Paradoxes of Intensionality
Version of: March 1, 2007

Dustin Tucker and Richmond H. Thomason
Philosophy Department, University of Michigan

Abstract

We identify a class of paradoxes that are neither set-theoretical or semantical, but that seem to depend on intensionality. In particular, these paradoxes arise out of plausible properties of propositional attitudes and their objects. We try to explain why logicians have neglected these paradoxes, and to show that, like the Russell Paradox and the direct discourse Liar Paradox, these intensional paradoxes are recalcitrant and challenge logical analysis. Indeed, when we take these paradoxes seriously, we may need to rethink the commonly accepted methods for dealing with the logical paradoxes.

1. Introduction

Contemporary work on the paradoxes of logic and set theory is framed by ideas that go back to the 1920s and 1930s. In 1925, Frank Ramsey divided the paradoxes into those (like the Russell Paradox) that arise within mathematics, and those (like the Liar Paradox) that do not.

Having made his division of the paradoxes, Ramsey took the Liar and related paradoxes to be extralogical, involving an "empirical" linguistic element. But later, in work published in 1936, Alfred Tarski showed that the Liar Paradox does arise even in rigorously presented semantic theories of formalized languages, making a compelling case that the enterprise of producing these theories belongs to logic. At the same time, he demonstrated how the paradox could be avoided by relativizing truth to a language and invoking a linguistic hierarchy in which no language could serve as its own semantic metalanguage.

These ideas were congenial to the spirit of the times. Ramsey's strategy, as we will see, depended on banishing intensionality from logic, an idea which fits well with the behaviorism and anti-mentalism that lasted well into the second half of the century.[1] And Tarski's construal of the Liar as metalinguistic follows the linguistic turn, transforming the traditional 'what I am now saying is false' into 'this sentence is false'. This "linguistic turn" was not only becoming popular at the time, but remains as part of the repertoire of contemporary philosophy.

[1] It is not clear how adamant Ramsey himself would have been about banishing intensionality from logic. Although much of [Ramsey, 1925] is concerned with eliminating intensionality from the logicist project of *Principia*, he refers to propositions throughout the work. And it is clear that he conceives of propositions intensionally, as the senses of sentences; for instance, in discussing the values of propositional functions, he proposes [p. 37] to "determine them by a description of their senses or imports." Later, he spends several pages [pp. 42–49] addressing the paradoxes he classified earlier as "not purely logical," making use of a meaning relation R that holds between symbols and functions (and symbols and propositions). Ramsey often speaks as if R was a relation between symbols and symbols; this may be due in part to confusion of use and mention.

Together, Ramsey and Tarski suggest an attractive picture of the general problem posed by the family of paradoxes resembling the Russell Paradox and the Liar. In the terminology that became current after 1936, the *set-theoretical* paradoxes belong to the foundations of mathematics, and are the proper concern of set theory and related areas of mathematics. On the other hand, the *semantic paradoxes* belong to the foundations of semantics. Along with the distinction goes a division of labor: most mathematicians working with semantic theories (and, in particular, most model-theorists) can afford to ignore the semantic paradoxes, but they remain a problem for a group of philosophers and mathematicians concerned with the foundations of semantics.

This picture, and the division of labor that goes along with it, had become well accepted by the 1960s,[2] and is still presupposed in contemporary work on the paradoxes. Since the publication of [Kripke, 1975] the Liar Paradox has received a great deal of attention; we know of ten books published after 1980 that deal with this topic. And almost all of this work on the Liar Paradox takes the metalinguistic formalization of the paradox for granted.[3]

The purpose of this paper is to question this cluster of assumptions. We believe that Ramsey's distinction is not exhaustive—for instance, there are versions of the Liar Paradox that do not fit into either of his categories. And indeed, reflection on these examples suggests that Ramsey's categories, and the division of labor that goes along with them, may be misguided. Although, of course, in the last sixty years we have learned much about specific formalizations of certain paradoxes, the general foundational problem presented by the paradoxes is rather neglected, and calls for radical reassessment. In fact, we may have to reset the clock back to 1900, and to rethink the entire problem in the light of what has been learned since.

2. The background to Ramsey's distinction

2.1. The Russell paradox

In 1901, Russell, prompted by Cantor's proof that there is no greatest cardinal number, discovered the paradox of the set of all sets that are not members of themselves, which came to be known as the *Russell Paradox*. Burali-Forti's paradox of the greatest ordinal number [Burali-Forti, 1897] had been discovered four years earlier, but Russell later stated [Russell, 1959][pp. 77–78] that, due to its complexity, he had allowed himself to take it rather lightly.

The Russell Paradox is remarkably simple. Moreover, it involves only logical operators and set membership. It therefore presented with devastating force a foundational problem concerning the nature of sets. And if set membership is formalized in terms of predication, as Russell formulated it, the paradox raises a foundational problem for logic itself.

The foundational importance of the paradox was immediately apparent: Russell reports [Russell, 1959][p. 76] that when he disclosed it to Frege, the latter "replied that arithmetic was tottering . . . [and] gave up the attempt to deduce arithmetic from logic." Russell himself "settled down to a resolute attempt to find a solution of the paradoxes" and "felt [it] as almost

[2]For documentation of this point, see [Fraenkel and Bar-Hillel, 1958][pp.5–14], [Beth, 1959][§171], [Quine, 1963][pp. 254-255], [Kneale and Kneale, 1962][pp. 664-665].

[3][Barwise and Etchemendy, 1987] is a notable exception to this generalization.

a personal challenge [which he] would, if necessary, have spent the whole of the rest of [his] life in an attempt to meet." [Russell, 1959][p. 79]

2.2. The vicious circle principle and the ramified theory of types

In Appendix B of [Russell, 1903], Russell proposed a sketch of "the doctrine of types," showing how it can provide a solution to the Russell Paradox. Russell later (in the introduction to [Russell, 1937]) characterized Appendix B as a "rough sketch" and a "crude form" of the theory of types. Five years later, Russell had completed a version of the theory of types that seemed satisfactory to him; this formulation, published in [Russell, 1908], is essentially the same as that of the first and second editions of *Principia Mathematica* [Whitehead and Russell, 1910–1913, Whitehead and Russell, 1925–1927].

A definitive version of Russell's views on the paradoxes prior to the publication of [Ramsey, 1925] can be found in Chapter II of *Principia*, entitled "The theory of logical types." The chapter begins with the comprehensive statement that the relevant paradoxes[4] "all result from a certain kind of vicious circle," and that these circles in turn "arise from supposing that a collection of objects may contain members which can only be defined by means of the collection as a whole." [Whitehead and Russell, 1910–1913][p. 39]. The *Vicious Circle Principle* is then stated as a rejection of "illegitimate totalities". It is formulated in two ways, which we will call the *direct* and the *converse* forms.

> (**VCP**) [The direct form:]"Whatever involves *all* of a collection must not be one of this collection"; or, conversely, [The converse form]: "If, provided that a certain collection had a total, it would have members only definable in terms of that total, then the said collection has no total." [Whitehead and Russell, 1910–1913][p. 40].

The principle is evidently meant to be understood in such a way that violations of it ("vicious circle fallacies") result in thoughts or uses of language that are meaningless. Thus, for instance, any statement about "all propositions," with 'all' understood in an unqualified way, could not express a proposition and so would have to be meaningless.

The vicious circle principle is problematic in many ways. (1) It is stated in very informal language and without much clarification. As Gödel pointed out many years later [Gödel, 1944], it can be interpreted in ways that are significantly different. (2) Its application even to the claim that the set of all sets belongs to itself is questionable. Does the collection of all sets involve all of the collection of all sets? This depends on what 'involve' means. If the converse form of **VCP** is indeed the converse of the direct form, then we have to ask whether the set of all sets is definable only in terms of the set of all sets. But it is not clear that it is only so definable, since we can use '$\{x \ / \ x = x\}$' to define the set of all sets.[5] (3)

[4]Seven such paradoxes are listed and analyzed later in in this chapter. They include: (1) The Liar Paradox, (2) The Russell Paradox, (3) a version of the Russell Paradox involving relations rather than sets, (4) the Burali-Forti paradox of the ordinal number of all ordinals, (5–7) the Richard Paradox, which concerns definable real numbers, and two variations on this paradox.

[5]The vicious circle principle, of course, is intended to render '$\{x \ / \ x = x\} \in \{x \ / \ x = x\}$' meaningless, but such a strong restriction is certainly not necessary to avoid paradoxes, and its appropriateness is debatable. Work in nonwellfounded set theory shows that systematic violations of **VCP** create no inconsistency (relative

The **VCP** is self-defeating. If we apply it comprehensively and ruthlessly as a criterion of meaninglessness, we will have to eliminate the **VCP** itself. (4) *Principia* is notoriously lax regarding use-mention distinctions. The formulation of **VCP** depends on this laxity, and seems to get much of its generality from blurring the distinction.

On the other hand, the **VCP** is admirably ambitious. It represents an attempt to survey the logical paradoxes and to find a single, comprehensive solution of them all, including an explanation of the fallacies on which they depend. And it motivates a formalized (or at least formalizable) logical theory in which a wide selection of paradoxes (including the seven examples listed at the end of [Whitehead and Russell, 1910–1913][Chapter II]) can be reproduced and exhibited as formally incorrect.

Except for approaches that embrace inconsistency,[6] it is hard to find contemporary work on the paradoxes that seeks to achieve this sort of generality.

2.3. Ramified type theory and the axiom of reducibility

The presentation of type theory in [Whitehead and Russell, 1910–1913] is notoriously careless regarding use and mention. The system of *Principia* can be seen equally as an account of a formal language based on typed variables, or as the characterization of the domains of entities (propositions and propositional functions) in terms of which the language is to be interpreted. Here, we will present it in the former way, as the definition of a language.

The very loose account of **VCP** becomes much clearer as soon as it is used to motivate the ramified theory of types. Crucially, the **VCP** emerges as a restriction on variables: that is, on the only mechanism available in the language to express general statements. It therefore automatically becomes a comprehensive constraint on the formulation of any generalization.

But in actually generating the types, the **VCP** plays out in two quite different ways, corresponding to the direct and the converse forms of the principle. The first of these ways amounts to what has become the standard method of avoiding the Russell Paradox in set theory; see, for instance, [Fraenkel *et al.*, 1973, Giaquinto, 2002]. To implement the direct form of **VCP**, you assign *levels*[7] to predicates (expressions, including variables, intended to denote classes) so that a predicate of level τ cannot be predicated of a term of level τ. The simplest way to do this is to treat such expressions as ill-formed. Thus, whatever level P receives, $\lambda P \neg P(P)$ will fail to be well-formed.

But, since this involves no typing of expressions of propositional type, it does not solve, for instance, the Liar Paradox, or the paradox of the cardinality of the set of all propositions. To avoid these problems, you assign *orders* to propositional variables and expressions in such a way that the orders of the expressions are strictly greater than the orders of all the bound variables occurring in them. The inferential rules for quantifiers then prevent instantiating a bound variable of order τ with an expression that receives an order greater than τ.[8] Thus,

to standard set theory), and makes a good case for their utility. See [Aczel, 1983, Barwise and Etchemendy, 1987].

[6]See, for instance, [Priest, 2005].

[7]In the following, we will use 'level' and 'order' in a technical sense. 'Type' will be used generally, for any hierarchical classification of expressions, until Section 5, where we will use it to refer to the types of a version of Simple Type Theory with a primitive type of propositions.

[8]The clearest formulation of the intensional ramified theory that we know of is presented in [Church, 1976].

whatever order a propositional variable p receives, $\lambda p\, \phi$ cannot be instantiated with a formula that contains a variable of this order.

These constraints interact, of course, since any predicate must receive a type of some sort, and yet when it is applied to appropriate terms it will form an expression that has propositional type (and which therefore must receive an order). Working out these interactions in a formal system leads to a version of the ramified theory of types, such as the system of *Principia* and more carefully formalized later versions, as in [Copi, 1971, Church, 1976].[9]

The ramified theory of types of [Whitehead and Russell, 1910–1913, Whitehead and Russell, 1925–1927] turned out to be notoriously weak as an instrument for carrying out the logicist program that inspired the project. In particular, the real numbers could not be located in a single type, but instead become spread out along a series of different orders,[10] and as a result Russell and Whitehead found themselves unable to prove fundamental results of analysis like the Least Upper Bound Theorem. To deal with these problems, *Principia* invokes the *Axiom of Reducibility*, which says that any functional expression of order τ is extensionally equivalent (and so "reducible") to a functional expression of order σ, where σ is any order less than τ.

From the start, the Axiom of Reducibility was always seen as an embarrassment for the program of formalizing analysis as a definitional extension of logic. The axioms of a logicist-inspired system like that of *Principia* must be plausible as axioms of logic (rather than as axioms of mathematics or any other domain of inquiry), but it is hard to motivate Reducibility except in terms of the utility of its consequences—that is, you need to motivate it in the same way you would motivate any nonlogical axiom. Moreover, the axiom is *ad hoc*. It violates the spirit, if not the letter, of the Vicious Circle Principle,[11] and (to some extent) it undermines its own motivation.[12]

That is not all. The apparatus of orders and the need for Reducibility produce a formulation of logic that struck many mathematicians as awkward, cluttered, and inelegant. Indeed, the mathematicians were probably more responsible for the demise of Ramified Type theory than the philosophers.[13]

Frank Ramsey, who was both a mathematician and a philosopher, set about to assist this trend by removing the need for orders in solving the paradoxes.

[9] This formulation of the ramified theory of types is intentionally sketchy. The details of the theory are complex, and not centrally relevant to the project of this paper.

[10] See [Copi, 1971][pp.91–99] and [Fraenkel *et al.*, 1973][pp. 150–153].

[11] Speaking of the incongruity of Reducibility, Quine says "Whatever sense of security we may have drawn from the constructional metaphor is now, therefore, forfeited." [Quine, 1963][p. 251].

[12] Quine remarks "If Russell's system with its axiom of reducibility is free from contradiction, then we may be sure that no contradiction would ensue if we were simply to repudiate all but predicative orders," following the remark up with a simple proof. [Quine, 1963][p. 253].

[13] Writing in 1937, for instance, Wilhelm Ackerman says (in the preface [Hilbert and Ackermann, 1937], "It was possible to shorten the fourth chapter inasmuch as it was no longer necessary to go into Whitehead and Russell's ramified theory of types, since it seems to have been generally abandoned." (Translated by L.M. Hammond, G.G. Leckie, and F. Steinhart.)

3. Ramsey's division of the paradoxes

In [Ramsey, 1925][pp. 20–25], Ramsey lists eight paradoxes, evidently taken from [Whitehead and Russell, 1910–1913][Chapter II]). He divides the eight into two so-called "fundamentally distinct groups," labeling them simply the "A paradoxes" and "B paradoxes." Ramsey's group A includes the Russell Paradox; his group B includes the Liar. He then characterizes the division as follows:

> Group A consists of contradictions which, were no provision made against them, would occur in a logical or mathematical system itself. They involve only logical or mathematical terms such as class and number, and show that there must be something wrong with our logic or mathematics. But the contradictions of Group B 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. If so, they would not be relevant to mathematics or to logic, if by 'logic' we mean a symbolic system.

Ascribing to Peano the idea that Group B paradoxes pertain to linguistics rather than to logic, Ramsey adds that he would prefer to say that they belong to "epistemology." The suggestion that these paradoxes are empirical is somewhat implausible and has not gained general acceptance. But as long as they do not seem to be formulated in purely logical terms or to require a solution in these terms, Ramsey does not need to locate these paradoxes precisely or even to point to a solution.[14] As long as the distinction asks logic to solve only the group A paradoxes, logic can get along with only the levels required by the Simple Theory of Types.

In the same essay [Ramsey, 1925][pp. 14–16],[15] Ramsey points out that the body of mathematics at which *Principia* aimed involves only extensional devices such as classes and relations-in-extension, so that the intensional foundations of *Principia* are not required for Russell and Whitehead's logicist program. The point is related to Ramsey's division of the paradoxes and to the dispensability of Reducibility, but Ramsey doesn't make the relation explicit and we are uncertain how closely he linked these two themes.

Quine is much more explicit than Ramsey about the connection between extensionality and the dispensability of Reducibility. In an extended discussion of these matters [Quine, 1963][§35], he says that the relevance of ramification to paradoxes concerning falsehood or denotation is dissolved in the presence of a careful distinction between use and mention, and especially between propositional functions and open sentences. And the foundations of mathematics do not need orders and ramification because they do not need intensionality. Although in this work, which is primarily concerned with the theory of sets, Quine does not have a great deal to say about intensionality, he evidently feels that intensional notions such as propositional functions are more problematic than sets and open sentences, that

[14] As we mentioned above, he nevertheless does attempt a solution. But his suggestion was ignored, perhaps because of its somewhat obscure formulation. In any case, this part of [Ramsey, 1925] seems to have had no historical impact whatever.

[15] In fact, this passage, which is part of a discussion of how to reduce all mathematical propositions to tautologies, comes immediately before Ramsey's classification of the paradoxes.

they are not required for the logical treatment of mathematics, and that the regrettable need for Reducibility was due to foolishly incorporating intensionality into the foundations of logic, and was compounded by a pervasive confusion of use and mention. In his later anti-intensional writings, Quine appeals to other considerations, but his discussion of reducibility in [Quine, 1963] makes it clear that systematic concerns having to do with the paradoxes and the formalization of mathematics also motivate the banishment of intensionality from logic.

4. Two kinds of truth

Otto Jespersen devotes a chapter of [Jespersen, 1965] to direct and indirect speech, introducing the topic as follows,

> When one wishes to report what someone else says or has said (thinks or has thought)—or what one has said or thought oneself on a previous occasion—two ways are open to one.
>
> Either one gives, or purports to give, the exact words of the speaker (or writer): *direct speech* (oratio recta).
>
> Or else one adapts the words according to the circumstances in which they are now quoted: *indirect speech* (oratio obliqua).
> [Jespersen, 1965][p. 290]

Thus, wishing to report what Bert said to her yesterday, Alice can say either (4.1) or (4.2).

(4.1) Bert said ‘I understand you’.

(4.2) Bert said that he understood me.

Jespersen presents the two forms as stylistic alternatives,[16] and gives evidence from several languages, showing that even in written language the two forms can blend and mingle. (Quotation marks can be omitted in constructions that are partly or entirely direct, and there are exceptions to the rule that tense and person are shifted in indirect but not in direct discourse.)

Philosophers and semanticists sharpen the distinction, and take it much more seriously than speakers of language seem to. During much of the Twentieth Century, analytic philosophers were frequently admonished not to confuse use and mention.

Is the adjective ‘true’ like the verb ‘say’, in supporting both direct and indirect usages? It is hard to say. Naturally occurring examples of ‘true’ with indirect discourse are easy to find, such as the following one from the Brown Corpus:[17]

[16]For linguistic data on the stylistic differences, see [Clark and Gerrig, 1990]. The evidence certainly seems to support a difference in logical form between direct and indirect discourse, although not necessarily a simple one according to which in direct discourse a linguistic expression is mentioned.

[17]This is a corpus collected in 1961, containing over a million words of representative English prose from various genres.

(4.3) It may be true that pool lighting dramatizes an evening scene, but . . . .

Natural examples of 'true' with direct discourse are, apparently, much rarer, if they exist at all. The Brown Corpus exhibits 110 occurrences of 'true' for which the distinction between direct and indirect discourse might arise. (The other 95 occurrences are adjectival or adverbial.) Of these 110, 26 are explicitly indirect. The most common usage (64 examples) makes reference to a previous or subsequent claim made in the text without anything explicit to indicate whether a sentence or a proposition is intended. A typical example is:

(4.4) High-level abstractions are always difficult to pin down with precision. That is particularly true of sovereignty when it is applied to democratic . . . .

In anaphoric cases like this, both the explicit words that have been used and the claim or proposition that has been made are salient, and there is no simple way to tell which of these is demonstrated by 'that'. The fact that an elaboration like 'that claim' sounds more natural than 'that sentence' here may provide weak evidence for a propositional interpretation.

Cases where 'true' is predicated of explicitly quoted material are rare: there are only 2 of them in the Brown Corpus. Here is an example.

(4.5) And a witty American journalist remarked over a century ago what is even more true today, "Many a writer seems to think he is never profound except when he can't understand his own meaning".

But in ordinary usage, quotes are not an unambiguous sign of reference to a linguistic expression, and even cases like (4.5) don't provide altogether convincing instances of 'true' predicated of a sentence: observe that (4.6a) is a much more natural elaboration of (4.5) than (4.6b).

(4.6a) And a witty American journalist remarked over a century ago what is even more true today, and what many contemporary journalists believe as well, . . .

(4.6b) And a witty American journalist remarked over a century ago what is even more true today, and what consists of eighteen words, . . .

As we said, however, the language that is used to deal with truth in philosophical and logical work since the 1970s is more regimented than this, and here at least you can clearly distinguish between indirect discourse forms like (4.7) and (4.8).[18]

[18](4.7a) and (4.7b) are synonymous. (4.7a) is the less natural form with a 'that' clause in subject position. (4.7b) is the more natural extraposed form with expletive 'it'.

(4.7a) That $5 + 7 = 12$ is true.

(4.7b) It is true that $5 + 7 = 12$.

(4.8) '$5 + 7 = 12$' is true.

When (4.7) and (4.8) are formalized, the differences between the two constructions are sharpened. The formalization of (4.8) is unproblematic and metalinguistic: when formalized, (4.8) has the form

(4.9) $T(\ulcorner\phi\urcorner)$,

where T is a first-order predicate, ϕ is a sentence, and $\ulcorner\phi\urcorner$ is an individual term serving as a canonical name of a linguistic expression—in this case, of '$5 + 7 = 12$'. This means not only that $\ulcorner$5+7=12$\urcorner$ names '$5 + 7 = 12$', but that $\ulcorner$5+7=12$\urcorner$ should integrate with syntactic predicates if they are present, or if they are added. So if a syntactic theory is added to the formalization language, we should expect the resulting theory to account for how the structure of the formula '$5 + 7 = 12$' can be recovered from the name $\ulcorner 5 + 7 = 12\urcorner$. For instance, there will be sentences of the theory involving $\ulcorner 5+7 = 12\urcorner$ saying that '$5+7 = 12$' consists of a certain number of symbols in a certain order; and if the axioms of the syntactic theory are adequate, such a sentence will be be provable if it is true, and disprovable if it is false.

There is no universally agreed-on policy for the formalization of indirect discourse forms like (4.7a) and (4.7b). If we follow the policy used in most versions of modal logic, where

(4.10) It is necessary that $5 + 7 = 12$

would be formalized as

(4.11) $\Box(5 + 7 = 12)$,

then we would formalize (4.7a) and (4.7b) as

(4.12) $T(5 + 7 = 12)$,

where now T is a modal operator rather than a one-place first-order predicate.

Formalizations like (4.12) are automatically consistent as additions to modal logics of the familiar sort. In a Kripke frame that uses the identity relation over worlds to interpret T, the analog (4.13) of Convention T is valid.

(4.13) $T(\phi) \leftrightarrow \phi$

The valid formulas of a theory that has models will be consistent. Kripke models of modal logics therefore provide a guarantee that indirect discourse versions of the Liar Paradox cannot arise when truth is formalized as in (4.13). This protection against paradox applies to any extension of propositional modal logic that has Kripke models; in particular, it applies to type-theoretic extensions like Montague's Intensional Logic [Montague, 1970, Gallin, 1975] with quantification over propositional types.

In Montague's Intensional Logic, (4.7) would be formalized as

$$(4.14)\ T({}^{\wedge}[5+7=12]),$$

where T has the type $\langle\langle s,t\rangle,t\rangle$ (and so would denote a function from sets of worlds[19] to truth-values), and ${}^{\wedge}\phi$ denotes the function that takes a possible world into the denotation of ϕ in that world. Formulas like (4.15) are admitted

$$(4.15)\ \forall x\,[T(x) \rightarrow {}^{\vee}x],$$

where the variable x has type $\langle s,t\rangle$ and so ranges over sets of worlds, and, where ϕ has type $\langle s,\tau\rangle$, ${}^{\vee}\phi$ denotes in a world w the denotation of ϕ in w.

As with ordinary modal logic, consistency is not a problem. Under the same interpretation of T using the identity relation,

$$(4.16)\ \forall x\,[T(x) \leftrightarrow {}^{\vee}x]$$

will be valid in this enriched logic. (Here, as before, x is a variable of type $\langle s,t\rangle$.)

Modal logic with propositions interpreted as sets of possible worlds therefore provides a paradox-free setting for intensional logic. However, the modal approach is committed to the closure of propositional attitudes under logical equivalence: where μ is any propositional operator of a modal logic, $\mu(\phi) \leftrightarrow \mu(\psi)$ is valid if $\phi \leftrightarrow \psi$ is valid. This makes modality unsatisfactory as a general treatment of propositional attitudes.

The discussion of this issue in the philosophical literature goes back to Kathleen Johnson Wu's critique of Jaakko Hintikka's defense of "logical omniscience" in epistemic logic; see [Wu, 1970, Hintikka, 1970]. Since then, an extensive literature on the topic has developed in philosophy and in computer science, where the problem is particularly acute because computer scientists are interested in applications of epistemic logic to agents with limited reasoning power.[20]

Some philosophers, especially Robert Stalnaker, have defended modal logic by claiming that propositional attitudes, if properly understood, do exhibit logical omniscience; see [Stalnaker, 1984]. But it is hard to see how to extend defenses of this kind, which depend on contextual parameters, to cases like (4.17), and approaches to intensionality based on modal logic over Kripke frames are clearly unsuited to many computational applications.

[19] Here, sets of worlds are themselves functions from worlds to truth-values.

[20] See, for instance, [Fagin *et al.*, 1995][Chapter 9].

(4.17a) Liz is aware that $5 + 7 = 12$.

(4.17b) Liz is aware that for all positive integers n, if there are positive integers i, j, and k such that $i^n + j^n = k^n$ then $n \leq 2$.

But if we are willing to work with formalizations of intensionality that, unlike modal logic with Kripke frames, fail to validate logical omniscience, the situation with regard to the paradoxes is no longer so straightforward, especially when we work without well-behaved models and introduce axioms that are inconsistent with logical omniscience. Here, we will have to reassess the question of freedom from paradoxicality.

To recapitulate: the direct and indirect formalizations of truth predications in (4.9) and (4.12) have much in common. They share the form

(4.18) $T(\langle\!\langle 5 + 7 = 12 \rangle\!\rangle)$,

where $\langle\!\langle\ \rangle\!\rangle$ creates a syntactic environment that is either like regimented quotation, or like the 'that' of indirect discourse. Also, both formalizations allow quantification into the argument position of the truth predicate.

But these two formalizations differ dramatically with regard to the Liar Paradox. In the presence of an adequate syntactic theory, (4.18) is inconsistent with Convention T when $\langle\!\langle\ \rangle\!\rangle$ is interpreted as quotation. But even in the presence of an adequate theory of sets of worlds, (4.18) is consistent with Convention T.

The concept of a proposition that comes to us from Frege and Russell is not syntactic, but it is clearly intended to admit the possibility of different propositions that are true in the same possible worlds. As we said, there are good reasons, motivated by the semantic behavior of propositional attitudes, for taking these more finely individuated propositions seriously.

Therefore, the question arises whether, when we interpret $\langle\!\langle\ \rangle\!\rangle$ as creating references to propositions of this sort, quantification over propositions and Convention T or similar schemes will produce paradoxes.

In at least some cases, we know that this can happen. In [Myhill, 1958], John Myhill showed that Alonzo Church's first formalization of the Logic of Sense and Denotation[21] was inconsistent, using an argument based on cardinalities. But the general case has not been much explored, and it is worthwhile to ask whether a formalization of propositions that does justice to the requirements of propositional attitudes and that allows unrestricted quantification over propositions can hope to avoid paradoxes like the Liar.

Unfortunately, it will turn out that paradoxes do arise very generally in formal settings of this type, and that in fact some of these paradoxes are very like traditional formulations of the Liar Paradox.

Furthermore, these "paradoxes of intensionality" lie outside of Ramsey's classification of the paradoxes and the toolkit of solutions that go along with this classification. Paradoxes of intensionality are not resolved by appeals to a metalinguistic hierarchy, unless we take

[21] In [Church, 1951].

an otherwise not well-motivated syntactic approach to propositional attitudes. Nor are they resolved by any of the set-theoretic solutions to the Russell Paradox.

5. Russell's intensional paradox

Russell, who included propositional quantifiers in his systems of logic, presents an example of an intensional paradox in Appendix B of the *Principles of Mathematics.* The problem has to do with the cardinality of the set of propositions. It is similar to the Russell paradox. But it is not addressed by Simple Type Theory, which has no propositional types.

> If m be a class of propositions, the proposition "every m is true" may or may not be itself an m. But there is a one-one relation of the proposition to m: if n be different from m, "every n is true" is not the same proposition as "every m is true." Consider now the whole class of propositions of the form "every m is true," and having the property of not being members of their respective m's. Let this class be w, and let p be the proposition "every w is true." If p is a w, it must possess the defining property of w; but this property demands that p should not be a w. On the other hand, if p be not a w, the p does possess the defining property of w and therefore is a w. Thus the contradiction appears unavoidable. ... To sum up: it appears that the special contradiction of Chapter X [the Russell Paradox] is solved by the doctrine of types, but that there is at least one closely analogous contradiction which is probably not soluble by this doctrine. The totality of all logical objects, or of all propositions, involves, it would seem, a fundamental logical difficulty. What the complete solution of the difficulty may be, I have not succeeded in 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.[22]

We will use the type framework of [Thomason, 1980] to formalize this argument. The type system resembles that of Montague's Intensional Logic, but intensionality is introduced with a primitive type p of propositions. Since this framework is neutral as to what propositions are (they could be truth-values, sets of possible worlds, sentences from a "language of thought," Fregean senses, or platonic abstractions of some other sort), it provides a conveniently general and ontology-neutral medium for this purpose. Where α is an expression of type p, ${}^{\vee}\alpha$ now denotes the truth-value of the proposition denoted by α.

In reproducing Russell's argument, we have to think of his "class" of propositions as a propositional function that inputs a proposition and returns a proposition: a function of type $\langle p, p\rangle$. If this "class" were a function of type $\langle p, t\rangle$, there would be no way to speak sensibly of *the* proposition that every member of the class is true. Of course, where x has type p, we can formulate the following condition on x:

$$(5.1)\quad {}^{\vee}x \;=\; \forall y\,[G(y) \rightarrow {}^{\vee}y].$$

Here, G has type $\langle p, t\rangle$ and y has type p.

[22] [Russell, 1903], pp. 527–528. The last sentence of the quotation is also the last of the *The Principles of Mathematics.*

In this sort of type theory, universal quantification appears as a typed operator, and formulas like (5.1) are more properly formulated using lambda abstraction:

$$(5.2)\quad {}^{\vee}x \;=\; \forall(\lambda y\,[G(y) \rightarrow {}^{\vee}y]).$$

In (5.2), $\forall$ has type $\langle\langle p, t\rangle, t\rangle$.

Condition (5.2) ensures that the proposition denoted by x will be true if and only if every member of the class denoted by G is true. But this is far from guaranteeing the uniqueness that is required for the paradox.

As a first step in representing Russell's argument, we intensionalize (5.2) to obtain a representation of the *proposition* saying that all propositions in the class denoted by G are true. For this purpose, we need intensional analogs of extensional logical operators. For example, to formulate *the proposition* that if $1 > 0$ then $2 > 0$ we will need a conditional operator $\rightsquigarrow$ of type $\langle p, \langle p, p\rangle\rangle$. If we use '$\phi \rightsquigarrow \psi$' to abbreviate '$[\rightsquigarrow(\phi)](\psi)$', then $[1 > 0 \rightsquigarrow 2 > 0]$ denotes the proposition that if $1 > 0$ then $2 > 0$, and ${}^{\vee}[1 > 0 \rightsquigarrow 2 > 0]$ denotes the truth value of this proposition. We will introduce the following suite of intensional operators.

Intensional = (over objects of type τ):	$\approx$, type $\langle\tau, \langle\tau, p\rangle\rangle$
Intensional $\neg$:	$\sim$, type $\langle p, p\rangle$
Intensional $\wedge$:	$\cap$, type $\langle p, \langle p, p\rangle\rangle$
Intensional $\vee$:	$\cup$, type $\langle p, \langle p, p\rangle\rangle$
Intensional $\rightarrow$:	$\rightsquigarrow$, type $\langle p, \langle p, p\rangle\rangle$
Intensional $\forall$: (over objects of type τ):	Ɐ, type $\langle\langle\tau, p\rangle, p\rangle$

Figure 1: Intensional operators

In the absence of a specific reification of propositions, there are few if any plausible intensional constraints to be placed on the intensional operators. However, it is reasonable to require that these operators reduce to the corresponding extensional operators. The following *extensional homomorphism* principles do this.

(5.3) Extensional isomorphism principles for $\approx, \sim, \rightsquigarrow, \cap, \cup$, and Ɐ:

(5.3a) $^{\vee}[\alpha \approx \beta] \leftrightarrow [\alpha = \beta]$

(5.3b) $^{\vee}[\sim\phi] \leftrightarrow [\neg{}^{\vee}\phi]$

(5.3c) $^{\vee}[\phi \rightsquigarrow \psi] \leftrightarrow [^{\vee}\phi \rightarrow^{\vee} \psi]$

(5.3d) $^{\vee}[\phi \cap \psi] \leftrightarrow [^{\vee}\phi \wedge {}^{\vee}\psi]$

(5.3e) $^{\vee}[\phi \cup \psi] \leftrightarrow [^{\vee}\phi \vee {}^{\vee}\psi]$

(5.3f) $^{\vee}[\text{Ɐ}x^{\tau}\phi] \leftrightarrow \forall x^{\tau}\, {}^{\vee}\phi$

In (5.3a–f), ϕ and ψ are arbitrary expressions of type p.

Returning now to the formalization of Russell's intensional paradox, we state the intensional assumption on which it depends: that the proposition that every proposition satisfying a propositional function is true depends uniquely on that propositional function. This assumption, which we formalize below as Assumption (5.4), belongs to a family of intensional constraints entailing that propositions cannot just be sets of possible worlds. It requires that propositions are individuated in ways that depend not just on their truth conditions, but on how they combine with propositional attitudes, or on the way these propositions are expressed. We will call such assumptions *Principles of fine-grained propositional individuation.*

At this point, we begin to use a convention of using superscripts to type the first occurrence of a constant or variable in a formula.

$$(5.4)\ \ \forall m^{\langle p,p\rangle}\forall n^{\langle p,p\rangle}[[\text{Ɐ}x^p[m(x)\rightsquigarrow x] = \text{Ɐ}x^p[n(x)\rightsquigarrow x]]\rightarrow m=n]$$

We define a functional expression w of type $\langle p,p\rangle$ and a formula P of propositional type as follows.

$$(5.5)\ \ w = \lambda x^p{\sim}\text{Ɐ}m^{\langle p,p\rangle}{\sim}[[x\approx\forall y^p[m(y)\rightsquigarrow y]]\cap{\sim}m(x)]$$

$$(5.6)\ \ P = \text{Ɐ}x^p[w(x)\rightsquigarrow x]$$

Russell's intensional paradox is that (5.4), (5.5), and (5.6) are inconsistent: a contradiction is forthcoming if ${}^{\vee}w(P)$ is assumed; but the contrary assumption is also contradictory. See Appendix A of this paper for a formalization of the argument.

We now define a "logical product" operation on propositional functions.

$$(5.7)\ \ L^{\langle\langle p,p\rangle,p\rangle} = \lambda f^{\langle p,p\rangle}\text{Ɐ}(\lambda x^p[f(x)\rightsquigarrow x])$$

As indicated, L has type $\langle\langle p,p\rangle,p\rangle$. Given an expression G of type $\langle p,p\rangle$, denoting a propositional function on propositions, $L(G)$ denotes the proposition that every proposition satisfying G is true.

Using logical product, we can obtain the following nasty corollary of Russell's intensional paradox, which dramatizes its unhappy consequences for the theory of propositional attitudes.

$$(5.8)\ \ \exists f^{\langle p,p\rangle}\exists g^{\langle p,p\rangle}[\exists x^p[f(x)\wedge\neg g(x)] \\ \wedge\ \Box\forall y^e[\mathbf{Believe}(y,L(f))\leftrightarrow\mathbf{Believe}(y,L(g))]]$$

There are propositional functions that differ. But necessarily, the proposition that every proposition satisfying the first function is true and the proposition that every proposition satisfying the second is function is true have the same belief conditions, for arbitrary agents.

This is hard to swallow, if you take propositional attitudes at all seriously. If propositional functions have different extensions, it ought to be possible to believe the logical product of one one of them without believing that of the other.

Russell's reaction to the difficulty, of course, was to develop the Ramified Theory of Types, which as we saw in §1 gave way after Ramsey's critique to the Simple Theory of Types. Within the Simple Theory of Types, however, there seem to be only three possible reactions to the problem: (i) embrace conclusions such as (5.8), or (ii) deny principles of fine-grained propositional individuation, or (iii) somehow change the logic without ramifying.

With respect only to this problem, the first alternative is entertainable. Maybe the discriminating capabilities of even idealized epistemic agents are limited, so that if two functions are sufficiently complex no agent will be able to distinguish their logical products. But if we widen the field of difficulties to include not only Russell's intensional paradox but a family of paradoxes discussed by Arthur Prior in a 1961 article, this alternative seems less hopeful.

6. The empirical paradoxes

Just as the preceding problem resembled the Russell paradox, there are intensional paradoxes that resemble the semantical paradoxes.[23] The most obvious of these paradoxes are "empirical" versions of the Liar Paradox.[24] Unlike formulations of the direct discourse Liar Paradox that rely only on the presence of a syntactic theory, the paradoxes of intensionality rely on the possibility of agents having certain attitudes: for instance, the possibility of a Cretan saying certain things, or believing certain things.

Arthur Prior provided an extended discussion of these paradoxes in [Prior, 1961]. This paper is unusual (almost unique) in concentrating on the intensional paradoxes.

It is (let us take this for granted) a matter of fact that Epimenides was a Cretan, and it *seems* to be a matter of fact that he said that everything a Cretan (ever) says is false. We may not believe *what* Epimenides said (we had better not, if we can reason with propositional quantifiers). But at least we believe that he said it.

Let us reiterate what we said about direct and indirect discourse in Section 4. Although the Epimenides Paradox, as we have stated it, is often mentioned in the literature on the Liar Paradox, it is usually not distinguished from the version of the Liar that runs "This sentence is false." But the two forms are not at all equivalent: (6.1) is (certainly) false, because Epimenides didn't speak English, whereas (6.2) is (probably) true, if we can trust the historical sources.

(6.1) Epimenides said 'Everything a Cretan said is false'.

(6.2) Epimenides said that everything a Cretan said is false.

Prior uses propositional quantification to formalize (6.2). In the framework we used to formalize Russell's intensional paradox, the paradox takes the following form:[25]

(6.3) $^{\vee}[\mathbf{Say}^{\langle p,p\rangle}(\forall x^{p}[\mathbf{Say}(x) \rightsquigarrow \sim x])]$

The general form of Prior's Epimenides Paradox is then:

(6.4) $^{\vee}[F^{\langle p,p\rangle}(\forall x^{p}[F(x) \rightsquigarrow \sim x])]$

Prior himself makes no distinction between expressions of type p (expressions that denote propositions) and expressions of type t (expressions that denote truth-values), or between

[23]Of course, neither of the analogies is perfect—the point of this paper is that the intensional paradoxes fall through the cracks of Ramsey's division, so they had better be disanalogous in important ways.

[24]In fact, it seems likely that the earliest formulations of the Liar that we know have this form: a Cretan says that everything a Cretan says is false.

[25]Prior assumes propositional quantification, in its unramified form, throughout [Prior, 1961].

boolean operators and the corresponding intensional operators. We believe that it is helpful to make these distinctions explicit, and will work with formalizations like (6.4).

(6.4) denotes a truth-value. If, for example, F denotes the propositional function of being said by a Cretan, (6.4) will denote the truth-value of the proposition that everything a Cretan says is false.

We continue to follow Prior's train of thought, using this notation to formalize it. First, (6.4) has the following two consequences:[26]

(6.5) $\neg^{\vee}[\forall x^p[F(x) \leadsto \sim x]]$

(6.6) $\exists x^p[{}^{\vee}F(x) \wedge {}^{\vee}x]$

These two formulas are logically equivalent; we mention them both only because the first makes it evident that the argument of F in (6.4) is false, while the second shows that some Cretan saying must be true.

Deriving (6.5) from (6.4) is a simple exercise, but it does involve instantiating the variable x in ${}^{\vee}F(\forall x^p[F(x) \leadsto \sim x])$ with this same formula, $\forall x^p[F(x) \leadsto \sim x]$.

The empirical paradoxes depend on a contingent premise, which nevertheless intuitively could be true, even if in fact it is false. In this case, suppose that as a matter of fact, Cretans are very laconic—the only other thing a Cretan ever says is that $7+5=11$. If (6.4) were true, then because of (6.6) some proposition a Cretan says must be true. But we have assumed that the only propositions a Cretan says are that $7+5=11$ and the proposition denoted by $\forall x^p[F(x) \leadsto \sim x]$, i.e. the proposition that everything a Cretan says is false. And we know independently that it is false that $7+5=11$, while in view of (6.5) it must be false that everything a Cretan says is false. That is, (6.4) is false.

In view of this argument, Prior concludes that, in a world in which a Cretan has said that $7+5=11$ and no Cretan has yet said anything else, it is impossible for a Cretan to say that everything a Cretan says is false. That is, there can be empirical situations which prevent a Cretan from saying something (from being in the appropriate relation to a proposition) or more generally, which can prevent an agent from having a propositional attitude to a proposition, even though the usual prerequisites for that circumstance are present. (We can assume, in this hypothetical situation, that Epimenides uttered the appropriate words.) Epimenides must not have said anything on the problematic occasion. Prior accepts this conclusion somewhat reluctantly, having this to say about it.

> ... I must confess that all I can say to allay the misgivings expressed in the past four sections is that so far as I have been able to find out, my terms are the best at present offering. I have been driven to my conclusion very unwillingly, and have as it were wrested from Logic the very most that I can for myself and others who feel as I do. So far as I can see, we must just accept the fact that thinking, fearing, etc., because they are attitudes in which we put ourselves in relation to the real world, must from time to time be oddly blocked by factors in the world, and we must just let Logic teach us

[26] That is, (6.5) and (6.6) must be true in any model of the type theory that satisfies (6.4). We are appealing here to the model theory of [Thomason, 1980].

where these blockages will be encountered. [Prior, 1961][p. 32].

It may be easier for us to accept this conclusion now than when Prior wrote his paper. Hilary Putnam, David Kaplan, and many other philosophers of language have urged that what you say or think depends on general on the circumstances, and that the "internal relations" of the speaker will not always suffice to fix a reference. If this is accepted, it is not surprising that *whether* anything is said or thought could also be risky.

You can even use Prior's techniques to construct Putnam-like examples, without having to resort to science-fiction-like hypotheticals. Imagine that for some reason Ralph, who is in Room 17 but doesn't realize that this is where he is, thinks to himself to the effect that[27] whatever anyone in Room 17 thinks to himself then is false. Unknown to Ralph, someone else—Annie—is hiding in the room. There are two cases: (1) Annie thinks to herself that $7 + 5 = 12$, and (2) Annie thinks to herself that $7 + 5 = 11$. According to Prior, Ralph is thinking something in case (1), but in case (2) he isn't. But nothing about Ralph's internal state will reveal this.

With this example, we begin to see the generality of (6.4) and its logical consequences as a source of problematic examples. F in ${}^{\vee}F^{\langle p,p\rangle}(\forall x^p\,[F(x) \leadsto \sim x])$ can be instantiated with any propositional attitude. We can start with a general attitude type, like thinking or expecting, and qualify it in any way we like—restricting the agent, the time, the place, and any other circumstances we care to choose. If we can do this in such a way that all the other instances in which the qualified attitude is instantiated are false, we have an empirical paradox.

Prior recounts [Prior, 1961][p. 29] an elaboration which he attributes to Michael Dummett. According to one popular view of what happened when Epimenides spoke, he uttered certain words (of Greek) that in virtue of the conventional rules of the language are associated in each context of utterance with a proposition. To simplify things, we can suppose that there are no indexicals in Epimenides' hypothetical sentence; then we can forget the context of utterance. But, although Epimenides' *words* are conventionally associated with the proposition that everything a Cretan says is false, we know that speech acts can misfire in various ways. Prior postulates a *logical* misfire in the case of Epimenides' utterance, which prevents him from saying anything when he makes the utterance.

As Prior presents it, Dummett's idea is to let F in (6.4) stand for 'Epimenides speaks words of Greek that conventionally signify (in Greek) ... '. (The dots here stand for an argument position of type p.)

It seems to follow that Epimenides *can't even utter the words.* This, of course, is unacceptable. Prior's response [Prior, 1961][p. 29] is to suggest that signifying "can't be infallibly effected by our conventions.' As far as we can see, this would rule out a theoretical approach to semantics. You can't put semantics on a proper footing without some way of drawing the encyclopedia/dictionary distinction—some way of making it possible to allow semantics to

[27]We use the awkward phrase "thinks to himself to the effect that" to indicate the motions that someone would go through normally in thinking something, and that would create the presumption that in going through these motions they had indeed thought something. The phrase sounds so awkward because there is no reason in the ordinary course of affairs to distinguish between going through the motions of thinking something and actually thinking it.

assign interpretations to phrases—and propositions to sentences—by local rules that are not forced to appeal to arbitrary and apparently irrelevant contingencies.

Dummett's example doesn't strike us as calling for such drastic measures at all, though in the present context it may give this appearance. We need to remember that whoever adopts Simple Type Theory is likely to have the Tarski hierarchy in his repertoire of puzzle solving devices. And this case is well suited to a Tarskian cure. Of course, '___ utters words of Greek that conventionally signify (in Greek) ... ' is a relation between an individual and a proposition: its type is $\langle p, \langle e, p\rangle\rangle$, the same as that of '___ believes ... '. But it is a *semantical* relation.[28] If we can convince ourselves that the **L**-expression relation, which relates an individual (a sentence) and a proposition, and so has type $\langle e, p\rangle$, is not definable in **L**, similar considerations should persuade us that Dummett's relation isn't definable in **L**.

But Prior considers another elaboration that has nothing metalinguistic about it and that is potentially much more damaging.[29] Consider an example in which Tarski thinks to himself: "Snow is white." Ordinarily, you'd suppose that Tarski has thought that snow is white. But unfortunately, someone else (whom we will call "Gödel") gets there first. Just before Tarski's act of thought, Gödel thinks to himself: "Either whatever I am now thinking is true and whatever Tarski will think immediately afterwards is false, or whatever I am now thinking is false and whatever Tarski will think immediately afterwards is true."

We can formalize the proposition that Gödel thinks as follows.

$$\begin{aligned}(6.7)\quad & [\boldsymbol{\forall} x^p\,[G(x) \rightsquigarrow x] \cap \boldsymbol{\forall} x^p\,[T(x) \rightsquigarrow\sim x]]\\ & \cup\ [\boldsymbol{\forall} x^p\,[G(x) \rightsquigarrow\sim x] \cap \boldsymbol{\forall} x^p\,[T(x) \rightsquigarrow x]]\end{aligned}$$

Suppose that as a matter of fact, Gödel thinks this, and nothing else, and that immediately afterwards Tarski thinks that snow is white, and nothing else. Also, we know that snow is white. We show first that the proposition expressed by (6.7) is false.

If this proposition is true, then (using Boolean Homomorphism), (6.8) denotes $\top$.

$$\begin{aligned}(6.8)\quad & [\forall x^p\,[{}^{\vee}G(x) \rightarrow {}^{\vee}x] \wedge \forall x^p\,[{}^{\vee}T(x) \rightarrow \neg{}^{\vee}x]]\\ & \vee\ [\forall x^p\,[{}^{\vee}G(x) \rightarrow \neg{}^{\vee}x] \wedge \forall x^p\,[{}^{\vee}T(x) \rightarrow {}^{\vee}x]]\end{aligned}$$

If $\forall x^p\,[{}^{\vee}G(x) \rightarrow \neg{}^{\vee}x] \wedge \forall x^p\,[{}^{\vee}T(x) \rightarrow {}^{\vee}x]$ denotes $\top$ then—since the only thing Gödel thinks is the proposition denoted by (6.7)—this proposition is false, contrary to assumption. On the other hand, $\forall x^p\,[{}^{\vee}G(x) \rightarrow {}^{\vee}x] \wedge \forall x^p\,[{}^{\vee}T(x) \rightarrow \neg{}^{\vee}x]$ can't denote $\top$, because the only thing Tarski thinks is the true proposition that snow is white. So our assumption that the proposition expressed by (6.7) is true (i.e., that (6.8) denotes $\top$) has led to a contradiction.

But if (6.8) denotes $\perp$, then (6.9) denotes $\top$.

$$\begin{aligned}(6.9)\quad & [\exists x^p\,[{}^{\vee}G(x) \wedge \neg{}^{\vee}x] \vee \exists x^p\,[{}^{\vee}T(x) \wedge {}^{\vee}x]]\\ & \wedge\ [\exists x^p\,[{}^{\vee}G(x) \wedge {}^{\vee}x] \vee \exists x^p\,[{}^{\vee}T(x) \wedge \neg{}^{\vee}x]]\end{aligned}$$

[28]Whatever this means. We are very much in need of tests (even relatively unreliable ones) that can help us to tell which predicates are semantical.

[29]Prior attributes examples of this kind to Jean Buridan.

But the second conjunct of (6.9) denotes $\perp$, since by assumption the only proposition Gödel thinks is false and the only proposition Tarski thinks is true. Therefore, our empirical assumptions are inconsistent. But it is difficult to say which of them is wrong, and it is cases like this that lead Prior to the uncomfortable solution he offers in the passage we quoted above.

In this example, Prior's explanation is that Tarski can't have managed to think anything after all, despite the apparently innocuous *content* of what he tried to think. Prior doesn't back his diagnosis up with a detailed account of the conditions under which agents can successfully have attitudes, but the general idea seems to be that here Gödel's act of thought trumps Tarski's because Gödel gets his thought in first. Perhaps the idea is that propositions are served out on a first-come first-served basis, and so a seemingly innocuous attempt to think something can be blocked by logic from being thought in paradoxical circumstances like this.

But if this sort of theory were right, elaborations of the Gödel-Tarski Paradox give malicious prior preemptors far too much scope. These elaborations don't even have to be hypothetical. For instance, you, the reader, may have felt as you read this paper that you were having thoughts, and that some of these were true. We can now reveal that you were mistaken. We are, of course, now writing this paper before you have had a chance to read it. And one of us is now thinking that either what he is thinking now or the logical sum of whatever is thought by anyone while reading this paper is false, but not both. On Prior's account, you can't succeed in having any true thoughts while reading this paper. And it is too late for you to do anything about this.

Indeed, for reasons like this, Prior's account seems to imply that we could never be sure, when we seek to engage a proposition with a propositional attitude, that we have actually managed to relate ourselves to the proposition we had in mind. Unlucky enough to have a malicious precursor, a person could go through an entire life without ever thinking, suspecting, or doubting anything.

Furthermore, our attempt to state what happens when someone attempts to engage a proposition but fails is subject to the same sort of paradoxical argument that any other attitude is. Prior wants to say that Epimenides didn't in fact say[30] anything. But (perhaps as part of an explanation of why he failed to say anything) we need to say what did he do.

We are tempted to say that Epimenides tried (unsuccessfully) to say something, or that he made as if to say something, or that he simulated saying something. In each case, we can reintroduce the paradox by substituting for F in (6.4) the predicate that we obtain by deleting 'something' from these formulations. This would lead to consider the case of a Cretan who, for instance, tries to say that everything a Cretan tries to say is false. Epimenides can't try to say that everything a Cretan tries to say is false. It can't seem to Prior that Epimenides can try to say that everything it seems to Prior that Epimenides can try to say is false. But then we are left with no very good way to describe what the person who is logically blocked from relating successfully to a proposition does do—or else we are left with a problematic regress of "trying to say" or "making as if to say."

We conclude that Prior's way out of the paradox is hopeless. Tentatively, or perhaps as a challenge to any philosopher who wants to work out such a theory, we suggest that

[30] This is the 'say' of indirect discourse.

attempts to develop, within the framework of a Simple Type Theory, a plausible theory of "propositional acts" or relations of epistemic agents to propositions that will resolve these paradoxes are likewise hopeless.

7. Some possible solutions

You can't help feeling that there is an asymmetry in content between sentences like 'Everything a schizophrenic fears is false' and '$7+5=12$', and that Prior's solution is flawed in allowing Tarski's relatively simple thought in the Gödel-Tarski Paradox to be blocked by Gödel's complex thought, which involves propositional quantification. The restricted comprehension axiom of Zermelo set theory suggests an approach that would do more justice to this difference. Just as the set theorist errs in assuming unrestricted set comprehension, in the form

$$(7.1)\quad \exists x \forall y\,[y \in x \leftrightarrow \phi]\,,$$

we could try to trace Russell's intensional paradox to the following principle, which says that every propositional function of type $\langle p, p\rangle$ possesses a logical product.

$$(7.2)\quad \forall f^{\langle p,p\rangle} \exists x^p\,[x = \forall\!\!\!\!\!\;\,\, y^p\,[f(y) \rightsquigarrow y]]$$

Of course, there are differences along with the set-theoretical analogy: (7.1) is not a principle of logic. Since (7.2) is validated by the semantics of quantification in Simple Type Theory, we can't do away with it without adjusting the logic of Simple Type Theory.

The following argument, whose last step is (7.2), indicates what will have to be discarded.

$$\begin{aligned}(7.3)\quad &(1)\ \forall y^p\,[f(y) \rightsquigarrow y] = \forall y^p\,[f(y) \rightsquigarrow y]\\ &(2)\ \exists x^p\,[x = \forall y^p\,[f(y) \rightsquigarrow y]]\\ &(3)\ \forall f^{\langle p,p\rangle} \exists x^p\,[x = \forall y^p\,[f(y) \rightsquigarrow y]]\end{aligned}$$

If we think of Russell's intensional paradox as arising from a discrepancy between the domain of propositions (the values of propositional variables) and the language's ability to form expressions of propositional type, the most natural object of suspicion is the inference from Step (1) to Step (2). Invalidating this inference[31] would result in a logic of *partial* propositional functions.

This might also help with the empirical paradoxes. If the proposition that everything a Cretan says is false doesn't exist in the paradoxical situation, then this can explain why Epimenides has said nothing. But if the facts are different, and on another occasion a Cretan has said that $7+5=12$, then a Cretan has said something true and we are inclined

[31] The mechanics of this are fairly straightforward, and we will not go into details here. It is only necessary to allow models in which the domain of propositions is not closed under all the operations definable in the logical language, and to adopt one of the standard policies for dealing with the resulting truth-value gaps.

to conclude that here, Epimenides has said something false. Surely, however, he couldn't have said something false without saying something.

We have been maneuvered at this point into saying that the existence of propositions is a contingent affair. This is likely to complicate our theory of set existence.

As usual, we can make matters worse by elaborating the empirical paradox. Let's go back to the original Epimenides, and imagine that a non-Cretan kibitzer says that everything a Cretan says is false. Epimenides, we agreed, said nothing. And the only thing that Cretans ever say is that $5 + 7 = 11$. (We have *supposed* this, in setting up the problem.) But it is false that $5 + 7 = 11$, and so it follows that everything a Cretan says is false. So the kibitzer has said something true, and unproblematic. But this is just the proposition that we had to rule out of existence, in order to prevent Epimenides from saying anything. Now the existence of propositions is not only contingent, but speaker-relative as well.

Also (and this is typical too of similar approaches to the direct discourse semantic paradoxes), if we develop a logic of partial propositional functions with a proof theory, it will be very difficult to avoid the existence of formulas ϕ of type p such that both ${}^{\vee}\phi$ and $\neg\exists x^p\, [x = \phi]$ are provable. Such cases tend to undermine the motivation of the theory.

This approach looks more promising than Prior's, but on the whole it is still pretty dismal.

An alternative approach, and one well worth considering, is to explore the idea that the problem with both Russell's intensional paradox and the empirical intensional paradoxes is unrestricted quantification over propositions. To do this is to reopen the intensional ramified theory of types as a serious logical alternative.[32]

This idea leads to a project that is beyond the scope of this paper. But we do wish to point out that rehabilitating Ramified Type Theory is not as hopeless a suggestion now as it would have been, say, in 1950. Attitudes towards intensionality are not as hostile now as they were then. Some people may perhaps take the logicist program as seriously as it was in Russell's day, but it is less easy nowadays to take it entirely seriously. And even if we do choose to be logicists, we can still be *extensional logicists*. We can take set-theoretical formalisms based on extensional logics to be adequate for the formalization of mathematics. This leaves us perfectly free to explore ramified type theory as a basis for formalizing intensional phenomena without having to invoke Reducibility or calling into question any of the work that has gone into formalizing analysis and other areas of mathematics.

Ramified Type Theory was also shunned because of its complexity, but since 1925 we have learned a great deal about how to develop complex logics in a way that makes them intelligible and even useful. Perhaps we can do the same for Ramified Type Theory.[33]

[32] Several authors have formulated extensional versions of ramified type theory and investigated their logical properties. This work does not address the problems that originally motivated ramification; in particular, it has nothing to say about the intensional paradoxes. Comparisons such as those found in [Kamareddine *et al.*, 2004][Chapter 3] of the ramified hierarchy with the hierarchy of partial truth predicates that is developed in [Kripke, 1975], for instance, apply only to the extensional case, and it is not clear whether results of this kind can be generalized to an intensional setting.

[33] [Church, 1976] provides a good starting point for this project. But Church does not provide a semantics for his reformulation of Ramified Type Theory.

8. Conclusion

We can now see Quine's rejection of intensionality as more than a philosophical abhorrence for things mental. It served a logical purpose, removing the threat of the paradoxes like those on which we have concentrated in this paper. But contemporary logicians, unlike Quine, can't really afford to deny legitimacy to intensional logics.

The status of logic has changed dramatically since the 1960s. Logic is no longer merely a "foundation" for mathematics (if it ever was such a thing), but is a source of formalisms that are widely used in philosophy, linguistics, economics, and computer science. Even if intensionality is marginal for "pure" mathematics, it is not marginal in these other areas.

In the absence, however, of a ramification revival or some alternative that has not occurred to us, we are not left with a comfortable strategy for dealing with the logical and set-theoretical paradoxes, particularly if we want a strategy that is supported by a rationale that makes it seem general as well as plausible. Despite the problems with the Vicious Circle Principle, there is a great deal to be said for Russell's attempt to diagnose the general cause of the paradoxes and to use this diagnosis to produce an equally general cure.

If we accept Ramsey's twofold classification of the paradoxes, along with generally accepted formalization methods for dealing with both of Ramsey's categories, then it may well seem unnecessary to seek a general, principled solution to the logical paradoxes. However, as we have seen, Ramsey's distinction not only fails to be exhaustive, but leaves out some particularly challenging paradoxes that have no very appealing solution method.

Even if a special-purpose method can be found for these paradoxes, the idea of replacing Ramsey's two-part distinction with a refined, many-fold distinction does not seem very appealing, in the absence of reasons to suppose that the distinction is exhaustive. But we see no way to produce such reasons without a general diagnosis of the logical paradoxes.

Appendix A

Suppose ${}^{\vee}w(P)$. By (5.5), the definition of w, and λ abstraction,

$$^{\vee}{\sim}\forall m^{\langle p,p\rangle}{\sim}[[\forall x^{p}[w(x)\rightsquigarrow x]\approx\forall x^{p}[m(x)\rightsquigarrow x]]\cap{\sim}m(\forall x^{p}[w(x)\rightsquigarrow x])].$$

By the extensional homomorphism conditions (5.3a–f), we have

$$\exists m^{\langle p,p\rangle}[[\forall x^{p}[w(x)\rightsquigarrow x]=\forall x^{p}[m(x)\rightsquigarrow x]]\wedge\neg^{\vee}[m(\forall x^{p}[w(x)\rightsquigarrow x])]].$$

Suppose

$$[\forall x^{p}[w(x)\rightsquigarrow x]=\forall x^{p}[m(x)\rightsquigarrow x]]\wedge\neg^{\vee}[m(\forall x^{p}[w(x)\rightsquigarrow x])].$$

In view of the principle of fine-grained propositional individuation—Condition (5.4)—we can conclude $m=w$. Therefore, $\neg^{\vee}[w(\forall x^{p}[w(x)\rightsquigarrow x])]$, i.e., $\neg^{\vee}w(P)$. By *reductio*, we can conclude $\neg^{\vee}w(P)$.

But then, by (5.5) and lambda abstraction, we have

$$\neg^{\vee}{\sim}\forall m^{\langle p,p\rangle}{\sim}[[\forall x^{p}[w(x)\rightsquigarrow x]\approx\forall x^{p}[m(x)\rightsquigarrow x]]\cap{\sim}m(\forall x^{p}[w(x)\rightsquigarrow x])].$$

By the extensional homomorphism conditions (5.3a–f), we have

$$\forall m^{\langle p,p\rangle}[[\forall x^{p}[w(x)\rightsquigarrow x]=\forall x^{p}[m(x)\rightsquigarrow x]]\rightarrow^{\vee}[m(\forall x^{p}[w(x)\rightsquigarrow x])]].$$

So, in particular,

$$[\forall x^{p}[w(x)\rightsquigarrow x]=\forall x^{p}[w(x)\rightsquigarrow x]]\rightarrow{}^{\vee}[w(\forall x^{p}[w(x)\rightsquigarrow x])].$$

From this, ${}^{\vee}w(\forall x^{p}[w(x)\rightsquigarrow x])$ follows, i.e., we have ${}^{\vee}w(P)$, a contradiction.

Bibliography

[Aczel, 1983] Peter Aczel. *Non-Well-Founded Sets*. Center for the Study of Language and Information, Stanford, California, 1983.

[Barwise and Etchemendy, 1987] Jon Barwise and John Etchemendy. *The Liar*. Oxford University Press, Oxford, 1987.

[Beth, 1959] Evert W. Beth. *The Foundations of Mathematics*. North-Holland Publishing Company, Amsterdam, 1959.

[Braithwaite, 1931] Richard B. Braithwaite, editor. *The Foundations of Mathematics and Other Essays*. Routledge and Kegan Paul, London, 1931. Collected papers of Frank P. Ramsey.

[Burali-Forti, 1897] Cesare Burali-Forti. Una questione sui numeri transfiniti. *Reconditi del Circolo matematico di Palermo II*, 11:154–164, 1897. Reprinted in [van Heijenoort, 1967], pp. 104–111.

[Church, 1951] Alonzo Church. A formulation of the logic of sense and denotation. In Paul Henle, Horace M. Kallen, and Susanne K. Langer, editors, *Structure, Method, and Meaning: Essays in Honor of Henry M. Scheffer*. Liberal Arts Press, New York, 1951.

[Church, 1976] Alonzo Church. Comparison of Russell's resolution of the semantical antinomies with that of Tarski. *Journal of Symbolic Logic*, 41:747–760, 1976.

[Clark and Gerrig, 1990] Herbert H. Clark and Richard J. Gerrig. Quotations as demonstrations. *Language*, 66(4):764–805, 1990.

[Copi, 1971] Irving M. Copi. *The Theory of Logical Types*. Routledge and Kegan Paul, London, 1971. Umich Tanner BC199 .T9 C78.

[Fagin *et al.*, 1995] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Y. Vardi. *Reasoning about Knowledge*. The MIT Press, Cambridge, Massachusetts, 1995.

[Fraenkel and Bar-Hillel, 1958] Abraham A. Fraenkel and Yehoshua Bar-Hillel. *Foundations of Set Theory*. North-Holland Publishing Co., Amsterdam, 1st edition, 1958.

[Fraenkel *et al.*, 1973] Abraham A. Fraenkel, Yehoshua Bar-Hillel, and Azrael Lévy. *Foundations of Set Theory*. North-Holland Publishing Co., Amsterdam, 2nd edition, 1973.

[Gallin, 1975] Daniel Gallin. *Intensional and Higher-Order Logic*. North-Holland Publishing Company, Amsterdam, 1975.

[Giaquinto, 2002] Marcus Giaquinto. *The Search for Certainty: A Philosophical Account of Foundations of Mathematics*. Oxford University Press, Oxford, 2002.

[Gödel, 1944] Kurt Gödel. Russell's mathematical logic. In Paul A. Schilpp, editor, *The Philosophy of Bertrand Russell*, pages 125–153. The Tudor Publishing Company, New York, 1944.

[Hilbert and Ackermann, 1937] David Hilbert and Wilhelm Ackermann. *Grundzüge der theoretischen Logik*. Springer-Verlag, Berlin, second edition, 1937.

[Hintikka, 1970] Jaakko Hintikka. Knowledge, belief, and logical consequence. *Ajatus*, 32:32–47, 1970.

[Jespersen, 1965] Otto Jespersen. *The Philosophy of Grammar*. W.W. Norton, New York, 1965. (This work was originally published in 1924.).

[Kamareddine *et al.*, 2004] Farouz Kamareddine, Twan Laan, and Rob Nederfelt. *A Modern Perspective on Type Theory—From its Origins until Today*. Kluwer Academic Publishers, Dordrecht, 2004.

[Kneale and Kneale, 1962] William Kneale and Martha Kneale. *The Development of Logic*. Oxford University Press, Oxford, 1962.

[Kripke, 1975] Saul Kripke. Outline of a theory of truth. *Journal of Philosophy*, 72:690–715, 1975.

[Montague, 1970] Richard Montague. Pragmatics and intensional logic. *Synthése*, 22:68–94, 1970. Reprinted in *Formal Philosophy*, by Richard Montague, Yale University Press, New Haven, CT, 1974, pp. 119–147.

[Myhill, 1958] John Myhill. Problems arising in the formalization of intensional logic. *Logique et Analyse*, 1:78–83, 1958.

[Priest, 2005] Graham Priest. *Doubt Truth to be a Liar*. Oxford University Press, Oxford, 2005.

[Prior, 1961] Arthur N. Prior. On a family of paradoxes. *Notre Dame Journal of Formal Logic*, 2:16–32, 1961.

[Quine, 1963] Willard V.O. Quine. *Set Theory and Its Logic*. Harvard University Press, Cambridge, Massachusetts, 1963.

[Ramsey, 1925] Frank P. Ramsey. The foundations of mathematics. *Proceedings of the London Mathematical Society, Series 2*, 25(5):338–384, 1925. Reprinted in [Braithwaite, 1931].

[Russell, 1903] Bertrand Russell. *The Principles of Mathematics*. Cambridge University Press, Cambridge, England 1903.

[Russell, 1908] Bertrand Russell. Mathematical logic as based on the theory of types. *American Journal of Mathematics*, 30:222–262, 1908.

[Russell, 1937] Bertrand Russell. *The Principles of Mathematics*. George Allen and Unwin, London, second edition, 1937.

[Russell, 1959] Bertrand Russell. *My Philosophical Development*. Allen & Unwin, London, 1959.

[Stalnaker, 1984] Robert C. Stalnaker. *Inquiry*. The MIT Press, Cambridge, Massachusetts, 1984.

[Thomason, 1980] Richmond H. Thomason. A model theory for propositional attitudes. *Linguistics and Philosophy*, 4:47–70, 1980.

[van Heijenoort, 1967] Jan van Heijenoort, editor. *From Frege to Gödel*. Harvard University Press, Cambridge, Massachusetts, 1967.

[Whitehead and Russell, 1910–1913] Alfred North Whitehead and Bertrand Russell. *Principia Mathematica*. Cambridge University Press, Cambridge, England, first edition, 1910–1913. Three volumes.

[Whitehead and Russell, 1925–1927] Alfred North Whitehead and Bertrand Russell. *Principia Mathematica*. Cambridge University Press, Cambridge, England, second edition, 1925–1927. Three volumes.

[Wu, 1970] Kathleen Johnson Wu. Hintikka and defensibility. *Ajatus*, 32:25–31, 1970.