of only a certain type -- the type of a formula is explained in §6 below. Although Russell makes this restriction clear in the Summary of Chapter *9, following his lead I shall largely bracket consideration of it in future discussion.

Grouping Devices:

$LT_{av}(PM)$ possesses the same grouping devices as $LT_d(PM)$ -- that is, dots, parentheses, and braces -- and employs them in the same way.

3. Rules of Formation:

The rules of formation of $LT_{av}(PM)$ are a generalisation of those of $LT_d(PM)$ and, as such, countenance the formulae of the latter theory as a proper subset of those of the former. Bracketing consideration of the circumstance that the logical constants are typed in the sense described in §2, I present $LT_{av}(PM)$'s rules of formation below.[202,203]

Atomic Formulae:

> Any propositional variable is an atomic formula. For $i>0$, if v_1 is a variable of type t_1, v_2 a variable of type t_2, ..., v_i a variable of type t_i, and f a propositional function variable whose pretype is $<t_1,t_2,\ldots,t_i>$, then $\ulcorner fv_1v_2\ldots v_i\urcorner$ is an atomic formula.

Compound Formulae:

> If ϕ is a formula, $\ulcorner\neg\phi\urcorner$ is a formula. If ϕ and ψ are formulae, $\ulcorner\phi\vee\psi\urcorner$ is a formula. If ϕ is a formula and α is a variable, then $\ulcorner(\alpha)\phi\urcorner$ is a formula.

---

[202]Here, I also bracket consideration of the mechanics of $LT_{av}(PM)$'s grouping devices as this was explained in Chapter 4.

[203]In contrast to the case with $LT_d(PM)$'s rules of formation, Russell in *Principia* puts forward only one primitive proposition concerning $LT_{av}(PM)$'s rules of formation -- proposition *10.121. The proposition is mentioned below. In any case, there is no serious doubt about what $LT_{av}(PM)$'s rules of formation are.

Three points are relevant here. First, as is well-known, although Russell never visibly superscripted any of the LT_{av}(PM) variables presented in *Principia*, in certain formulae containing predicative variables he indicated that such variables are predicative by placing an exclamation point after some of their occurrences. Secondly, insofar as LT_{av}(PM)'s rules of formation do not countenance as formulae both $\ulcorner fu_1u_2...u_j\urcorner$ and $\ulcorner fv_1v_2...v_j\urcorner$ where for some i, $1 \leq i \leq j$, u_i and v_i are of different types, they require that the types in general be noncumulative in the sense discussed in Chapter 3. Interestingly, in *Principia* Russell puts forward only one primitive proposition that concerns LT_{av}(PM)'s rules of formation, *10.121, and what it states is just this requirement:

> *10.121. If "ϕx" is significant, then if a is of the same type as x, "ϕa" is significant, and vice versa.[204]

With regard to *10.121, Russell writes:

> It follows from this proposition that two arguments to the same function must be of the same type; for if x and a are arguments to $\phi\hat{x}$, "ϕx" and "ϕa" are significant, and therefore x and a are of the same type.[205]

Thirdly, LT_{av}(PM)'s rules of formation guarantee that LT_{av}(PM) abides by VCP1 in the sense that, for any $i>0$, there is no sequence of i formulae respectively having the following forms: $f_1(f_2)$, $f_2(f_3)$, ..., $f_{i-1}(f_i)$, and $f_i(f_1)$. The reason is that LT_{av}(PM)'s types as defined above realise a well-founded ordering.

4. <u>Definitions:</u>

In addition to LT_d(PM)'s definitions concerning the implication, conjunction, and equivalence

---

[204] *Principia*, p. 140.

[205] *ibid.*

signs, LT_{sv}(PM) possesses definitions concerning the existential quantifier and identity signs. The definitions concerning the existential quantifier signs are represented in *Principia* in typically ambiguous fashion by proposition *10.01:

*10.01. $(\exists x).\phi x.=.\neg(x).\neg\phi x$ Df[206]

Here, an occurrence of 'ϕ' is an occurrence of a propositional function variable, not an occurrence of a well-formed context containing occurrences of the variable 'x'.

The definitions concerning the identity signs -- there is one such sign for every type -- are represented in *Principia* in typically ambiguous fashion by proposition *13.01:

*13.01. $x=y.=:(\phi):\phi!x.\rightarrow.\phi!y$ Df[207]

One should note carefully that although *13.01 *qua* particular instance resembles the standard definition of identity given in second-order logic, it differs from such a definition in one important respect: its 'ϕ' does not range over every propositional function that may apply with sense to x and y. This difference gives rise to the following problem. There appears the possibility that, although every ϕ that applies to x likewise applies to y and, so, x is identical to y in the sense of *13.01, there exists a non-predicative propositional function ψ that applies to x but not to y and, so, x cannot really be identical to y. As we shall see, LT_{sv}(PM) has a special group of axioms -- the axioms of reducibility -- that guarantee that such a problematic possibility cannot obtain.[208]

---

[206] *Principia*, p. 140.

[207] *Principia*, p. 169.

[208] The problem at issue here may be stated more formally as follows: If identity is defined in terms of *13.01, then, unless appeal may be made to principles of a rather special sort (such as the axioms of reducibility), some instances of Leibniz's Law will be beyond derivation.

5. Axioms:

$LT_{av}(PM)$ has four kinds of axiom. They are the propositional axioms, the axioms of universal instantiation, the axioms of confinement of quantifier, and the axioms of reducibility. Russell explicitly put forward the axioms of universal instantiation and confinement of quantifier in Chapter *10 and the axioms of reducibility in Chapter *12.

The Propositional Axioms:

$LT_{av}(PM)$'s propositional axioms are precisely those of $LT_d(PM)$.

The Axioms of Universal Instantiation:

$LT_{av}(PM)$'s axioms of universal instantiation are represented in *Principia* in typically ambiguous fashion by proposition *10.1:

*10.1. $\vdash: (x).\phi x. \supset .\phi y$[209]

Here, as in *10.01, 'ϕ' is a propositional function variable, not a well-formed context. Accordingly, the standard requirement that 'y' be free for 'x' is satisfied so long as both be of the type required by 'ϕ'. As we shall see in the next section, $LT_{av}(PM)$'s rules of substitution for propositional function variables allow the substitution of well-formed contexts for propositional function variables provided that certain requirements are satisfied. As one might have expected, one of the requirements is that of a variable's being free for another and, in this instance, its satisfaction is non-trivial.

[209] *Principia*, p. 140.

The Axioms of Confinement of Quantifier:

LT_{av}(PM)'s axioms of confinement of quantifier govern the prenexing behaviour of quantifiers. They are represented in *Principia* in typically ambiguous fashion by proposition *10.12:

$$*10.12 \quad \vdash:.(x).p \vee \phi x. \rightarrow :p. \vee .(x).\phi x$$ [210]

Here, 'p' is a propositional variable, not a formula in general. The analogous remark applies to 'ϕ'. Since there is no sense in which a propositional variable may contain a free occurrence of another variable, the standard requirement that no free occurrence of 'x' occur in 'p' is trivially satisfied. As we shall see in the next section, LT_{av}(PM)'s rules of substitution for propositional variables allow the substitution of formulae in general for propositional variables provided that certain requirements are satisfied. One of the requirements is analogous to the above one, *viz.*, that a substituent formula contain no free occurrences of variables that would be captured upon substitution. In contrast to the case with the above requirement, the satisfaction of this one is non-trivial.

The Axioms of Reducibility:

LT_{av}(PM)'s axioms of reducibility are represented in typically ambiguous fashion by propositions *12.1 and *12.11:

$$*12.1. \qquad \vdash:(\exists f):\phi x. \leftrightarrow_{x}. f!x$$
$$*12.11. \qquad \vdash:(\exists f):\phi(x,y). \leftrightarrow_{x,y}. f!(x,y)$$ [211]

In *12.1, 'ϕ' is a monadic propositional function variable and, in *12.11, 'ϕ' is a dyadic propositional function variable. Although Russell does not explicitly say so, there must be axioms of reducibility concerning propositional function variables of every arity.

---

[210] *ibid.*

[211] *Principia*, p. 167.

Roughly speaking, these axioms claim that, for every propositional function, there is a predicative propositional function coextensive to it. The axioms apart, no element of $LT_{av}(PM)$ guarantees that its ontology is such as they claim. The import of their claim may be illustrated by the following two considerations. First, consider a particular axiom of reducibility:

$$\ulcorner\exists f\forall v_1 \ldots v_n(f!v_1 \ldots v_n \leftrightarrow \phi v_1 \ldots v_n)\urcorner \qquad (1)$$

As we shall see, $LT_{av}(PM)$'s rules of substitution allow the substitution of any formula F containing free occurrences of $v_1, \ldots, v_n$ but no free occurrences of f for the atomic formula $\ulcorner\phi v_1 \ldots v_n\urcorner$. To this extent, for any such F, the comprehension formula

$$\ulcorner\exists f\forall v_1 \ldots v_n(f!v_1 \ldots v_n \leftrightarrow F)\urcorner \qquad (2)$$

may be inferred from (1). In *Principia* Russell indeed employs comprehension formulae derived by means of $LT_{av}(PM)$'s axioms of reducibility in order to comprehend wanted propositional functions.[212]

The significance of the circumstance just pointed up -- *viz.*, that the axioms of reducibility allow the derivation of comprehension formulae -- might appear mitigated by consideration of the following further circumstance: For any formula F of the kind in question, the axioms of universal instantiation and the rules of substitution -- as we shall see -- allow the derivation of a comprehension formula similar to (2) above:

$$\ulcorner\exists G\forall v_1 \ldots v_n(G!v_1 \ldots v_n \leftrightarrow F)\urcorner \qquad (3)$$

(2) and (3), however, differ in the following important respect. Whereas f is of the least order appropriate to the orders of its arguments -- that is, it is predicative -- G need not be. In fact, depending on the construction of F, G may be of arbitrarily high order. In virtue of this difference -- as we shall see in the next section -- depending on the construction of F, (2) may

---

[212]See *Principia*, p. 180, the demonstrations concerning *14.17 and *14.171; p. 191, the demonstration concerning *20.112.

be impredicative -- in the modern sense -- while (3) may never be.

Secondly, as I said with respect to $LT_{av}(PM)$'s definition of identity in §4, the axioms of reducibility guarantee that the problematic possibility there envisaged cannot obtain. For suppose that x is identical to y in the sense of *13.01. Thus, $(\phi)(\phi!x \rightarrow \phi!y)$. Suppose further that there is some impredicative propositional function ψ such that ψx. Then, in virtue of an appropriate axiom of reducibility, there must be some predicative propositional function π coextensive to ψ and, so, πx. Since $(\phi)(\phi!x \rightarrow \phi!y)$, πy and, so, ψy.[213]

In *Principia* and elsewhere Russell observed that, in contra... the other $LT_{av}(PM)$ axioms, the axioms of reducibility do not indisputably present themselves as logical truths. Notwithstanding, he included them among $LT_{av}(PM)$'s axioms in order to make good $LT_{av}(PM)$'s definition of identity and, more importantly, in order to enable $LT_{av}(PM)$ to simulate enough class theory to accomplish the reduction of arithmetic. To justify their inclusion, Russell used his *regressive method*. In this respect, he wrote:

> That the axiom of reducibility is self-evident is a position which can hardly be maintained. But in fact self-evidence is never more than a part of the reason for accepting an axiom, and is never indispensable. The reason for accepting an axiom, as for accepting any other proposition, is always largely inductive, namely that many propositions which are nearly indubitable can be deduced from it, and that no equally plausible way is known by which these propositions could be true if the axiom were false, and nothing which is probably false can be deduced from it. If the axiom is apparently self-evident, that only means, practically, that it is nearly indubitable; for things have been thought to be self-evident and have yet turned out to be false. And if the axiom itself is nearly indubitable, that merely adds to the inductive evidence derived from the fact that its consequences are nearly indubitable: it does not provide new evidence of a radically different kind. Infallibility is never attainable, and therefore some element of doubt should al··ays attach to every axiom and to all its consequences. In formal logic, the element of doubt is less than in most sciences, but it is not absent, as appears from the fact that the paradoxes followed from premisses which were not previously known to require

[213]Formally speaking, if identity is defined in terms of *13.01, the axioms of reducibility allow the derivation of all instances of Leibniz's Law.

> limitations. In the case of the axiom of reducibility, the inductive evidence in its favour is very strong since the reasonings which it permits and the results to which it leads are all such as appear valid.[214]

6. Rules of Inference:

$LT_{av}(PM)$'s rules of inference consist of modus ponens, interchange of definitional equivalents, universal generalisation, alphabetic change of bound variables, and substitution. They are described as follows.

(a) Modus Ponens:

$LT_{av}(PM)$ inherits this rule unchanged from $LT_d(PM)$. It is restated by proposition *9.12:

> *9.12. What is implied by a true premiss is true.[215]

(b) Interchange of Definitional Equivalents:

$LT_{av}(PM)$ inherits this rule unchanged from $LT_d(PM)$. For convenience, it is restated as follows.

> For any two formulae ϕ and ϕ' such that with respect to some $LT_{av}(PM)$ definition ϕ is a definiens and ϕ' a definiendum, if the two formulae ψ and ψ' are such that ψ contains one or more occurrences of ϕ (ϕ') and ψ' is like ψ except for containing ϕ' (ϕ) at some places where ψ contains ϕ (ϕ'), infer ψ' from ψ.

This rule is implicit in *Principia*. One clear application of it occurs in the demonstration of *10.253 (the proposition that $\neg\{(x).\phi x\}.\equiv.(\exists x).\neg\phi x$):[216]

---

[214]*Principia*, p. 59.

[215]*Principia*, p. 132.

[216]*Principia*, p. 143.

From: ⊢: ¬{(y).¬(¬ϕy)}.→.¬{(x).ϕx}:

To: ⊢: (∃y).¬ϕy. →.¬{(x).ϕx}

(c) Universal Generalisation:

LT_{av}(PM)'s rule of universal generalisation is stated by proposition *10.11:

> *10.11. If ϕy is true whatever possible argument y may be, then (x).ϕx is true.[217]

Here, 'ϕy' and 'ϕx' are metavariables respectively ranging over formulae containing one or more free occurrences of the object variables 'x' and 'y'. The rule is stated more precisely as follows:

> If the formulae ϕ and ϕ' are such that ϕ contains free occurrences of a variable α and ϕ' is like ϕ except for containing a free occurrence of a variable α' of the same type as the variable α in every place where ϕ contains a free occurrence of α, from ϕ, infer ⌜(α')ϕ'⌝.

A clear application of the rule of universal generalisation occurs in the demonstration of *10.2 (the proposition that (x).pVϕx.≡:p.V.(x).ϕx):[218]

From: ⊢:. p.V.(x).ϕx:→.pVϕy:.

To: ⊢:. (y):.p.V.(x).ϕx:→.pVϕy:.

(d) Alphabetic Change of Bound Variables:

LT_{av}(PM)'s rule of alphabetic change of bound variables is implicit in *Principia*. It is stated as follows:

> For any two formulae ⌜(α)ϕ⌝ and ⌜(α')ϕ⌝ such that the variable α is of the same type as the variable α' and ϕ' is like ϕ except for containing a free

---

[217]*Principia*, p. 140.

[218]*Principia*, p. 141.

> occurrence of α' in every place where ϕ contains a free occurrence of α, if the formulae ψ and ψ' are such that ψ contains one or more occurrences of ϕ and ψ' is like ψ except for containing ϕ' in some places where ψ contains ϕ, from ψ, infer ψ'.

Needless to say, this rule may be derived as a metatheorem given $LT_{av}(PM)$'s axioms and other rules of inference.

Two clear applications of the rule occur in the demonstration of *10.22 (the proposition that (x).φx.ψx. ↔ :(x).φx:(x).ψx):[219]

From:	⊢:.	(x).φx.ψx. → :(y).φy:(z).ψz	(4)
and:	⊢:.	(x).φx:(x).ψx: → . (y).φy.ψy	(5)
To:	⊢:.	(x).φx.ψx. ↔ :(x).φx:(x).ψx	
		['y' and 'z' in (4) and (5) are replaced by 'x'.]	

(e) Substitution:

$LT_{av}(PM)$'s rules of substitution are a superset of those of $LT_d(PM)$ and, as such, are very important. Like those of $LT_d(PM)$, $LT_{av}(PM)$'s rules are implicit in *Principia*. They are stated as follows:

Rule of Substitution for Individual Variables:

> If α and α' are individual variables and the formulae ϕ and ϕ' are such that ϕ contains one or more free occurrences of α and ϕ' is like ϕ except for containing a free occurrence of α' in every place where ϕ contains a free occurrence of α, from ϕ, infer ϕ'.

A clear application of this rule of substitution occurs in the demonstration of *13.193 (the proposition that φx.x=y. ↔ .φy.x=y):[220]

---

[219] *Ibid.*

[220] *Principia*, p. 171.

From: $\vdash:\quad \phi x.x{=}y. \to .\phi y.x{=}y$

To: $\vdash:\quad \phi y.y{=}x. \to .\phi x.y{=}x \qquad [y/x,\ x/y]$

First Rule of Substitution for Propositional Function Variables:

> If α and α' are propositional function variables of the same type and the formulae ϕ and ϕ' are such that ϕ contains one or more free occurrences of α and ϕ' is like ϕ except for containing a free occurrence of α' in every place where ϕ contains a free occurrence of α, from ϕ, infer ϕ'.

A clear application of this rule of substitution occurs in the demonstration of *11.7 (the proposition that $(\exists x)(\exists y):\phi(x,y).\vee.\phi(y,x): \leftrightarrow .(\exists x)(\exists y).\phi(x,y)$):[221]

From: $\vdash:.(\exists x)(\exists y):\phi(x,y).\vee.\psi(x,y):\leftrightarrow:(\exists x)(\exists y):\phi(x,y):\vee:(\exists x)(\exists y).\psi(x,y):$

To: $\vdash:.(\exists x)(\exists y):\phi(x,y).\vee.\phi(y,x):\leftrightarrow:(\exists x)(\exists y):\phi(x,y):\vee:(\exists x)(\exists y).\phi(y,x):$

$$[\lambda x\lambda y\{\phi(y,x)\}/\lambda x\lambda y\{\psi(x,y)\}]$$

Another clear application occurs in the demonstration of *10.414 (the proposition that $\phi x \leftrightarrow_x \chi x.\psi x \leftrightarrow_x \theta x. \to :\phi x \leftrightarrow \psi x.\leftrightarrow_x.\chi x \leftrightarrow \theta x$):[222]

From: $\vdash:.\phi x \leftrightarrow_x \chi x.\psi x \leftrightarrow_x \theta x. \to :\phi x \to \psi x.\leftrightarrow_x.\chi x \to \theta x$

To: $\vdash:.\psi x \leftrightarrow_x \theta x.\phi x \leftrightarrow_x \chi x. \to :\psi x \to \phi x.\leftrightarrow_x.\theta x \to \chi x$

$$[\psi/\phi,\ \phi/\psi,\ \theta/\chi,\ \chi/\theta]$$

Before stating the next rules of substitution, I must define both the *type of a formula with respect to an argument-variable sequence* and the *substitution of a formula with respect to an argument-variable sequence for an atomic formula*.[223] First, the type of a formula with respect to an argument-variable sequence: For $n \geq 0$, let ϕ be a formula containing free occurrences of

---

[221] *Principia*, p. 159.

[222] *Principia*, p. 148.

[223] The reader may want to skip over the next few pages as they are rather hirsute.

n variables. For i such that $0 \leq i \leq n$, select i of the n variables in some arbitrary order without repetition: $v_1, v_2, \ldots, v_i$.[224] Call the variables $v_1, v_2, \ldots, v_i$ *argument variables* of ϕ and their sequence an *argument-variable sequence* of ϕ. Call the other n-i free variables of ϕ *parameter variables* of ϕ. Where v_1 is of type t_1, v_2 is of type t_2, ..., v_i is of type t_i, call the sequence $<t_1, t_2, \ldots, t_i>$ the *pretype* of ϕ with respect to $v_1, v_2, \ldots, v_i$. Where o is the least number greater than the orders of $v_1, v_2, \ldots, v_i$, greater than the orders of the quantified variables of ϕ, and greater than *or* equal to the orders of the parameter variables of ϕ, call o the *order* of ϕ with respect to $v_1, v_2, \ldots, v_i$. Finally, call the ordered pair of the pretype of ϕ with respect to $v_1, v_2, \ldots, v_i$ and the order of ϕ with respect to $v_1, v_2, \ldots, v_i$ the *type* of ϕ with respect to $v_1, v_2, \ldots, v_i$. In this respect, it should be clear that, for any LT_{av}(PM) formula having an argument-variable sequence, there is an LT_{av}(PM) propositional or propositional function variable such that the variable and the formula with respect to the sequence are of the same pretype, order, and type.

Secondly, the substitution of a formula with respect to an argument-variable sequence for an atomic formula: Let $\ulcorner fa_1a_2 \ldots a_i \urcorner$ be an atomic formula and let the formula ϕ with an argument-variable sequence $v_1, v_2, \ldots, v_i$ be such that f and ϕ with respect to $v_1, v_2, \ldots, v_i$ are of the same pretype. Then, the substitution of ϕ with respect to $v_1, v_2, \ldots, v_i$ for $\ulcorner fa_1a_2 \ldots a_i \urcorner$ is formed from ϕ by putting a_1 for every free occurrence of v_1, a_2 for every free occurrence of v_2, ..., a_i for every free occurrence of v_i.

Second Rule of Substitution for Propositional Function Variables:

> Let the propositional function variable f and the formula ϕ with respect to an argument-variable sequence $v_1, v_2, \ldots, v_i$ be of the same type. Let the formulae ψ and ψ' be such that ψ contains one or more free occurrences of f in *predicate position* only and ψ' be like ψ except for containing, for any variables $a_1, a_2, \ldots, a_i$ of appropriate type, the substitution of ϕ with respect to $v_1, v_2, \ldots, v_i$ for

[224] Note that 0 variables may be selected.

$\ulcorner fa_1a_2...a_i \urcorner$ in every place where ψ contains an occurrence of $\ulcorner fa_1a_2...a_i \urcorner$ whose occurrence of f is free. If (i) ψ contains no subformula $\ulcorner(\omega)\Gamma\urcorner$ such that ω is a parameter variable of φ and Γ contains a free occurrence of f in predicate position; and if (ii) for any variables $a_1, a_2, ..., a_i$ for which ψ contains an occurrence of $\ulcorner fa_1a_2...a_i \urcorner$ whose occurrence of f is free, φ contains no subformula $\ulcorner(a_1)\Gamma_1\urcorner$ that contains a free occurrence of v_1, no subformula $\ulcorner(a_2)\Gamma_2\urcorner$ that contains a free occurrence of v_2, ..., and no subformula $\ulcorner(a_i)\Gamma_i\urcorner$ that contains a free occurrence of v_i; then, from ψ, infer ψ'.

A clear application of this rule of substitution occurs in the demonstration of *13.22 (the proposition that $(\exists z)(\exists w).z=x.w=y.\phi(z,w). \equiv .\phi(x,y)$):[225]

From: $\vdash :.(\exists z)(\exists w).\phi z.\psi(z,w). \equiv :(\exists z):\phi z:(\exists w).\psi(z,w):$

To: $\vdash :.(\exists z)(\exists w).z=x.w=y.\phi(z,w). \equiv :(\exists z):z=x:(\exists w).w=y.\phi(z,w):$

$[\lambda z\{z=x\}/\lambda z\{\phi z\}, \lambda z\lambda w\{w=y.\phi(z,w)/\lambda z\lambda w\{\psi(z,w)\}]$

Another clear application occurs in the demonstration of *13.21 (the proposition that $z=x.w=y.\supset_{z,w}.\phi(z,w): \equiv .\phi(x,y)$):[226]

From: $\vdash ::\phi z.\psi(z,w).\supset_{z,w}.\chi(z,w): \equiv :.\phi z.\supset_z:\psi(z,w).\supset_w.\chi(z,w):$

To: $\vdash ::z=x.w=y.\supset_{z,w}.\phi(z,w): \equiv :.z=x.\supset_z:w=y.\supset_w.\phi(z,w):.$

$[\lambda z\{z=x\}/\lambda z\{\phi z\}, \lambda z\lambda w\{w=y\}/\lambda z\lambda w\{\psi(z,w)\}, \lambda z\lambda w\{\phi(z,w)\}/\lambda z\lambda w\{\chi(z,w)\}]$

First Rule of Substitution for Propositional Variables:

If α and α' are propositional variables of the same type and the formulae φ and φ' are such that φ contains one or more free occurrences of α and φ' is like φ except for containing a free occurrence of α' in every place where φ contains a free occurrence of α, from φ, infer φ'.

Second Rule of Substitution for Propositional Variables:

Let the propositional variable p and the formula φ with respect to the empty argument-variable sequence be of the same type. Let the formulae ψ and ψ' be

---

[225]*Principia*, p. 171.

[226]*Ibid.*

> such that ψ contains one or more free occurrences of p none of which is in *subject position* and ψ' be like ψ except for containing ϕ in every place where ψ contains a free occurrence of p. If ψ contains no subformula $\ulcorner(\omega)\Gamma\urcorner$ such that ω is a parameter variable of ϕ and Γ contains a free occurrence of p, then from ψ, infer ψ'.

Clear applications of this rule and the previous one occur throughout Chapters *1 to *5 -- as we saw in Chapter 4.

There are several points to note about LT_{av}(PM)'s rules of substitution. First, given how complicated some of the rules are, it is curious that *none* of them received explicit statement in *Principia*. In his 1919 *Introduction to Mathematical Philosophy*, Russell acknowledged that he ought to have but failed to offer such statement in this work.[227] Russell, however, is not alone as one who ought to have but failed. Hilbert and Ackermann in their first and second editions of *Grundzüge der Theoretische Logik* intended to offer explicit statement of the rules of substitution but did so only inadequately. In their first edition (1928),[228] they offered only a very rough statement of the second rule of substitution for propositional function variables and, although they offered a better statement of this rule in their second edition (1938),[229] here they omitted mention of condition (ii) above. Carnap also intended to offer explicit statement of the rules of substitution in his *Logical Syntax of Language* as did Quine in his *A System of Logistic*, although neither fully succeeded.[230,231] Indeed, Church and Quine have claimed that

---

[227]See *Introduction to Mathematical Philosophy*, p. 151.

[228]See *Grundzüge der Theoretische Logik*, first edition, p. 54.

[229]See *Grundzüge der Theoretische Logik*, second edition, pp. 56-7.

[230]In his statement of the second rule of substitution for propositional function variables, Carnap also omitted mention of condition (ii) above. See *Logical Syntax of Language*, pp. 90f.

[231]See Church, *Introduction to Mathematical Logic*, pp. 288-9.

an adequate explicit statement was not forthcoming until the publication of Hilbert and Bernay's *Grundlagen der Mathematik*.[232]

Secondly, as I said in connection with LT_{av}(PM)'s axioms of reducibility, LT_{av}(PM)'s rules of substitution together with its axioms of universal instantiation allow the derivation of comprehension formulae. To see that this is so, choose any formula ϕ and any argument-variable sequence $v_1, v_2, ..., v_n$ appropriate to ϕ. Let t be the type of ϕ with respect to $v_1, v_2, ..., v_n$ and let F and G be n-ary propositional function variables of type t. Then, we have:

$Fv_1v_2...v_n \leftrightarrow Fv_1v_2...v_n$	Tautology
$\exists G(Gv_1v_2...v_n \leftrightarrow Fv_1v_2...v_n)$	EG[233]
$\exists G(Gv_1v_2...v_n \leftrightarrow \phi)$	Substitution

In this respect, the conjunction of the rules of substitution with the axioms of universal instantiation carries a serious ontological commitment. Quine has observed that such a commitment is 'strangely inconspicuous'.[234] That it is must in part be owing to the circumstance that neither the rules without the axioms nor the axioms without the rules carry the ontological commitment in question. Indeed, the rules without the axioms are ontologically 'harmless' in the sense that their addition to any first-order theory merely brings about a *conservative extension* since they can be derived as metatheorems.

Thirdly, recall that the order o of a formula ϕ with respect to an argument-variable sequence $v_1, v_2, ..., v_i$ is defined as the least number greater than the orders of the variables $v_1, v_2, ..., v_i$, greater than the orders of the quantified variables of ϕ, and greater than or equal to

---

[232]See Church, *ibid.*; Quine, *Methods of Logic*, p. 181. Richard Heck claims that Frege offered an adequate explicit statement in *Grundgesetze der Arithmetik*.

[233]In *Principia* Russell derives '$\phi y . \rightarrow . (\exists x).\phi x$' (*10.24) by appealing only to tautologies and an axiom of universal instantiation.

[234]See *Set Theory and its Logic*, p. 257.

the orders of the parameter variables of ϕ. The reasons for such a definition should be made apparent. The order o must be greater than the orders of ϕ's argument variables $v_1, v_2, \ldots, v_i$ in order to abide by VCP1 -- recall that VCP1 requires that the types realise a well-founded ordering. The order o must be greater than the orders of ϕ's quantified variables in order to abide by VCP2. If o were otherwise, comprehension formulae could be derived in the way described above whose s-formulae contained quantified variables some of whose ranges contained the terms whose existence such comprehension formulae would affirm. The order o must be greater than or equal to the orders of ϕ's parameter variables in order to abide, again, by VCP2. For suppose that o were just one less than the highest of the orders of the parameter variables concerned. Then the following comprehension formula (4) could be derived (here, the superscripts indicate orders):

$$\exists m^3(x^2)(y^2)\{m^3(x^2,y^2) \leftrightarrow [(a^1)(x^2(a^1)\rightarrow y^2(a^1))\wedge\psi^4(n^3)]\} \qquad (4)$$

Since the comprehension formula (5) could be derived apart from the supposition in question:

$$\exists\psi^4(j^3)\{\psi^4(j^3) \leftrightarrow (l^3)[\phi^4(l^3)\rightarrow\pi^4(l^3,j^3)]\} \qquad (5),$$

the comprehension formula (6) could be derived in turn:

$$\exists m^3(x^2)(y^2)\{m^3(x^2,y^2) \leftrightarrow [(a^1)(x^2(a^1)\rightarrow y^2(a^1)) \wedge (l^3)[\phi^4(l^3)\rightarrow\pi^4(l^3,n^3)]]\} \qquad (6).$$

Since (6)'s s-formula contains a quantified variable, l^3, whose range contains the term whose existence (6) affirms, (6) violates VCP2.

Fourthly, note that $LT_{av}(PM)$ differs from a simple-type-theoretic version of itself only in that its second rule of substitution for propositional function variables and its second rule of substitution for propositional variables both ultimately make essential appeal to the notion of the order of a formula with respect to an argument-variable sequence which is defined as described just above. As we saw, if this notion were modified in any of various ways, comprehension formulae whose s-formulae violated VCP2 could be derived.

7. <u>Abstraction:</u>

The description of $LT_{av}(PM)$ at this point is complete. In this section I describe a certain conservative extension of it, $LT_{av+}(PM)$, which in *Principia* Russell partially defines and on occasion employs.[235] This extension differs from $LT_{av}(PM)$ in point of countenancing abstraction.[236]

(a) Vocabulary:

In addition to the vocabulary of $LT_{av}(PM)$, $LT_{av+}(PM)$ possesses abstracts as singular terms. The abstracts may be thought of as partitioned into an ω-sequence of classes that are defined inductively as follows:

Base Step:

> The class of *first-level* abstracts consists of all the expressions obtained by carrying out the operations (i) and (ii) below for every $LT_{av}(PM)$ formula ϕ and argument-variable sequence s appropriate to ϕ.
> (i) If s is empty, prefix ϕ with a single lambda operator λ and call the result a *propositional abstract*; otherwise, if s consists of $v_1, v_2, \ldots, v_i$, $i>0$, prefix ϕ with the sequence $\ulcorner\lambda v_1 \lambda v_2 \ldots \lambda v_i\urcorner$ and call the result a *propositional function abstract*.
> (ii) Assign the result of (i) the type of ϕ with respect to s.

Inductive Step:

> The class of *(n+1)st-level* abstracts consists of all the expressions obtained by the way that the first-level abstracts are obtained except for starting with those formulae that belong to the result of extending $LT_{av}(PM)$ by adding abstracts of every level below the (n+1)st level.

---

[235]See *Principia*, proposition *10.122, p. 140, as well as Chapters *20 and *21, especially p. 200.

[236]As such, the extension's conservativeness may be shown be methods similar to those employed to show that NB is a conservative extension of ZF.

Note that, according to the above description, a formula ϕ may give rise to two or more abstracts whose type and orders are different one from the other. Consider, for instance, $\ulcorner\lambda y^1.f^2x^1y^{1}\urcorner$ and $\ulcorner\lambda f^2\lambda y^1.f^2x^1y^{1}\urcorner$ whose superscripts indicate orders. The former abstract is of order 2 whereas the latter abstract is of order 3.

(b) Rules of Formation:

The rules of formation of LT_{av+}(PM) are a generalisation of those of LT_{av}(PM):

Atomic Formulae:

> Any propositional variable is an atomic formula. For i>0, if e_1 is an LT_{av+}(PM) expression of type t_1, e_2 an LT_{av+}(PM) expression of type t_2, ..., e_i an LT_{av+}(PM) expression of type t_i, and f an LT_{av+}(PM) expression whose pretype is $<t_1,t_2,\ldots,t_i>$, then $\ulcorner fe_1e_2\ldots e_i\urcorner$ is an atomic formula.

Compound Formulae:

> The compound formulae are obtained from the atomic formulae in the same way in which they are in LT_{av}(PM).

(c) Rules of Inference:

LT_{av+}(PM)'s rules of substitution are the result of extending LT_{av}(PM)'s rules of substitution in the obvious way so as to countenance the substitution of abstracts for variables of like type and the substitution of LT_{av+}(PM) formulae with respect to argument-variable sequences for variables of like type. Excluding its rules of substitution, LT_{av+}(PM)'s rules of inference are those of LT_{av}(PM).

(d) Axioms:

Every axiom of LT_{av}(PM) is an axiom of LT_{av+}(PM). In addition, LT_{av+}(PM) possesses *axioms* *of concretion* which fix the logical behaviour of the propositional function abstracts. Each of these axioms has the following form (where the usual conditions on a 'substituens' being free for a 'substituendum' obtain):

$$\lambda v_1 \ldots \lambda v_i . \phi(a_1,\ldots,a_i) . \leftrightarrow . \phi[v_1,\ldots,v_i/a_1,\ldots,a_i] \qquad (7)$$

Here, '$\lambda v_1 \ldots \lambda v_i . \phi(a_1,\ldots,a_i)$' represents an atomic formula whose predicate position contains a propositional function abstract having the form '$\lambda v_1 \ldots \lambda v_i . \phi$'.

Interestingly, the axioms of concretion function in LT_{av+}(PM), roughly speaking, as comprehension axioms and do so in at least two ways. First, from any of such axioms having the form (7), a comprehension formula having the following form may be derived by universal and existential generalisation:

$$\exists f(w_1)\ldots(w_i)\{f(w_1,\ldots,w_i) . \leftrightarrow . \phi[v_1,\ldots,v_i/w_1,\ldots,w_i]\} \quad (8)$$

Secondly, these axioms may actually be employed in derivations *in lieu of* any comprehension formulae whatsoever. To illustrate this, I present below two derivations of the statement that identity defined as $\lambda x \lambda y.(\phi)(\phi x \rightarrow \phi y)$ is symmetric. The first derivation employs a comprehension axiom in a representative way and the second derivation is like the first one except for employing an axiom of concretion in lieu of the first one's comprehension axiom.

Derivation 1: $x{=}y \rightarrow y{=}x$

{1}	(1)	$x{=}y$	Assumption
{1}	(2)	$(\phi)(\phi x \rightarrow \phi y)$	Definition, 1
{1,3}	(3)	ϕy	Assumption
{1,3,4}	(4)	$\neg\phi x$	Assumption
{1,3,4}	(5)	$\exists G(x)(Gx \leftrightarrow \neg\phi x)$	Comprehension Axiom
{1,3,4}	(6)	$(x)(Gx \leftrightarrow \neg\phi x)$	EI, 5
{1,3,4}	(7)	$Gx \leftrightarrow \neg\phi x$	UI, 6
{1,3,4}	(8)	$Gx \rightarrow Gy$	UI, 2
{1,3,4}	(9)	Gx	BE, 4, 7
{1,3,4}	(10)	Gy	MP, 8, 9
{1,3,4}	(11)	$Gy \leftrightarrow \neg\phi y$	UI, 6
{1,3,4}	(12)	$\neg\phi y$	MP, 10, 11 (no further use of 'G')

{1,3}	(13)	ϕx	RAA, 3, 12
{1}	(14)	$\phi y \rightarrow \phi x$	CP, 3, 13
{1}	(15)	$(\phi)(\phi y \rightarrow \phi x)$	UG, 14
{1}	(16)	$y=x$	Definition, 15
{}	(17)	$x=y \rightarrow y=x$	CP, 1, 16

Derivation 2: $x=y \rightarrow y=x$.[237]

{1}	(1)	$x=y$	Assumption
{1}	(2)	$(\phi)(\phi x \rightarrow \phi y)$	Definition, 1
{1,3}	(3)	ϕy	Assumption
{1,3,4}	(4)	$\neg\phi x$	Assumption
{1,3,4}	(5)	$[\lambda a.\neg\phi a](x) \leftrightarrow \neg\phi x$	Axiom of Concretion
{1,3,4}	(8)	$[\lambda a.\neg\phi a](x) \rightarrow [\lambda a.\neg\phi a](y)$	UI, 2, $[\lambda a.\neg\phi a/\phi]$
{1,3,4}	(9)	$[\lambda a.\neg\phi a](x)$	BE, 4, 5
{1,3,4}	(10)	$[\lambda a.\neg\phi a](y)$	MP, 8, 9
{1,3,4}	(11)	$[\lambda a.\neg\phi a](y) \leftrightarrow \neg\phi y$	Substitution, 5
{1,3,4}	(12)	$\neg\phi y$	MP, 10, 11
{1,3}	(13)	ϕx	RAA, 3, 12
{1}	(14)	$\phi y \rightarrow \phi x$	CP, 3, 13
{1}	(15)	$(\phi)(\phi y \rightarrow \phi x)$	UG, 14
{1}	(16)	$y=x$	Definition, 15
{}	(17)	$x=y \rightarrow y=x$	CP, 1, 16

(e) Russell's Abstracts:

Strictly speaking, my formulation of LT_{av+}(PM)'s abstracts differs somewhat from that of Russell's. According to his formulation, roughly speaking, an abstract is obtained by taking a formula ϕ whose free variables occur in alphabetical order, circumflexing such variables, and then interchanging them in ϕ so that they occur in the order in which they are to be replaced under the rules of substitution. In this regard, Russell writes in Chapter *21:

> For substitution in $\phi!(\hat{x},\hat{y})$ and $\phi!(\hat{y},\hat{x})$, we adopt the convention that when a function (as opposed to its values) is represented in a form involving $\hat{x}$ and $\hat{y}$, or any other two letters of the alphabet, the value of this function for the arguments *a* and *b* is to be found by substituting *a* for $\hat{x}$ and *b* for $\hat{y}$, while the value for the arguments *b* and *a* is to be found by substituting

[237]Note that Derivation 2's line numbers correspond to those of Derivation 1.

> *b* for $\hat{x}$ and *a* for $\hat{y}$. That is, the argument mentioned first is to be substituted for the letter which comes first in the alphabet, and the argument mentioned second for the later letter; thus the mode of substitution depends upon the *alphabetical* order of the letters which have circumflexes and the *typographical* order of the other letters.[238]

Immediately afterward, Russell offers the following axioms of concretion to illustrate the substitution conventions in question:[239]

$$a\{\phi!(\hat{x},\hat{y})\}b\,.{=}.\,\phi!(a,b)$$
$$b\{\phi!(\hat{x},\hat{y})\}a\,.{=}.\,\phi!(b,a)$$
$$a\{\phi!(\hat{y},\hat{x})\}b\,.{=}.\,\phi!(b,a)$$
$$b\{\phi!(\hat{y},\hat{x})\}a\,.{=}.\,\phi!(a,b)$$

There are of course a few problems with Russell's formulation. First, it says nothing about how to obtain abstracts containing free parameter variables. Moreover, any straightforward extension of the formulation that would say how to do so would likely give rise to ambiguities of scope. For instance, any such extension would likely represent both:

$$\lambda x^1[\pi^3(\lambda y^1.x^1L^2y^1)]$$

and:

$$\lambda y^1[\pi^3(\lambda x^1.x^1L^2y^1)]$$

by:

$$\pi^3(\hat{x}^1L^2\hat{y}^1)$$

Secondly, what it is for a formula ϕ to have its free variables occur in alphabetical order must spelled out in order to speak to the situation in which ϕ contains several occurrences of several free variables. Thirdly, the very requirement that such a formula ϕ's free variables occur in alphabetical order appears to conflict with the point that variables in general "serve merely to indicate cross-references to various positions of quantification"[240] as well as to create undue

---

[238]*Principia*, p. 200.

[239]*ibid.*, *21.02. Curiously, Russell does not regard the axioms of concretion as axioms proper. Rather, he regards them as *definitions* concerning propositional function abstracts.

[240]Quine, *Mathematical Logic*, p. 70.

restriction on the mechanics of derivation.

I should note that none of these problems caused Russell trouble because in *Principia* he contextually defined class and relation abstracts which function in the way in which LT_{av+}(PM)'s abstracts as conceived of by my formulation do. Russell employed the former abstracts in most of *Principia*'s technical work.

8. Fined-Grained Types:

At the beginning of this chapter I said that, apart from Russell's official definition of *being of the same type*, LT_{av}(PM) may be taken as having any of an entire spectrum of notions of type ranging from the most coarse-grained to the most fine-grained. So far, I have considered LT_{av}(PM) as having a rather coarse-grained notion. In this section I consider it as having a rather fined-grained notion, one that fits Russell's official definition. For convenience, call the former logical theory $LT_{avc}(PM)$ and the latter logical theory $LT_{avf}(PM)$.

(a) Russell's Official Definition of *Being of the Same Type:*

Russell states his definition of the relation λxλy(x is of the same type as y) in *9.131:

> *9.131. *Definition of "being of the same type."* The following is a step-by-step definition, the definition for higher types presupposing that for lower types. We say that *u* and *v* "are of the same type" if (1) both are individuals, (2) both are elementary functions taking arguments of the same type, (3) *u* is function and *v* is its negation, (4) *u* is ϕ^*x* or ψ^*x*, and *v* is ϕ^*x* $\vee$ ψ^*x*, where ϕ^*x* an ψ^*x* are elementary functions, (5) *u* is (*y*). ϕ(^*x*,*y*) and *v* is (*z*). ψ(^*x*,*z*), where ϕ(^*x*,^*y*), ψ(^*x*,^*y*) are of the same type, (6) both are elementary propositions, (7) *u* is a proposition and *v* is ¬*u*, or (8) *u* is (*x*). ϕx and *v* is (*y*). ψy, where ϕ^*x* and ψ^*x* are of the same type.[241]

Several considerations are relevant here. First, in contrast to the level at which I have been speaking, the definition speaks at the object level. As such, the entities to which the relation

---

[241]*Principia*, p. 133.

that it defines applies are not elements of LT_{av}(PM)'s vocabulary but, rather, terms in LT_{av}(PM)'s ontology. In particular, they are LT_{av}(PM)'s individuals, matrices, propositional functions, and propositions. Secondly, although it is obvious that the definition is inductive, it is perhaps less obvious that the induction in question moves over the relation of *being an immediate part of*, where x is an immediate part of y if and only if (i) y is the negation of x, (ii) for some z, y is the disjunction of x and z, or (iii) for some u, y is the universal quantification of x with respect to u. With regard to LT_{av}(PM)'s ontology, this relation is well-founded[242] and, as such, the induction in question is legitimate. There are two points to note in this connection. One is that, in light of the fact that this induction does move over such a relation, Russell's wording in *9.131's first sentence is, to a certain extent, misleading. In writing "the definition for higher types presupposing that for lower types", Russell meant -- contrary to what one might have expected -- something like 'the definition for terms having a certain degree of complexity presupposing that for terms having a smaller degree of complexity'. The other point is that, in virtue of the nature of inductive definitions in general, the base clause of this definition must suppose that the relation that it defines, $\lambda x \lambda y$(x is of the same type as y), be already defined for the LT_{av}(PM) terms that are minimal with respect to the relation that the induction in question moves over -- that is, the individuals and matrices. Conditions (1) and (2) of *9.131, which approximate the base clause, make such a supposition clear.

Thirdly, Russell's statement of the definition in *9.131 is incomplete. He omits mention of the case involving the disjunction of non-elementary terms. One should note that, here, by

---

[242]Recall that every LT_{av}(PM) term is either a simple or a complex ultimately built up from simples by well-founded compositional operations -- as such, LT_{av}(PM) respects VCP1. In particular, LT_{av}(PM)'s individuals and matrices are simples and its propositional functions and propositions are complexes ultimately built up from individuals, matrices, and variables by the well-founded
operations of predication, negation, disjunction, and universal quantification.

"elementary", Russell means *simple*, not *first-order*. In this respect, an elementary propositional function is, roughly speaking, a matrix[243] and an elementary proposition is an atomic proposition as explained in Chapter 2.

In the light of these three considerations, one might spell out Russell's statement of the definition in order to see which notions of type fit the relation that it defines. Here is a rough spelling-out: x is of the same type as y if and only if the simples from which x is ultimately built are of the same type as those from which y is ultimately built and, for any parts $x_1, x_2, \ldots, x_i$ of x that combine in some way to form another part x_{i+1} of x, there are parts $y_1, y_2, \ldots, y_i$ of y that combine in the same way to form another part y_{i+1} of y and vice versa. More precisely: Assign the terms of $LT_{av}(PM)$'s ontology some reasonable degree of complexity. For any such term u, let t(u) be the type of u, let P(u) be the class of all the parts of u, and let $P_i(u)$ be the class of all the parts of u having degree of complexity less than i. Then, t(x) = t(y) if and only if there exists a bijection f from P(x) onto P(y) such that, where the degree of complexity of x is l, (1) for all i, $0 \leq i \leq l$, f bijectively maps $P_i(x)$ onto $P_i(y)$, (2) for all $u \in P_0(x)$, t(u) = t(f(u)), and (3) for all j, $0 \leq j < l$, if $v \in P_{j+1}(x)$ and $v_1, \ldots, v_k$ combine in some way -- that is, by predication, negation, disjunction, or universal quantification -- to form v, then $f(v_1), \ldots, f(v_k)$ combine in the *same* way to form f(v).

There are two points to note about this spelling-out of Russell's statement *9.131. First, it makes up for the incompleteness that the statement suffers and does so by requiring of any two compound terms u and v of the same type that, roughly speaking, the immediate parts of u be of the same type as those of v and that these parts of u combine in the same way as

[243] To the extent that Russell construes a propositional function in general as the result of taking a proposition and replacing one or more of its parts with one or more variables, an elementary propositional function is, strictly speaking, the result of taking an *atomic* proposition and .... .

those of v. Obviously, there are less stringent ways to make up for the incompleteness. Secondly, the above spelling-out is somewhat unfaithful to Russell's statement in the sense that whereas, according to the latter, u and ¬u are of the same type as are u and u∨v if u and v are, according to the former, neither u and ¬u nor u and u∨v are of the same type. The above spelling-out, however, is simpler than one faithful to Russell's statement would be and serves the expository purposes of this section equally well.

It should be clear that the notion of type defined in §2 is far too coarse-grained to fit the relation λxλy(x is of the same type as y) as defined by the above spelling-out. It should also be clear that any of several rather fine-grained notions do fit it. In the next subsection I describe one such notion and LT_{av}(PM) as having it -- that is, LT_{avf}(PM).

(b) LT_{avf}(PM):

LT_{avf}(PM) will be described in the way in which LT_{avc}(PM) was described above.

(1) LT_{avf}(PM)'s Vocabulary:

LT_{avf}(PM)'s Orders:

Orders are natural numbers including zero.

LT_{avf}(PM)'s Types:

Types are ordered triples of *s-formula/matrix pretypes*, *argument pretypes*, and orders. These are defined recursively as follows:

Base Step:

The empty sequence <> is an s-formula pretype as well as an argument pretype. The ordered triple <<>,<>,0> is a type.

Inductive Step:

Matrix Pretype and Matrix Type

If t_1, t_2, ..., t_i are types, i>0, then the sequence M of these types is a *matrix pretype*. If O is the least number greater than the orders of t_1, t_2, ..., t_i, then the ordered triple <M,M,O> is a *matrix type* -- call it the *corresponding matrix type* of the matrix pretype M.

S-Formula Pretype

Let C_1 be the function $\lambda x.<0,x>$; let C_2 be the function $\lambda x\lambda y.<1,x,y>$; and let C_3 be the function $\lambda x\lambda y.<2,x,y>$. C_1, C_2, and C_3 are required in order to define a notion of type that fits the above spelling-out of *9.131. Heuristically speaking, given a certain constituent feature of the type of an LT_{av}(PM) term u, C_1 yields the analogous feature of the type of u's negation; given certain constituent features of the types of terms u and v, C_2 yields the analogous feature of the type of their disjunction; and given certain constituent features of the types of terms u and v, C_3 yields the analogous feature of the type of the universal quantification (u).v.

An item S is an s-formula pretype if and only if there is a finite sequence σ -- call it *S's witnessing sequence* -- such that: (1) for any member n of σ, (i) n is a matrix type, or (ii) for some i occurring earlier in σ, n = C_1(i), or (iii) for some i, j occurring earlier in σ, n = C_2(i,j), or (iv) for some i occurring earlier in σ and for some e that is either a matrix type occurring earlier in σ or a type that is a member of a matrix pretype whose corresponding matrix type occurs earlier in σ, n = C_3(e,i) (any such e is said to be *bound in* S); (2) S is not a matrix type; (3) S is the last member of σ; and (4) σ is the smallest possible witnessing sequence of S (there are no extraneous matrix types occurring in it). Heuristically speaking, an s-formula pretype represents some constituent feature of the type of a term u where u is specified by means of an s-formula and the feature identifies the s-formula's compositional structure.

Argument Pretype

An item A is an argument pretype for a matrix pretype M if and only if it is a sequence, empty or non-empty, of any of the members of M in any order without repetition. Call the members of M not in A *parameters of M with respect to A*. An item A is an argument pretype for an s-formula pretype S if and only if it is a non-repeating sequence, empty or non-empty, such that for every member n of A, either (1) n is a matrix type occurring in S's witnessing sequence σ, or (2) n is a type which is a member of a matrix pretype whose corresponding matrix type occurs in σ and n is not bound in S. Call the non-members of A that satisfy (1) and (2) *parameters of S with respect to A*. An item A is an argument pretype if and only if it is an argument pretype for some matrix pretype or an argument pretype for some s-formula pretype. Heuristically speaking, an argument pretype represents some constituent feature of the type of an $LT_{av}(PM)$ term u that identifies which of u's free variables -- *qua* terms in $LT_{av}(PM)$'s ontology -- are argument variables, as opposed to parameter variables, as well as the order in which they occur.

Order

The *order of a matrix pretype M with respect to an argument pretype A* (where A is an argument pretype for M) is the least number greater than the orders of the members of A and greater than or equal to the orders of the parameters of M with respect to A. The *order of an s-formula pretype S with respect to an argument pretype A* (where A is an argument pretype for S) is the least number greater than the orders of the members of A, greater than the orders of the types that are bound in S, and greater than or equal to the orders of the parameters of S with respect to A. Note that here, in contrast to the situation with the coarse-grained notion of type defined in §2 where several orders are appropriate to any particular pretype, only one order is appropriate to a matrix/s-formula pretype with respect to an argument pretype.

Type

An item T is a type if and only if it is an ordered triple $\langle S, A, O\rangle$ such that S is either a matrix pretype or an s-formula pretype, A is an argument pretype for S, and O is the order of S with respect to A. As such, a matrix type as defined above is a type. Heuristically speaking, a type, thus understood, of an $LT_{av}(PM)$ term u identifies u's compositional structure as well as its argument-variable sequence.

It should be clear that the notion of type just defined fits the relation $\lambda x \lambda y$(x is of the same type as y) as defined by the above spelling-out of *9.131. One may prove that it does by a straightforward, if not tedious, induction.

Variables:

For every type t, $LT_{avf}(PM)$ has an infinite alphabet of variables. Concerning the semantics of these variables, if a type $t = \langle\langle\rangle, \langle\rangle, 0\rangle$, the variables of type t are taken to range over the individuals and, as such, are called *individual variables*. If $t_1, \ldots, t_i$ are types, $i>0$, $M=\langle t_1,\ldots,t_i\rangle$ is a matrix pretype and, for some appropriate order O, $t=\langle M,M,O\rangle$ is a matrix type. The variables of type t are taken to range over atomic (elementary) propositional functions[244] of that type where the arguments of such functions are in the relevant sense of types $t_1, \ldots, t_i$. These variables correspond to the matrices about which Russell writes in Chapter *12 of *Principia* and, as such, are called *matrix variables*. They are also called *propositional function*

[244]See Chapter 2, §2 (a.ii).

variables. If $A=\langle t_{a1},\ldots,t_{aj}\rangle$, $a_j\leq i$, is a non-empty argument pretype for M, for some appropriate order O, $t=\langle M,A,O\rangle$ is a type. The variables of type t are taken to range over propositional functions of that type where the arguments of such functions are of types $t_{a1}, \ldots, t_{aj}$. As such, the variables of type t are called *propositional function variables*. If $A=\langle\rangle$ is the empty argument pretype for M, for some appropriate order O, $t=\langle M,\langle\rangle,O\rangle$ is a type. The variables of type t are taken to range over atomic (elementary) propositions and, as such, are called (*atomic*) *propositional variables*. If S is an s-formula pretype and $A=\langle t_1,\ldots,t_i\rangle$ a non-empty argument pretype for S, for some appropriate order O, $t=\langle S,A,O\rangle$ is a type. The variables of type t are taken to range over compound propositional functions of that type where the arguments of such functions are of types $t_1, \ldots, t_i$. As such, the variables of type t are called (*compound*) *propositional function variables*. If $A=\langle\rangle$ is the empty argument pretype for S, for some appropriate order O, $t=\langle S,\langle\rangle,O\rangle$ is a type. The variables of type t are taken to range over compound propositions of that type and, as such, are called (*compound*) *propositional variables*.

LT_{avf}(PM)'s predicative variables are all and only the matrix variables. These variables correspond to the predicative functions about which Russell writes in Chapter *12 in *Principia*.[245]

Logical Constants and Grouping Devices:

LT_{avf}(PM) possesses the same logical constants and grouping devices as LT_{avc}(PM).

---

[245]Note that others of LT_{avf}(PM)'s variables could have been chosen to play the role of predicative variable. In particular, for a given argument pretype A, any variables of type $\langle S,A,O\rangle$ whose order O is the least possible with respect to A could have been chosen.

(2) LT$_{avf}$(PM)'s Rules of Formation:

LT$_{avf}$(PM)'s rules of formation are similar to those of LT$_{avc}$(PM):

Atomic Formulae:

Any propositional variable is an atomic formula. For $i>0$, if v_1 is a variable of type t_1, v_2 a variable of type t_2, ..., v_i a variable of type t_i, and f a propositional function variable whose argument pretype is $<t_1,t_2,...,t_i>$, then $\ulcorner fv_1v_2...v_i\urcorner$ is an atomic formula.

Compound Formulae:

The compound formulae are obtained from the atomic formulae in the same way in which they are in LT$_{avc}$(PM).

(3) LT$_{avf}$(PM)'s Axioms and Rules of Inference:

Except for the necessary changes, the axioms and the rules of inference of LT$_{avf}$(PM) are those of LT$_{avc}$(PM).

It is noteworthy that there are natural injective mappings from the class of LT$_{avc}$(PM)'s types into the class of LT$_{avf}$(PM)'s types and, correlatively, from the class of LT$_{avc}$(PM)'s vocabulary and formulae into the class of LT$_{avf}$(PM)'s vocabulary and formulae. Moreover, if f is such a mapping, then for any LT$_{avc}$(PM) proof P that makes no appeal to LT$_{avc}$(PM)'s second rule of substitution for propositional function variables nor to its second rule of substitution for propositional variables, there is an LT$_{avf}$(PM) proof P' that likewise makes no such appeal such that P is isomorphic to P' via f. By contrast, there are LT$_{avc}$(PM) proofs Q that do make appeal of the kind in question such that, for any $LT_{avf}(PM)$ proof Q', Q is not isomorphic to Q' via f. This is not to say that, if ϕ is the $LT_{avc}(PM)$ formula that Q proves, $f(\phi)$ is beyond proof in $LT_{avf}(PM)$.

References

Boolos, G. "Whence the Contradiction?" in *Aristotelian Society*, Suppl. vol. 67 (1993) pp. 213-234.

Carnap, R. *The Logical Syntax of Language* (London: Harcourt Brace, 1937).

---- *Der Logische Aufbau der Welt* (1928) translated as *The Logical Structure of the World*, by Rolf A. George (Berkeley: University of California Press, 1967).

Cartwright, R. "Speaking of Everything", 1989 ASL lecture.

Chihara, C. *Ontology and the Vicious-Circle Principle*, (Ithaca: Cor. University Press, 1973).

Church, A. "Comparison of Russell's Resolution of the Semantical Antinomies with that of Tarski" in *Journal of Symbolic Logic* 41 (1976) pp. 747-60.

---- *Introduction to Mathematical Logic*, vol. 1 (Princeton: Princeton University Press, 1956).

Copi, I. *The Theory of Logical Types* (London: RKP, 1971).

Davidson, D. "Radical Interpretation" in *Truth and Interpretation*, (Oxford: Clarendon, 1985).

Dummett, M. *Frege: Philosophy of Language*, 2nd edn. (Cambridge, Mass.: HUP, 1981).

Frege, G. *Begriffsschrift* (1879) in van Heijenoort.

---- *Collected Papers on Mathematics, Logic and Philosophy*, B. McGuinness, ed. (Oxford: Basil Blackwell, 1984).

---- *Grundgesetze der Arithmetik* (Jena: Pohle, vol. 1, 1893; vol. 2, 1903).

---- "Über die Begriffsschrift des Herrn Peano und meine eigene," in *Berichte über die Verhandlungen der Königlich Sachsischen Gesellschaften der Wissenschaften zu Leipzig*, Mathematischephysische Classe, 48 (1897). Translated as "On Mr. Peano's Conceptual Notation and My Own," in *Collected Papers*, pp. 234-248.

Gentzen, Gerhard "Investigations into Logical Deduction" in *The Collected Papers of Gerhard Gentzen*, ed. M.E. Szabo (Amsterdam: North-Holland Publishing Company, 1969).

Gödel, K. "On Formally Undecidable Propositions of *Principia Mathematica* and Related Systems I" (1931) in van Heijenoort, pp. 596-616.

Goldfarb, W. "Logic in the Twenties: The Nature of the Quantifier", in *Journal of Symbolic Logic* 44 (1979) pp. 351-368.

---- "Russell's Reasons for Ramification" (RRR) in Savage and Anderson, pp. 24-40.

Hallett, M. *Cantorian Set Theory and the Limitation of Size* (Oxford: Clarendon Press, 1984).

van Heijenoort, J. *From Frege to Gödel* (Cambridge, Mass.: HUP, 1967).

---- "Logic as Calculus and Logic as Language" in *Synthese* 17 (1967), pp. 324-30.

Hilbert, D., and W. Ackermann *Grundzüge der Theoretischen Logik* (Berlin: Springer, 1928; 2d ed., 1938; 3d ed., 1949).

---- and P. Bernays *Grundlagen der Mathematik* (Berlin: Springer, vol. 1, 1934; vol. 2, 1939).

Hylton, P. *Russell, Idealism, and the Emergence of Analytic Philosophy* (RIAP), (Oxford: OUP, 1990).

Kleene, S. *Introduction to Metamathematics* (Amsterdam: North-Holland, 1952).

Kneale, W. and Kneale, M. *The Development of Logic* (Oxford: Clarendon, 1984).

von Neumann, J. *Zur Einführung der transfiniten Zahlen*. Acta Sci. Math. (Szeged) *1*, pp. 199-208.

Parsons, C. "Sets and Classes" in *Mathematics in Philosophy*, (Ithaca, New York: Cornell University Press, 1983).

Poincaré, H. "Les mathématiques et la logique", *Revue de Métaphysique et de Morale*, 1906, 14.

Ramsey, "Foundations of Mathematics" (FM), in *Foundations*, (Atlantic Highlands, NJ: Humanities Press,1978).

Ricketts, T. "Frege, the *Tractatus* and the Logocentric Predicament", *Nous* 19 (1985) pp. 3-16.

Russell, B. *Essays in Analysis* (EA) D. Lackey, ed. (New York: George Braziller, 1973).

---- "Mathematical Logic as Based on the Theory of Types" (1908) (ML) in van Heijenoort.

---- "Mathematics and the Metaphysicians" (MM) in *Mysticism and Logic*.

---- *My Mental Development* in Schilpp.

---- *Mysticism and Logic* (Totowa, NJ: Barnes and Noble, 1976).

---- *My Philosophical Development* (MPD) (London: Unwin Books, 1975).

---- "On 'Insolubilia' and their Solution in Symbolic Logic" ("On 'Insolubilia'") in EA, pp. 190-214.

---- "On Some Difficulties in the Theory of Transfinite Numbers and Order Types" ("On Some Difficulties") in EA, pp. 135-164.

---- *The Philosophy of Logical Atomism* (LA), (La Salle, Illinois: Open Court, 1985).

---- *The Principles of Mathematics* (*Principles*) (London: CUP, 1903; second edition: New York: Norton, 1938).

---- "The Relation of Sense-Data to Physics" in *Mysticism and Logic,* pp. 108-131.

Quine, W.V. *Elementary Logic*, (Cambridge Massachusetts: HUP, 1980).

---- "Epistemology Naturalized" in *Ontological Relativity.*

---- "Existence and Quantification" in *Ontological Relativity.*

---- *From a Logical Point of View* (Cambridge, Mass.: HUP, 1980).

---- *Mathematical Logic* (Cambridge, Mass.: HUP, Rev. ed., 1951).

---- *Methods of Logic* (Cambridge, Mass.: HUP, 4th ed., 1982).

---- *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1969).

---- "Ontological Relativity" in *Ontological Relativity.*

---- "On What There Is" in *From a Logical Point of View.*

---- *Set Theory and its Logic* (Cambridge, Mass.: HUP, 1969).

---- *A System of Logistic* (Cambridge, Mass.: 1934).

---- "Two Dogmas of Empiricism" in *From a Logical Point of View.*

---- *Word and Object* (Cambridge, Mass.: MIT, 1960).

Savage, C. and Anderson, C. eds. *Rereading Russell* (Minneapolis: University of Minnesota Press, 1989).

Schilpp, P. ed. *The Philosophy of Bertrand Russell* (La Salle, Illinois: Open Court, 1989).

Tarski, A. "On the Concept of Logical Consequence" (1936) in *Logic, Semantics, and*

Metamathematics, (Indianapolis: Hackett, 1983).

---- "The Concept of Truth in Formalized Languages" ("Wahrheitsbegriff") (1931) in *Logic, Semantics, and Metamathematics*, (Indianapolis: Hackett, 1983).

Whitehead, A.N. and Russell, B. *Principia Mathematica* (*Principia*), vols. 1-3 (Cambridge: CUP, 1910-12; 2nd edn. 1927)