Worlds and Propositions Set Free*

OTÁVIO BUENO
University of Miami

CHRISTOPHER MENZEL
Texas A&M University

EDWARD N. ZALTA
Stanford University

Abstract

The authors provide an object-theoretic analysis of two paradoxes in the theory of possible worlds and propositions stemming from Russell and Kaplan. After laying out the paradoxes, the authors provide a brief overview of object theory and point out how syntactic restrictions that prevent object-theoretic versions of the classical paradoxes are justified philosophically. The authors then trace the origins of the Russell paradox to a problematic application of set theory in the definition of worlds. Next the authors show that an object-theoretic analysis of the Kaplan paradox reveals that there is no genuine paradox at all, as the central premise of the paradox is simply a logical falsehood and hence can be rejected on the strongest possible grounds — not only in object theory but for the very framework of propositional modal logic in which Kaplan frames his argument. The authors close by fending off a possible objection that object theory avoids the Russell paradox only by refusing to incorporate set theory and, hence, that the object-theoretic solution is only a consequence of the theory's weakness.

 This paper is published in *Erkenntnis*, 79 (2014): 797–820. The 2nd and 3rd authors would like to express their gratitude to Hannes Leitgeb and the Munich Center for Mathematical Philosophy (MCMP), where both spent time as visiting fellows during the 2011–12 academic year, and to the Alexander von Humboldt Foundation, for its support of the MCMP. All three authors wish to express their indebtedness to two anonymous reviewers for detailed and perceptive comments on an earlier version of this paper. Their comments led to substantive changes.

Introduction

Paradoxes deriving from Russell (1903) and Kaplan (1995) show that paradoxes can arise quickly in an ontology that includes both sets and either propositions or possible worlds. Potent versions of these paradoxes arise for particular philosophical theories that identify worlds with sets of propositions (Adams 1981, Plantinga 1974,[1] and Pollock 1984) or propositions with sets of worlds (Lewis 1986). Moreover, these paradoxes threaten to arise as well for semantic theories that assume a certain amount of set theory and take worlds as primitive, e.g., in the interpretation of modal languages (Lewis 1970, Montague 1974). The paradoxes suggest that the foundations for such assumptions are not secure.

In sorting through these issues, important questions arise, for example, whether to take possible worlds as primitive and define propositions, or take propositions as primitive and define worlds. On the issue of possible worlds and propositions, Stalnaker (1976) comes down in favor of taking possible worlds as primitive and defining propositions. As part of his argument for this claim, Stalnaker notes:

> Whatever propositions are, if there are propositions at all then there are sets of them, and for any set of propositions, it is something determinately true or false that all the members of the set are true.[2] (*ibid.*, 73)

Our argument in what follows, if correct, shows that the first sentence is in error. We develop a metaphysical theory that can do the work one expects of a foundational theory but that asserts the existence of propositions, and derives the existence of worlds, without requiring that there are sets of any of these entities. We try to show that it is better to take propositions as primitive 0-place relations, axiomatize them (as part of a general theory of n-place relations), and then define worlds in terms of a background theory of objects. At least, we will show that one can do all this in a way that is paradox-free and which provides a metaphysical foundation for using possible worlds when doing the semantics of modal logic.

[1] Plantinga defines worlds to be complex states of affairs but also claims that there is, for each world w, the *book on* w, i.e., the set of propositions that would be true if w were actual.

[2] This passage is preserved in Stalnaker 1984 (55) and Stalnaker 2003 (36).

The argument in what follows also undermines a claim of Sider's (2002, 307) concerning world theories and cardinality problems:

> *Every* theory of worlds encounters trouble in this area. The linguistic ersatzist, for example, may admit arbitrarily large worlds, but cannot admit a world with so many individuals that they cannot all be members of a set (except in special cases where the objects display symmetries allowing simpler description); for a linguistic-ersatz world is a maximal consistent set of sentences, and sentences themselves are also sets. That is a limitation on possibility, although a bit less severe than an upper bound on world size. Similar problems confront other views that identify possible worlds with abstract entities other than sets of sentences.

We shall be defending a view that identifies possible worlds as abstract entities, but we shall establish that no obvious size-based argument that places limits on possibility affects our theory. Moreover the view we shall defend avoids a similar problem that arises for Lewis (1986, 101–104) that forces him to restrict his principle of recombination in such a way that there must be a cardinal upper bound on the number of worlds and *possibilia*.

In §1 below, we lay out the Russell and Kaplan paradoxes in detail. In §2, we present the formal theory of objects, propositions, and worlds and describe why the classical paradoxes, as well as object-theoretic versions of them, do not arise. In §3, we trace the origins of Russell's 1903 paradox to the identification of possible worlds with sets of propositions and point out how the paradox simply cannot arise in the object-theoretic analysis of possible worlds, which does not rely upon the apparatus of set theory. In §4, we show that the Kaplan paradox does not arise for object theory — despite the fact that the allegedly paradoxical principle **K**, introduced in §1.2 below, can be formulated in it. We then argue that **K** is not paradoxical and, indeed, that the reason it is not paradoxical can be adopted even for friends of possible world semantics who define propositions as sets of worlds. Moreover, we argue, the reason **K** is not paradoxical is obscured when the paradox is analyzed from within the set-theoretic framework of possible world semantics itself. Finally, in §5, we address a potential objection, namely, that object theory avoids the 1903 paradox at the cost of eliminating set theory.

Our moral is that an uncritical introduction of sets into the foundations of metaphysics is perilous. It can lead, in the worst case, to outright paradox and, even when it does not, it can lead to a perception of paradox where, in fact, none exists. As we demonstrate, object theory provides an alternative framework that allows the formulation of a rigorous logic of properties, relations, and propositions and a comprehensive theory of possible worlds without invoking sets and, hence, without the corresponding perils. Worlds and propositions can be set free.

1 The Paradoxes

1.1 Worlds as Sets of Propositions: Russell's *PoM* Paradox

Not long after Russell discovered the famous set-theoretic paradox that now bears his name, he discovered another paradox involving propositions that he published in 1903 in Appendix B (§500) to *The Principles of Mathematics* (hereafter, *PoM*).[3] The premises are as follows:

(1) There is a set P of all propositions.

(2) For every set S of propositions, there is a uniquely identifiable proposition p_S.[4]

(3) For sets S, S' of propositions, if $p_S = p_{S'}$, then $S = S'$.

Now, for any set S of propositions, since p_S is a proposition, we have either $p_S \in S$ or $p_S \notin S$. Let $R \subseteq P$ consist of exactly those of the latter sort, that is:

B $\quad q \in R \leftrightarrow \exists S \subseteq P(q = p_S \wedge q \notin S)$.

Since $R \subseteq P$, by (2) we have a corresponding proposition p_R and the obvious question: $p_R \in R$? Suppose so. Then by **B**, for some $S \subseteq P$, $p_R = p_S$ and $p_R \notin S$. By (3), $R = S$. Hence, $p_R \notin R$. So suppose not, i.e., that $p_R \notin R$. Then by **B** once again, for any $S \subseteq P$, if $p_R = p_S$, then

[3] Today the latter paradox is usually taken to show that there can be no set of all propositions although, at the time, prior to Zermelo's axiomatization of set theory and his own ramified type theory, Russell did not fix upon a single premise as particularly problematic.

[4] Russell identifies p_S with S's "logical product", i.e., the proposition *that every member of S is true*, but this is entirely incidental to the proof. All that matters is that the propositions p_S satisfy (3).

$p_R \in S$. So, in particular, since $R \subseteq P$ and, obviously, $p_R = p_R$, it follows that $p_R \in R$. Contradiction. Call this Russell's *PoM* paradox.

Say that a set of propositions is *maximal* if it contains, for every proposition p, either p or its complement $\neg p$. The set P of all propositions, if it exists, is of course maximal. It is, moreover, obviously inconsistent, in the sense that it is impossible that its members be simultaneously true. Interestingly, however, replacing 'proposition' with 'truth' uniformly throughout the above argument yields a corresponding paradox for the set of all *truths* which, of course, if it exists, is just as obviously consistent as P is inconsistent.

This is significant, of course, as it shows that there are serious problems with the otherwise intuitive idea of defining possible worlds as maximal, consistent sets of propositions — a definition proposed initially by Adams (1974, 225) that persists in modern expositions, e.g., Stalnaker and Oderberg (2009, 48). The argument above shows that, given minimal assumptions, postulating a particularly significant instance of one such world — the *actual* world, the world that contains all and only true propositions — leads to contradiction.

The *PoM* paradox cannot be generalized to arbitrary possible worlds without introducing irreducibly modal notions into the proof.[5] However, with a bit of additional set theory and a somewhat stronger account of propositions, a variation on the paradox can be formulated that applies to any maximal set of propositions and hence, in particular, to any possible world in the sense at hand. The premises of this paradox are as follows:

(4) There is a maximal set of propositions.

(5) Every proposition q has a complement $\neg q$.

(6) For every set S of propositions, there exists a uniquely identifiable proposition p_S.

(7) If S and S' are distinct sets of propositions, then p_S, $p_{S'}$, and their complements are pairwise distinct.

(5)–(7) seem reasonable on any "fine-grained" understanding of propositions. Indeed (6) is rather weak, for all that is required for its truth is that every set be correlated with only one proposition, say, the proposition *that S has at least one member*. Intuitively, this proposition is specifically about S and, hence, is distinct from the proposition *that S' has at least one member*, when $S \neq S'$. However, assuming some basic set theory, these premises are inconsistent.

By (4), let S^* be a maximal set of propositions. By (5), (6), and the definition of maximality, for every member S of the power set $\wp(S^*)$ of S^*, either $q_S \in S^*$ or $\neg q_S \in S^*$ (perhaps both). By the axiom of Separation, let R consist of all such propositions, that is, let

$$R = \{p \in S^* : \exists S \in \wp(S^*)(p = q_S \vee p = \neg q_S)\}.$$

Since $R \subseteq S^*$, it follows that R is no larger than S^*. However, R contains, for every $S \in \wp(S^*)$, q_S or $\neg q_S$ and, by (7), all such propositions are pairwise distinct. Hence, we can map $\wp(S^*)$ one-to-one[6] into R and, hence, $\wp(S^*)$ is no larger than R. It follows that $\wp(S^*)$ is no larger than S^*, contradicting Cantor's theorem that, for all sets S, $\wp(S)$ is strictly larger than S.[7]

Call this the *generalized PoM* paradox. This paradox challenges a number of other theories of possible worlds. Plantinga (1974) defines worlds to be *states of affairs* of a certain sort, but also postulates that, corresponding to each possible world w is the *book* on w, that is, the set of all propositions true at w (*ibid.*, 44–46).[8] It is easy to demonstrate that every such book is a world in exactly the sense of Adams (1974) and, hence, that it is subject to the paradox above.[9]

Adams (1981) provides a more sophisticated version of his earlier theory. In this account, he first defines a possible world (or "world story") as above to be a maximal consistent set of propositions (*ibid.*, 21ff). He then

[5]For example, one can alter (2) to stipulate that, for every $S \subseteq P$, there is a *necessarily true* proposition p_S.

[6]The preceding fact guarantees a one-to-many mapping from $\wp(S^*)$ into R. To derive a one-to-one mapping f it is necessary only that, for $S \in \wp(S^*)$, f "choose" q_S or $\neg q_S$ in those cases where R contains both. For example, by the Powerset and Separation axioms, we can define the mapping $f : \wp(S^*) \longrightarrow R$ such that $f(S) = q_S$ if $q_S \in R$ and $f(S) = \neg q_S$, otherwise. This mapping "chooses" q_S when both it and $\neg q_S$ are in R.

[7]This is a somewhat tighter and more general version of the paradox in Bringsjord (1985). See also follow-up discussions by Menzel (1986), Grim (1986), and Menzel (2012).

[8]Where a proposition p is true at a world w just in case, had w obtained, p would have been true.

[9]See also Chihara (1998, 126-7), who reconstructs basically the paradox here directly in terms of Plantinga's states of affairs.

proceeds to qualify the definition so as to reflect his "existentialism", that is, his view that singular propositions are ontologically dependent on the individuals they are about.[10] Specifically, worlds in which an individual *a* fails to exist (that is, worlds lacking the proposition *that a exists*) contain no propositions involving *a* as a "constituent". Given this qualification, some possible worlds turn out not to be maximal in the sense above. However, it has no effect on the nature of worlds containing the same individuals as the actual world — notably, of course, the actual world itself, that is, the set of true propositions. Such worlds are still maximal in the sense above and, hence, the paradox above still applies to Adams' account.

Another theory that also faces the paradox has been developed by Lycan and Shapiro (1986). The theory explicitly identifies worlds with sets of propositions (*ibid*., 345). There are constraints, however, on the language in which the theory is formulated so that the identification of worlds with sets of propositions cannot be made in the object language. No such constraints are found in the metalanguage, and the paradox immediately arises. Alternatively, at the object level, by extending the language so that sets are included, the paradox can be immediately formulated. Additional restrictions would then need to be included, but it is unclear how that could be done in a principled way.

1.2 Propositions as Sets of Worlds: Kaplan's Paradox

Paradoxes deriving from Kaplan (1995) concern those theories — notably, David Lewis's theory of concrete worlds — that follow possible world semantics in defining propositions to be sets of worlds.[11] Kaplan expresses the paradox in terms of the unsatisfiability of a certain intuitively possible sentence in a modal language L^+ with propositional quantifiers and nonlogical sentence operators, viz.:

K $\forall p \Diamond \forall q (Eq \leftrightarrow q = p).$

Intuitively, **K** implies that there is a property E of propositions such that, for every proposition p, it is possible that p, and only p, has E. The problem is that, if, as in standard possible world semantics (PWS), the modal operators are quantifiers over worlds and every set of worlds is a proposition, **K** is unsatisfiable, as there will not be enough worlds to go around; there will always be a proposition that falsifies **K**.

Following the sketch in Davies (1981, 262), Lewis (1986, 104-5) expresses the paradox in more informal terms with, in particular, a definite example of the property E — the property of being *entertained* (at some fixed time t). On this interpretation, then, **K** says — not implausibly, on the face of it — that every proposition could be *uniquely* entertained, i.e., every proposition could be the only proposition that is entertained by anyone (at the given time t). However, as Kaplan argues, whether or not we can identify a "natural" property to play the role of E shouldn't matter:

> [I]f PWS is to serve for intensional *logic*, we should not build [any] metaphysical prejudices into it. We logicians strive to *serve* ideologies not to constrain them. Thus, insofar as possible, our intensional logic should be neutral with respect to such issues.[12]

Thus, as there seems to be nothing *logically* amiss with **K**, that is, with the idea of a property such that any proposition could be the only thing that has it, Kaplan appears to be arguing that a proper semantics for modal logic should not render it logically false and any semantics that *does* must therefore exhibit some built in metaphysical prejudices. Unfortunately, given a bit of set theory, the unsatisfiability of **K** in standard PWS follows directly.

To see this, note first that the following are standard principles of PWS, extended to languages like L^+ (*ibid*., 43):

(8) There is a set W of all possible worlds.

(9) p is a proposition $=_{df}$ $p \in \wp(W)$.

(10) Modal operators are quantifiers that range over W.

[10] See Menzel (2008, §4.2.2) for a detailed exposition of Adams' account.

[11] The 1995 publication date is much later than Kaplan's actual discovery of the paradox, which he had communicated to a number of philosophers in the late 1970s. See, e.g., Davies (1981, 262).

[12] Lewis's response to the paradox only addresses the particular instance that arises on his informal interpretation of the property E. See Lewis 1986, 106–107, where he in effect argues that, on that interpretation, and given his functionalist theory of mind, $\mathbf{K}_W$, the formulation of **K** in terms of possible worlds (discussed below), is implausible. Perhaps so, but it is clear from the passage quoted here that Kaplan would take no comfort whatever in this response.

(11) Propositional quantifiers range over $\wp(W)$.

(12) Sentence operators O are assigned properties of propositions, that is, functions mapping each world $w \in W$ to a set O_w of propositions.

With some innocuous abuse of notation, the PWS truth condition for **K** is spelled out explicitly in terms of (8)–(12) as follows:

$\mathbf{K}_{\boldsymbol{W}}$ $\quad \forall p \in \wp(W) \exists w \in W \forall q (q \in E_w \leftrightarrow q = p)$.

But, given some set theory, this condition cannot hold. Kaplan himself does not provide a detailed proof but instead (as in Davies' and Lewis's reconstructions) sketches the following simple cardinality argument (*ibid.*, 44). Understood as $\mathbf{K}_{\boldsymbol{W}}$, **K** asserts that, for every proposition p, there is at least one world where p alone has property E. So there have to be at least as many worlds as propositions, i.e., at least as many members of W as $\wp(W)$, in violation of Cantor's Theorem.

We can express this argument a bit more formally so as to make the underlying set-theoretic machinery involved in this argument explicit. Let P be the set $\wp(W)$ of propositions. By the Union and Powerset axioms and an instance of Separation, the Cartesian product $W \times P$ exists[13] and hence, by Separation, we have the relation $f = \{\langle w, p\rangle \in W \times P : \forall q (q \in E_w \leftrightarrow p = q)\}$ that holds between a world w and a proposition p just in case p alone has E in w. It is easy to show that f is a function mapping its domain $U \subseteq W$ onto $P = \wp(W)$,[14] which is of course impossible, by Cantor's Theorem.

So Kaplan's sobering concern "that there is a problem in the conceptual/mathematical foundation of possible world semantics" (*ibid.*, 41) appears to be validated.

[13]Assuming Kuratowski ordered pairs: $W \times P = \{z \in \wp(\wp(W \cup P)) : \exists w \in W \exists p \in P(z = \{\{a\}, \{a, b\}\})\}$. Alternatively, the existence of $W \times P$ follows from Pairing, Union, and Replacement.

[14]To see that f is a function, suppose $\langle w, p\rangle \in f$. Then, by definition of f, $\forall q (q \in E_w \leftrightarrow p = q)$ and, hence, as $p = p$, $p \in E_w$. Suppose then $\langle w, p'\rangle \in f$. Then we have $\forall q (q \in E_w \leftrightarrow p' = q)$ and hence, as $p \in E_w$, it must be that $p' = p$. To see f is onto, suppose $p \in P = \wp(W)$. By $\mathbf{K}_{\boldsymbol{W}}$, there is at least one world w such that $\forall q (q \in E_w \leftrightarrow p = q)$ and, by definition of f, $\langle w, p\rangle \in f$, i.e., $f(w) = p$.

2 Object Theory and the Classical Paradoxes

Our diagnosis will trace the source of the above paradoxes to the volatile interplay of worlds and propositions with sets, particularly when the former entities are *identified* with sets of one sort or another. The framework of our analysis is a basic version of object theory as developed in Zalta 1983, 1993, and elsewhere. We begin with a relatively brief presentation of the theory so that the paper is self-contained. Readers familiar with the basic theory may skip ahead to the next subsection.

2.1 Review of Object Theory

The Language of Object Theory. The theory of objects, propositions and worlds (henceforth 'object theory') that we shall be discussing is couched in a syntactically second-order modal language.[15] Thus, it uses primitive variables ranging over *objects* $(x, y, z, \dots)$ and primitive variables ranging over n-place *relations* ($F^n, G^n, H^n, \dots$, for $n \geq 0$), where *properties* are identified as 1-place relations and *propositions* are identified as 0-place relations. A simultaneous definition of *term* and *formula* is given so that the language includes two atomic forms of predication ($F^n x_1 \dots x_n$ and xF^1) and complex predicates ($[\lambda x_1 \dots x_n\, \varphi]$, for $n \geq 0$).[16] The second atomic formula, xF^1 (hereafter xF) is to be read: x *encodes* F. These encoding formulas express a second mode of predication that has been motivated and applied in a variety of other publications. As we shall see, such formulas, and the axioms stated in terms of them, will make up for the loss of the notion of set membership and the axioms of set theory. For reasons to be discussed below, the complex predicates (λ-expressions) of the language, may not contain any encoding subformulas in φ. Finally, the language of object theory includes a distinguished predicate $E!x$ ('*x is concrete*'). This predicate used to define a critical distinction between *ordinary* objects and *abstract* objects, i.e., between objects that are possibly concrete and those that are not:

[15]We say 'syntactically second-order' because the theory doesn't require full second-order logic. Although we won't pursue the matter here, it should suffice to mention that the theory can be interpreted using *general* models, in which the domain of properties is *not* the full power set of the domain of individuals.

[16]The complete language of object theory also includes definite descriptions, $\imath x \varphi$. As these expressions are not relevant for purposes here, we omit them.

O! $O!x =_{df} \Diamond E!x$

A! $A!x =_{df} \neg\Diamond E!x$.

Basic Logic: Modality, Identity, and λ-Conversion. The basic logic of object theory builds upon a classical second-order S5 quantified modal logic.[17] Thus, both the first- and second-order Barcan formulas are theorems.

Identity is not a primitive in object theory but rather defined: x and y are identical iff they are both ordinary objects and necessarily exemplify the same properties, or they are both abstract objects and necessarily encode the same properties. In other words, we may define:

Id $x = y =_{df} (O!x \wedge O!y \wedge \Box\forall F(Fx \leftrightarrow Fy)) \vee (A!x \wedge A!y \wedge \Box\forall F(xF \leftrightarrow yF))$.

Moreover, identity is defined for properties, relations, and propositions as well. We present here only the definition of identity for properties and propositions, where F, G range over properties and p, q range over propositions:

Id₀ $p = q =_{df} [\lambda x\, p] = [\lambda x\, q]$

Id₁ $F = G =_{df} \Box\forall x(xF \leftrightarrow xG)$.

In other words, properties F and G are identical whenever they are necessarily encoded by the same objects, and propositions p and q are identical whenever the property of *being such that* p is identical to the property of *being such that* q. This reduces proposition identity to that of property identity. Identity for n-place relations ($n \geq 2$) is also reducible to identity for properties, but we shall not present the details here.[18] Though the above definitions allow us to prove $\alpha = \alpha$, for any object variable or relation variable α, object theory takes the traditional principle of the indiscernibility of identicals as an axiom:

Ind $\alpha = \beta \rightarrow (\varphi \rightarrow \varphi')$, where α and β are variables of the same type, β is free for α in φ, and φ' is the result of replacing one or more free occurrence of α in φ with an occurrence of β.

[17] The full theory also includes axioms for rigid definite descriptions.

[18] The definition of relations asserts that F^n and G^n are identical ($n \geq 2$) iff all of the pairwise 1-place relational properties that result from the various ways of 'plugging' $n-1$ arbitrarily chosen objects into F^n and G^n are identical.

Finally, the usual logical axioms governing λ-predicates apply. For purposes here, we highlight only the λ-conversion principle:[19]

LC $[\lambda y_1 \ldots y_n\, \varphi]x_1 \ldots x_n \leftrightarrow \varphi^{x_1,\ldots,x_n}_{y_1,\ldots,y_n}$.

This principle yields comprehension principles for properties, relations, and propositions.[20] Note that we have now identified the most important principles of our theory of properties, relations and propositions, namely, their existence and identity conditions. This theory asserts the existence of a wide variety of complex properties, relations, and propositions, including complex propositional properties of the form $[\lambda y\, p]$ (*being such that* p) and complex propositions of the form $[\lambda\, p]$ (*that* p), where p is a proposition. These latter are governed by special instances of the 0-place case of λ-conversion, namely, $[\lambda\, p] \leftrightarrow p$, which assert: the proposition *that* p is true iff p.

The Logic of Encoding. The encoding axioms of object theory include two axioms and an axiom schema:

Enc $\Diamond xF \rightarrow \Box xF$

O! $O!x \rightarrow \Box\neg\exists F\, xF$

OC $\exists x(A!x \wedge \forall F(xF \leftrightarrow \varphi))$, where x is not free in φ.

The first axiom **Enc** guarantees that encoded properties are rigidly encoded; the properties that an abstract object encodes are not relative to

[19] The other usual axioms for λ-predicates will be assumed: namely that interchange of bound variable makes no difference to the denotation of the λ-predicate ($[\lambda y_1 \ldots y_n\, \varphi] = [\lambda y'_1 \ldots y'_n\, \varphi']$) and elementary λ-expressions denote the *governing* relation ($[\lambda y_1 \ldots y_n\, F^n y_1 \ldots y_n] = F^n$).

[20] For example, after n applications of universal generalization, an application of the rule of necessitation, and an application of existential generalization, we derive the following comprehension principle for relations:

$\exists F^n \Box\forall x_1 \ldots \forall x_n(F^n x_1 \ldots x_n \leftrightarrow \varphi)$, where φ has no free occurrences of F^n and no encoding subformulas

As special cases, we have comprehension for properties and propositions:

$\exists F \Box\forall x(Fx \leftrightarrow \varphi)$, where φ has no free occurrences of F and no encoding subformulas

$\exists p \Box(p \leftrightarrow \varphi)$, where φ has no free occurrences of p and no encoding subformulas any circumstance. The second axiom **O!** tells us that ordinary objects necessarily fail to encode properties. Finally the schema **OC** is a comprehension principle that tells us the conditions under which abstract objects exist. More specifically, **OC** asserts that, for every condition on properties, there is an abstract individual that encodes just the properties satisfying the condition. Given the principle of identity for objects described earlier, we can derive a strengthened version of this principle as a theorem:

(13) $\exists! x(A!x \wedge \forall F(xF \leftrightarrow \varphi))$, where x is not free in φ.

That is, for any condition φ, there is a *unique* abstract object that encodes exactly the properties satisfying φ. There couldn't be two distinct abstract objects encoding exactly the properties satisfying φ since distinct abstract objects must differ by one of their encoded properties.

2.2 How Object Theory Avoids Classical Paradoxes

To see that object theory is immune to the classical paradoxes, it suffices to note that since the theory is formulated in the language of second-order logic, which types the relation and argument places of atomic sentences, Russell's classic paradox does not arise: one cannot assert of properties that they can or cannot exemplify themselves, nor formulate a property that is exemplified by all and only those properties that do not exemplify themselves. However, it is also important to explain: (i) why the syntactic restrictions on the formation of λ-expressions block an object-theoretic version of Russell's paradox, and (ii) why those restrictions are justified by the underlying conception of properties, relations, and propositions. For without at least a reasonably compelling justification, object theory's avoidance of the *PoM* and Kaplan paradox would be a rather hollow victory.

Recall that λ-expressions may not be constructed from formulas φ with encoding subformulas. The reason for this restriction concerns a Russell-style paradox (the "Clark paradox") that arises in the foundations of object theory.[21] This paradox is unrelated to the Russell-Kaplan

[21]This paradox was first reported in the literature in Clark 1978 (184), rehearsed in Rapaport 1978, and formulated more precisely in object theory in Zalta 1983. It was developed independently in Boolos 1987 (17).

paradoxes with which we began our paper, and so it would serve well to rehearse it briefly here.

Suppose we could form the λ-expression '$[\lambda x\, \exists G(xG \wedge \neg Gx)]$' expressing, intuitively, the property *being an object that encodes a property it does not exemplify*. Then we could formulate the following instance of the comprehension principle for abstract objects:

(14) $\exists z(A!z \wedge \forall F(zF \leftrightarrow \forall y(Fy \leftrightarrow [\lambda x\, \exists G(xG \wedge \neg Gx)]y)))$.

With a little bit of reasoning, a contradiction follows.[22] So by banishing encoding formulas from the formation of λ-expressions and relation comprehension, we forestall the Clark paradox.

A related paradox was discovered by Alan McMichael.[23] Suppose that identity ('=') were taken as a primitive. Then one could formulate the λ-expression '$[\lambda y\, y = z]$' expressing, intuitively, the property *being identical with* z. Call such a property a 'haecceity of z'. If such λ-expressions were legitimate, then so would be the following simplified instance of comprehension for abstract objects:

M $\exists x \forall F(xF \leftrightarrow \exists z(F = [\lambda y\, y = z] \wedge \neg zF))$.

M asserts the existence of an object that encodes exactly those haecceities that are not encoded by their instances. But as with (14), from **M** a contradiction quickly ensues.[24]

[22]Let L be the property $[\lambda x\, \exists G(xG \wedge \neg Gx)]$ and let b be a witness to the existential claim (14). Then b encodes all and only the properties F which are materially equivalent to L. Either Lb or $\neg Lb$. Suppose the former. Then by λ-conversion, $\exists G(bG \wedge \neg Gb)$, i.e., there is some property, say Q, such that bQ and $\neg Qb$. But from the former, it follows that $\forall y(Qy \leftrightarrow Ly)$, by definition of b. But from $\neg Qb$, it then follows that $\neg Lb$, contrary to hypothesis. So suppose $\neg Lb$. Then by λ-conversion once again it follows that $\forall G(bG \rightarrow Gb)$, and in particular, $bL \rightarrow Lb$. But, by the definition of b, we know that $bL \leftrightarrow \forall y(Ly \leftrightarrow Ly)$. Since the right-hand side of the biconditional is derivable from logic alone, it follows that bL. Hence, Lb. Contradiction.

[23]This paradox was first reported in McMichael and Zalta (1980). It was discovered independently and reported in Boolos (1987, 17).

[24]Let b be a witness to this existential claim **M**. By definition of b, we know:

M′ $\forall F(bF \leftrightarrow \exists z(F = [\lambda y\, y = z] \wedge \neg zF))$

Now consider the property $[\lambda y\, y = b]$ and suppose $b[\lambda y\, y = b]$. Then by **M′**, it follows that $\exists z([\lambda y\, y = b] = [\lambda y\, y = z] \wedge \neg z[\lambda y\, y = b])$. Call such an object c. So, $[\lambda y\, y = b] = [\lambda y\, y = c] \wedge \neg c[\lambda y\, y = b]$. Note independently that $b = b$ by the laws of identity, from which it follows by λ-conversion that $[\lambda y\, y = b]b$. Since $[\lambda y\, y = b] = [\lambda y\, y = c]$, it follows that $[\lambda y\, y = c]b$. So

Now, the same restriction provided earlier in the context of the Clark paradox, namely, the banishment of encoding subformulas from λ-expressions and comprehension, also solves the McMichael paradox. For in object theory, in the above definition of object identity, the defined notation '$x=y$' is given in terms of a definiens containing encoding subformulas. Thus, '$[\lambda xy\, x=y]$', '$[\lambda y\, y=z]$', and '$[\lambda y\, z=y]$' are all ill-formed, as are the corresponding instances of comprehension.[25] The paradox does not get off the ground.[26]

The fact that banishing encoding formulas from λ-expressions avoids the Clark paradox and the McMichael paradox provides a strong, practical justification for the proscription. However, we believe there are powerful *theoretical* justifications for the move as well.

First, the fact that proscribing the occurrence of encoding subformulas provides a *unified* solution to both the Clark and the McMichael paradoxes itself offers theoretical warrant. Object theory thereby helps to illuminate and unify two paradoxes that may otherwise seem to be importantly different. The paradoxes differ in their formulation but are fundamentally similar in their solution. And by categorizing paradoxes in terms of the way in which they can be solved, object theory offers an understanding of what these paradoxes challenge.

---

by λ-conversion, it follows that $b=c$. But since $\neg c[\lambda y\, y=b]$, it follows that $\neg b[\lambda y\, y=b]$, contrary to hypothesis. So suppose instead $\neg b[\lambda y\, y=b]$. Then, by $\mathbf{M}'$, it follows that $\neg\exists z([\lambda y\, y=b]=[\lambda y\, y=z] \wedge \neg z[\lambda y\, y=b])$, i.e., $\forall z([\lambda y\, y=b]=[\lambda y\, y=z] \rightarrow z[\lambda y\, y=b])$. By instantiating the universal claim to b, we get $[\lambda y\, y=b]=[\lambda y\, y=b] \rightarrow b[\lambda y\, y=b]$. And since the antecedent is true by the laws of identity, it follows that $b[\lambda y\, y=b]$. Contradiction.

[25]Note, however, that the classical, Leibnizian definition of indiscernibility can still be treated as a *bona fide* relation. For the λ-expression '$[\lambda xy\, \forall F(Fx \leftrightarrow Fy)]$' is well-formed. Moreover, in object theory, indiscernibility plays a role in the definiens of the notion of identity for ordinary objects, $x=_E y$, which is defined as: $O!x \wedge O!y \wedge \Box\forall F(Fx \leftrightarrow Fy)$. Thus, the λ-expression '$[\lambda xy\, x=_E y]$' is also well-formed. It denotes a relation that is well-behaved on the ordinary objects — it is provably an equivalence relation on the ordinary objects. So object theory does allow for a relation of identity, but as with all other relations, one can prove that there are abstract objects that are indiscernible with respect to $x=_E y$, as explained in the next footnote.

[26]It is to interesting note that object theory yields the theorem that some distinct abstract objects can't be distinguished by the traditional notion of exemplification. It is provable that for any relation R, there are at least two distinct abstract objects a, b such that $[\lambda x\, Rxa] = [\lambda x\, Rxb]$. Consider the following instance of comprehension:

$$\exists x(A!x \wedge \forall F(xF \leftrightarrow \exists y(F=[\lambda z\, Rzy] \wedge \neg yF)))$$

Call such an object k. From the assumption that k doesn't encode the property $[\lambda z\, Rzk]$, one can prove that k does encode that property. Then, from the definition of k, the fact that k does encode this property yields that there is a distinct object, j, such that $[\lambda z\, Rzk] = [\lambda z\, Rzj]$.

Now from this result, by letting R be $[\lambda xy\, \forall F(Fx \leftrightarrow Fy)]$, one can prove that there are distinct abstract objects a, b such that $\forall F(Fa \leftrightarrow Fb)$. This establishes that there are too many abstract objects for the traditional notion of exemplification to distinguish. Readers who wish to see the reasoning spelled out in detail should consult Zalta (1999, 626 and footnote 16).

Second, a solution to a theoretical problem accrues some theoretical justification if it has no detrimental impact on the power and applicability of the theory. An illustrative case is set theory itself. With set theory beset by paradox, Zermelo (1908) endeavored to axiomatize the theory in a way that preserved its already impressive array of extraordinary results (see, e.g., Kanamori 1996, 2004). Although it would be some twenty years before it was sufficiently understood how the fine structure of the cumulative hierarchy explained the effectiveness of the restrictions on naive comprehension that Zermelo had introduced in the axiom schema of Separation, the power of the theory even with those restrictions provided some warrant for believing that they reflected deeper structural facts about sets. The proscription on encoding subformulas in object theory is analogous — it avoids the paradoxes without compromising the breadth and power of the theory.

However, this analogy is not perfect, as we believe that we can already identify a deeper structural justification for the proscription on encoding subformulas in λ-predicates. We begin with an informal observation about the basic philosophical idea of object theory, namely, that there is a domain of abstract objects that encode properties with which we are already familiar. The underlying intention is simply to assert the existence of new *objects* that are *constituted* by familiar properties; it is not to assert the existence of new and unfamiliar "encoding properties" — properties that, if anything, emerge solely as artifacts of the increased expressive power that encoding predication brings to the language. On this basis alone one is led to disallow encoding formulas from λ-expressions and comprehension.

But there is an even deeper theoretical justification underlying this basic idea. Object theory introduces a special type of connection — encoding — between properties and objects of a special sort. This connection is indeed often considered akin to exemplification — most fruitfully, perhaps, in applications where a philosophical problem can be solved by appealing to an ambiguity in the copula between encoding and exemplification.[27] Parallels aside, however, encoding is completely distinct from exemplification. Encoding is a unique connection between objects (of a special sort) and properties (not n-place relations generally) whose theoretical role is to provide a mechanism whereby properties satisfying certain conditions are unified into a single conceptual object. There is no independent reason to think this metaphysical connection plays a role in the logical structure of fine-grained relations. There is, in fact, reason to think it does not.

On object theory's fine-grained conception of relations, to *be* a relation is to be either a structurally simple n-place universal or to be built up from such by means of a series of natural logical operations: predication, negation, conjunction, quantification, etc. The effect of these operations is reflected in the syntactic structure of λ-predicates (though not necessarily exactly reflected). Structurally simple n-place relations correspond to primitive n-place predicates such as 'R'. Now the relation R that this predicate expresses just *is* the property *exemplifying* R (or *standing in* R) and this is reflected in an object theoretic axiom that governs elementary λ-expressions, namely, $[\lambda x_1...x_n\, Rx_1...x_n] = R$. The argument places in the predicate reveal, as Frege (1891) put it, the unsaturated (*ungesättigt*) or "gappy" nature of relations, which indicates their predicability of any n or fewer objects. No separate operation is needed for a property to be suitably "prepared" for combining logically with objects or other relations: Given F, i.e., $[\lambda x\, Fx]$, we have its negation $[\lambda x\, \neg Fx]$; given $[\lambda y\, Gy]$ we then have the conjunctive relation $[\lambda xy\, \neg Fx \wedge Gy]$; and given an object b, a (2-place) predication operation yields the proposition $[\lambda\, \neg Fb \wedge Gb]$ (the proposition that b is not F but G); and so on.

By contrast, by virtue of what would an encoding property be suitably prepared to combine logically in the same sort of way with objects and other relations? In order to provide a general semantics for λ-predicates containing encoding formulas involving free variables, one must initially have, for properties F, atomic encoding properties $[\lambda x\, xF]$, i.e., properties of the form *encoding* F.[28] But unlike F itself, as $[\lambda x\, xF] \neq F$, $[\lambda x\, xF]$ must be a complex property; that is, its existence must be due to some sort of transformation on F — specifically, a primitive transformation that takes F as input and yields the encoding property $[\lambda x\, xF]$ as output. But unlike all the other operations that yield complex properties, the proposed transformation is *not* a logical operation; unlike predication, negation, conjunction, etc, it is a purely metaphysical transformation corresponding to no logical intuition whatsoever.

To be clear, then, what is being denied here is that there is a logical operation that *transforms* the property F — that is, the property *exemplifying* F — into the property *encoding* F, as such a transformation, we claim, corresponds to no natural logical operation. There is, therefore, simply no philosophical warrant for postulating such a transformation and, hence, no syntactic warrant for permitting encoding formulas to occur in the λ-predicates of object theory. Indeed, the reflections here suggest that permitting them so to occur, and thereby enabling encoding to bleed into the logical structure of relations, would be a type of category mistake.

Far from an *ad hoc* maneuver introduced to avoid paradox, then, the proscription on encoding subformulas in λ-predicates is independently and antecedently motivated by object theory's account of fine-grained relations.

3 Object Theory and the *PoM* Paradoxes

To show that object theory is immune to the *PoM* paradox and its generalization we have to review its theory of possible worlds. In Zalta 1993, worlds are defined as situations, where these in turn are defined as abstract objects that encode only propositional properties:

Sit $\mathit{Situation}(x) =_{df} A!x \wedge \forall F(xF \rightarrow \exists p(F = [\lambda y\, p]))$.

If we let s range over situations, then to say that p *is true in* s ('$s \models p$') is to say that s encodes the propositional property *being such that* p:

TI $s \models p =_{df} s[\lambda y\, p]$.

Finally, possible worlds are defined as situations s that might be such that all and only true propositions are true in s:[29]

[27] For example, in the analysis of fiction, object theory analyzes ordinary natural language claim such as "Holmes is a detective" as ambiguous between an exemplification reading (Dh), which is false, and an encoding reading (hD), which is true. See Zalta 2000a.

[28] Thus, in particular, if there *were* such a property as *encoding* F, $[\lambda x\, xF]$, the predication operation applied to that property and an object b would reasonably be thought to return a proposition, viz., *that* b *encodes* F.

[29] In the definiens of the following definition, the symbol $\leftrightarrow$ dominates $\models$. So the definiens is to be parsed as $\forall p((s \models p) \leftrightarrow p)$, not as $\forall p(s \models (p \leftrightarrow p))$.

PW $World(s) =_{df} \Diamond\forall p(s \models p \leftrightarrow p)$.

In what follows, when s is a world w, we read $w \models p$ as p is *true at* w.

These definitions, plus the definitions of *maximal*, *consistent*, *actual*, etc., allow one to derive the basic principles of world theory as theorems (Zalta 1983, 1993). It is provable that worlds are maximal, consistent, modally-closed, that there is a unique actual world *at* which all and only the truths are true, that a proposition is necessarily true iff true at all possible worlds, and that a proposition is possible iff it is true at some world. The latter theorem constitutes a conditional existence principle for worlds: whenever we add to the theory a proposition p such that $\Diamond p$ and $\neg p$, the principle ensures that there exists a possible world distinct from the actual world at which p is true.[30] Indeed, the definition of identity here is crucial to this latter fact, for it is provable that if there is a proposition, say p, that is true at w and not at w', then $w \neq w'$.[31]

Moreover, in contrast to those that would model worlds as sets of propositions, the worlds of object theory are not *ersatz* worlds. The propositions true at a world are the ones they encode. But encoding is a mode of predication, and as such, the predicate in encoding formulas *characterizes* the subject. Given that object theory treats both encoding and exemplification forms of predication as ways of disambiguating the copula 'is', the definiens of 'p is true in w' (namely, $w[\lambda y\, p]$), can be understood as asserting that *w is such that p*. So the propositions true at a world do in fact *characterize* those worlds.

Given this analysis of possible worlds, and the underlying theory of propositions, we can see why object theory is immune to the Russell paradox. Object theory doesn't employ any set theory in its theory of propositions. In particular, it is not committed to the existence of maximal consistent *sets* of propositions. So neither the *PoM* paradox nor its generalization can even get started.

Because object theory (a) does not have primitive sets, and (b) does not reconstruct worlds as sets of propositions, it is immune to the reasoning in the *PoM* paradoxes. Propositions are instead axiomatized, as a

[30]See Menzel and Zalta 2014 for a study of the smallest models that are required to make this principle, and the axioms used to derive it, true.

[31]Suppose $w \models p$ and $w' \not\models p$ (to show that $w \neq w'$). Then by the definition of $\models$, w encodes a property, namely, $[\lambda y\, p]$, that w' doesn't encode. Since worlds are situations, they are both abstract. So by the definition of identity for abstract objects, w and w' are distinct.

subtheory of the theory of relations, and worlds are defined as abstract objects. No paradoxical set-theoretic correlations between worlds and propositions can be established. So no similar paradoxes by way of set-theoretic considerations threaten to arise.

4 Object Theory and Kaplan's Paradox

Turning to the Kaplan paradox, note that object theory is also not committed to sets of possible worlds or to the analysis of propositions as sets of worlds. So it is not committed to two of the principles used to derive the paradox in §1.2:

(8) There is a set W of all possible worlds.

(9) p is a proposition $=_{df}$ $p \in \wp(W)$.

Hence, the Kaplan paradox too might appear to fall by the wayside. But the full object-theoretic analysis of Kaplan's paradox is both more subtle and more interesting than this. For an argument to the falsity of (an object-theoretic version of) **K** can be given within object theory no less than in PWS, showing thereby that matters here do not depend upon an ontology of sets, primitive worlds, or upon a reduction of propositions to sets of worlds. But, as we will now argue, the object-theoretic argument reveals that there is no actual paradox, not even for the defender of PWS.

4.1 Principle K is Inconsistent with Object Theory

Recall Kaplan's contradictory principle:

K $\forall p \Diamond \forall q(Eq \leftrightarrow q = p)$.

The most natural argument to a contradiction from **K** is blocked in object theory. The reason for this is that the argument requires the construction of a λ-predicate derived from **K** involving identity between propositions.[32] However, identity between propositions is not primitive in object theory but rather is defined in terms of encoding (see $\mathbf{Id_0}$

[32]See fn 36 for further detail.

and $\mathbf{Id_1}$), and the presence of encoding formulas in λ-predicates is proscribed, for reasons defended at length in §2.2. As it happens, however, we can squeeze a contradiction from a weaker principle suggested by Kit Fine,[33] viz.,

K* $\forall p \Diamond (Ep \wedge \forall q(Eq \rightarrow \Box(q \leftrightarrow p)).$

On our informal interpretation of 'E' (which we will continue to use for expository purposes), **K*** says that every proposition could be such that it, and perhaps some propositions necessarily equivalent to it, are the only propositions that are entertained — for short, that every proposition could be *almost-uniquely* entertained.[34] Moreover, it should be clear that, under any interpretation of 'E', **K** entails **K*** and, indeed, **K*** is derivable from **K** in object theory by a simple bit of reasoning in its underlying modal predicate logic.[35] Hence, a contradiction follows from

[33]Fine pointed this out in conversation with one of the present authors at the Australasian Association of Philosophy Conference in Melbourne, Australia, in July 2009.

[34]As Kaplan (1995, fn 9) notes, on the assumption that propositions are identical if necessarily equivalent, **K** can be re-expressed as

K′ $\forall p \Diamond \forall q(Eq \leftrightarrow \Box(q \leftrightarrow p))$

and essentially the same PWS argument as the one we have laid out in §1.2 will still go through. Of course, we do not make the assumptions involved in that argument, as just pointed out above, nor do we assume that necessarily equivalent propositions are identical. Nonetheless, a contradiction is still derivable from **K′** in object theory — essentially the one given for **K***. But, under the informal interpretation of E as the property of being entertained (at some given fixed time t), **K′** expresses the proposition that every proposition could be such that *all* (and only) propositions necessarily equivalent to it are entertained — a proposition that friends of fine-grained propositions, who reject the thesis that necessarily equivalent proposition are identical, will find much less plausible than either **K** or **K***. However, as **K*** is obviously derivable from **K′**, the object-theoretic analysis of **K** will apply no less to **K′**.

[35] By **Ind** and the theorem $\Box(p \leftrightarrow p)$, we have

(A) $p = q \rightarrow \Box(p \leftrightarrow q).$

By (A) and some basic predicate logic we have:

(B) $\forall q(Eq \rightarrow p = q) \rightarrow \forall q(Eq \rightarrow \Box(p \leftrightarrow q))$

and hence also

(C) $(Ep \wedge \forall q(Eq \rightarrow p = q)) \rightarrow (Ep \wedge \forall q(Ep \rightarrow \Box(p \leftrightarrow q))).$

So, by some basic modal predicate logic, we have

(D) $\forall p \Diamond (Ep \wedge \forall q(Eq \rightarrow p = q)) \rightarrow \forall p \Diamond (Ep \wedge \forall q(Ep \rightarrow \Box(p \leftrightarrow q))).$

The antecedent of (D) is easily shown to be logically equivalent to **K** and the consequent is, of course, **K***.

K in object theory no less than in possible world semantics, albeit by a different route.

To show this, we begin by simplifying the expression of **K*** by means of the following definition:

(15) $Up =_{df} Ep \wedge \forall q(Eq \rightarrow \Box(q \leftrightarrow p)).$

This yields the following abbreviated expression of **K***:

K* $\forall p \Diamond Up$

Now define the proposition q_0 as follows:

(16) $q_0 =_{df} [\lambda\, \exists p(\neg p \wedge Up)].$[36]

So, informally understood, q_0 is the proposition that there exists a false proposition p that is almost-uniquely entertained. But from **K*** and the existence of q_0 a contradiction quickly ensues. For the claim that q_0 is not almost-uniquely entertained, i.e.,

(17) $\neg Uq_0,$

is a theorem of object theory. By Necessitation, so is $\Box \neg Uq_0$, i.e., $\neg \Diamond Uq_0$. But the latter proposition is obviously inconsistent with **K***.

To show that (17) is indeed a theorem of object theory, let us note first that, by the definition (16) of q_0, (0-place) λ-conversion, and **Ind**, we have

(18) $q_0 \leftrightarrow \exists p(\neg p \wedge Up).$

Assume (for *reductio*) that (17) is false, i.e., assume Uq_0. Unpacking the latter according to (15), we have:

(19) $Eq_0 \wedge \forall q(Eq \rightarrow \Box(q \leftrightarrow q_0)).$

Now, either q_0 or $\neg q_0$. Suppose the former. Then from (18) it follows that $\exists p(\neg p \wedge Up)$. Let s be such a p; then we have

(20) $\neg s \wedge Us.$

[36]It is at this point that a direct, analogous argument for the inconsistency of **K** in object theory is blocked, as the definition (15) of 'Up', which would instead have been extracted from **K** instead of **K***, would have involved identity (hence encoding formulas) rather than necessary equivalence and, hence, could not be used to construct the λ-predicate in (16).

By unpacking Us, it follows from (20) that $\forall q(Eq \rightarrow \Box(q \leftrightarrow s))$ and so, in particular, $Eq_0 \rightarrow \Box(q_0 \leftrightarrow s)$. By (19) we have Eq_0 and hence $\Box(q_0 \leftrightarrow s)$. By the T schema ($\Box\varphi \rightarrow \varphi$), we have $q_0 \leftrightarrow s$ and, by (20) once again, $\neg s$. Hence, $\neg q_0$, contradicting our assumption that q_0.

So suppose instead that $\neg q_0$. By (18), it follows by a bit of predicate logic that $\forall p(Up \rightarrow p)$ and hence, in particular that $Uq_0 \rightarrow q_0$. But we assumed at the outset that Uq_0, so it follows that q_0, contradicting our assumption $\neg q_0$. Either way, we get a contradiction from the assumption that Uq_0.

So our theorem (17) shows that a contradiction can be derived from **K*** in object theory and, hence, as **K*** is derivable from **K**, **K** is inconsistent with object theory.

4.2 Analysis: Principle K is *Logically False*

The crux of Kaplan's argument that **K** leads to a genuine paradox is that, on the face of it, there is nothing logically amiss with **K**. And, indeed, under the informal interpretation we've been considering, *prima facie*, **K** seems, at the least, logically consistent, perhaps even plausible: intuitively, for any given proposition p, p alone could have been the only proposition entertained by any rational agent. Thus, Kaplan's argument appears to reveal a genuine paradox, a genuine inconsistency between, on the one hand, the (modal) logical intuitions that ground **K** and, on the other, the set-theoretic assumptions and semantic intuitions that ground the framework of possible world semantics: the intuitions in question are all intuitively compelling and the set-theoretic assumptions are, at the least, theoretically necessary; but not all of them can be true.

However, as the early history of set theory has shown us, intuitions are not always a reliable guide to consistency. *Prima facie*, the Naive Comprehension principle of naive set theory is extremely compelling:

NC $\forall G\exists y\forall x(x \in y \leftrightarrow Gx)$.

What could be more obvious than the principle that the things satisfying some condition jointly constitute a *set* of things satisfying that condition? If you have the things, how do you not have the set? However, by instantiating $\forall G$ to $[\lambda z\, z \notin z]$ in **NC** and performing λ-conversion, we have that Russell's set of all non-self-membered sets exists:

R $\exists A\forall x(x \in A \leftrightarrow x \notin x)$.

However, the problem here was not so much with the notion of *set* — though, of course, there *was* a problem, later resolved (in the minds of many) with the development of Zermelo-Frankel set theory (ZF) and, subsequently, its natural iterative models (Zermelo 1930, Boolos 1971) — but rather with the fact that framing matters in term of sets obscured the more fundamental problem, viz., that **NC** is in fact a logical falsehood. Intuitions about sets notwithstanding, the falsity of **NC** is an instance of a general theorem of second-order logic:[37]

(21) $\forall R\exists G\neg\exists y\forall x(Rxy \leftrightarrow Gx)$.

The object theoretic analysis reveals that the situation with regard to **K** is analogous. For note that, although we have framed the argument in the preceding section as an argument for **K**'s inconsistency with *object theory*, **K** is in fact *logically* inconsistent — the argument is given entirely in object theory's underlying second-order modal predicate logic. No distinctively object-theoretic principles are involved. Rather, it is simply a theorem of the second-order modal logic underlying object theory that, for *any* property F of propositions, there will be a proposition p that can't possibly be the only proposition that has F,[38]

(22) $\forall F\exists p\neg\Diamond\forall q(Fq \leftrightarrow q = p)$,

Hence, there is in particular a proposition that can't possibly be the only proposition that has property E, under any interpretation and, hence, that **K**, like **NC**, is logically false — a fact initially obscured, perhaps, by framing matters in terms of a naive notion of entertainment that endowed **K** with an initial measure of plausibility, just as the logical falsity of **NC** was obscured by framing matters in terms of a naive notion of set.

Now, importantly, note that, not only is the proof of **K**'s logical falsity not in any way distinctively object theoretic, the full framework of object

[37] To see this is a theorem, suppose its negation, for *reductio*. Then $\exists R\forall G\exists y\forall x(Rxy \leftrightarrow Gx)$. Let S be an arbitrary such R. Then $\forall G\exists y\forall x(Sxy \leftrightarrow Gx)$. Now instantiate to $[\lambda z\, \neg Szz]$ and perform λ-conversion to obtain: $\exists y\forall x(Sxy \leftrightarrow \neg Sxx)$. Let a be such a y: $\forall x(Sxa \leftrightarrow \neg Sxx)$. By instantiating to a, we get: $Saa \leftrightarrow \neg Saa$.

[38] Since we have derived a contradiction from **K*** in the preceding section and have shown (see fn 35) that **K*** follows from **K**, it follows that $\neg$**K** — i.e., $\neg\forall p\Diamond\forall q(Eq \leftrightarrow q = p)$ — is a theorem and, hence, by some elementary quantifier logic, so too is $\exists p\neg\Diamond\forall q(Eq \leftrightarrow q = p)$. Since we've assumed no special facts about 'E', it can be regarded as arbitrary and so we have our theorem (22) by universally generalizing on 'E'.

theory's underlying second-order modal predicate logic is far more apparatus than is necessary. Notably, the proof can be reconstructed in a simple proof theory for the very framework of modal logic on which Kaplan himself bases his alleged paradox: one needs only the simple modal logic that includes a quantified version of the T axiom $\forall p(\Box p \to p)$, standard axioms for propositional quantifiers and identity, and an obviously valid comprehension schema for propositions:[39]

C $\exists r(r \leftrightarrow \varphi)$

for any formula φ of the language L^+ of the framework in which r does not occur free.[40]

This realization casts an entirely different light on Kaplan's alleged paradox. For one is saddled with a paradox only if there are no compelling grounds for rejecting any of the jointly inconsistent propositions that drive it. But the preceding analysis reveals that there are in fact the most compelling grounds imaginable: **K** is false as a matter of logic alone — its initial intuitive plausibility notwithstanding, its falsity is completely and satisfyingly explained by its refutability from compelling logical principles of (extended) modal propositional logic.[41] Hence, there simply is no genuine paradox.

[39]The requisite proposition r for each instance of **C** will of course simply be the set of worlds in which φ is true.

[40]The proof of **K**'s inconsistency runs as follows. As an instance of **C**, we have:

(E) $\exists r(r \leftrightarrow \exists p(\neg p \wedge \forall q(Eq \leftrightarrow p = q)))$.

Let q_0 be one such proposition r (the only such, of course, in the context of PWS). Then we have:

(F) $q_0 \leftrightarrow \exists p(\neg p \wedge \forall q(Eq \leftrightarrow p = q))$.

It can now be shown by a *reductio* argument very similar to the one in the previous section (starting with proposition (18)) that

(G) $\neg\forall q(Eq \leftrightarrow q = q_0)$

is a theorem of the logic in question and, hence, by necessitation and existential generalization that

(H) $\exists p \Box \neg\forall q(Eq \leftrightarrow q = p)$,

which, of course, is equivalent to the negation of **K**.

[41]Anderson (2009) similarly points out that the negation of **K** (which he calls *anti-(A)*, *ibid.*, 92) is a theorem in an extended modal propositional logic like the one we have sketched here and concludes as we do that the issue at root has nothing to do with PWS *per se*. Curiously, however, he does not draw any parallels with the lessons of naive set theory but, rather, seems to agree with Kaplan that **K**'s initial plausibility under some interpreta-

Why might one miss this simple analysis? The answer, we contend, is that by analyzing the situation within the *set-theoretic* framework of PWS instead of a *logical* framework, one can easily overlook the intrinsic inconsistency of **K** and, instead, rely upon naive intuitions and conclude Kaplan's argument poses a genuine paradox. The fact that the logical analysis we have provided extends to Kaplan's own framework shows that the problem does not have anything to do with PWS *per se* — given that **K** is internally inconsistent, a "paradox" would be expected. But, at the least, an overreliance upon that framework — notably, the assumption that there is a set of all worlds and the identification of propositions with sets of worlds — led to a misdiagnosis regarding the true source of the problem.

5 *PoM* Paradoxes Redux: Whither Set Theory?

Our analysis has defused the Kaplan paradox by revealing the intrinsic, if not intuitively transparent, logical inconsistency of principle **K**. This, in turn, has shown that the "paradox" has nothing essentially to do with either set theory in general or possible world semantics in particular. However, it might be objected that our solution to Russell's *PoM* paradox and its generalization is unsatisfying: it avoids brewing a paradox simply by removing the essential ingredient of set theory. The "solution", then, according to this objection, only comes courtesy of a crippled object theory incapable of supporting a robust theory of sets.

In response to this objection we note that nothing prevents one from taking membership as a primitive and adding the axioms of, say, ZFU (i.e., ZF+*Urelemente*) to object theory. But note that sets of propositions aren't guaranteed to exist in this setting. The reason for this is that propositions in object theory are not in the range of the first-order quantifiers of ZFU. Hence, the axioms of ZFU wouldn't entail the existence of sets of propositions. So this response allows for set theory, but without the problematic sets of propositions needed for the *PoM* paradox.

tions of 'E' means that there is some cost involved in taking **K** to be logically false (*ibid.*). For our part, the *provability* of **K**'s inconsistency from obviously valid logical principles is simply yet another reminder that naive intuitions are unreliable indicators of the logical properties of propositions, particularly when they rest upon such non-logical notions as *set* and *entertainment*.

That noted, there is no obvious reason why one could not formulate a *typed* version of ZFU that quantifies over sets of each logical type of entity in the range of the object-theoretic quantifiers, notably, sets of propositions. Still, however, there would be no *PoM* paradox. For the premises of the *PoM* paradox entail that, for every set S, there is a proposition p_S corresponding uniquely to S. Hence, it is provable in ZFU that premise (1) of the original paradox is false, that there is no set P of all propositions (or alternatively, all truths). For otherwise, by correlating each proposition $p_S \in P$ with its corresponding set S, the members of a subset of P could be correlated one-to-one with the members of the class V of all sets. Hence, by the axiom of Replacement, it would follow that V is a set, which would of course lead quickly to contradiction. Similar reasoning applies to premise (4) of the generalized paradox, i.e., the premise that there is a maximal set of propositions.[42]

This of course raises the possibility of a modification of ZFU — notably, of Replacement — so as to allow proper-class-sized sets as in, for example, the set theories ZFCU$'$ and ZFCU* of Menzel 2014. However, the arguments for the *PoM* paradoxes will not necessarily survive such modifications — in particular, in ZFCU$'$ and ZFCU* they do not.[43] Absent an actual modification of ZFU in which the *PoM* paradoxes can be reconstructed, the mere possibility of such a modification is a toothless speculation.

So the object theorist can meet this objection; there is no reason that ZFU cannot be adjoined to object theory. Of course, this leaves open the question of the consistency of doing so. Object theory itself is known to be consistent relative to a weak fragment of Z set theory, as models by

[42]The idea that the incompatibility of ZFU with set-theoretic worlds provides a satisfying dissolution of the *PoM*-style paradoxes is of course not new. See, in particular, Grim 1991, Ch. 4.

[43]Even though ZFCU$'$ allows proper-class-sized sets, it is still a theorem of that theory that there are no sets the size of the entire universe and, hence, given (2) and (3), that there is no set P of all propositions, contrary to premise (1). (Hence, in ZFCU$'$, unlike ZFU (with classes), "proper class sized" is not equivalent to "as large as the universe".) And given premises (5), (6) and (7) it is for the same reason a theorem of ZFCU$'$ that there is no maximal set of propositions, contrary to premise (4). By contrast, while the premises of both *PoM* paradoxes appear to be consistent with ZFCU*, that theory adopts a generalized Powerset axiom that will not do as a substitute for ZFU's Powerset axiom in the arguments for both paradoxes.

Scott, Aczel, and Menzel & Zalta show.[44] The consistency of full ZFU, of course, can only be proved in systems that are even stronger than ZFU. But, given that object theory is consistent (in Z) and the fact that almost no one doubts the consistency of ZFU after over a century of fruitful, unproblematic use, there is no reason to think that adding (some version of) ZFU to object theory would lead to any problems. So, assuming that some combination of set theory and object theory is consistent, it should be clear that the *PoM* paradoxes do not arise. This, we claim, fully addresses the objection.

However, we think we can develop our response to this objection further. The objection presupposes that sets must be assumed among the philosophical primitives of a correct ontology; that without set theory, object-theoretic foundations are somehow compromised or crippled. Such a presupposition seems to underlie Stalnaker's claim, quoted at the outset, that if there are propositions, then there are sets of them. If we can successfully challenge this assumption, and argue that object-theoretic foundations have no need for set theory, then the *PoM* paradoxes might be seen to be based on the mistaken idea that sets are required in a foundational ontology.

We can't hope to develop a full argument here, but the idea is clear in the following observations. Object theory's perspective is that the axioms of set theory need not be taken as fundamental ontological principles, but rather can be analyzed. The object-theoretic analysis allows mathematicians, scientists, linguists, and philosophers to use the language and axioms of set theory for their theoretical and applied purposes, but identifies the sets they use as abstract objects that encode their set-theoretic properties.[45] Although the details would take us too far afield, the object-theoretic analysis doesn't endorse the set-theoretic existence principles *per se* needed to formulate the paradoxes, but rather endorses them prefaced by the operator *In the theory T* (for the relevant theory T).[46] If the set-theoretic principles are consistent, the abstract objects

[44]A model by Dana Scott was presented in Zalta 1983; a model by Peter Aczel was described in Zalta 1999, and another model was developed in Menzel and Zalta 2014.

[45]This object-theoretic analysis of the language and axioms of set theory has been developed in Zalta 2000b, 2006, Linsky & Zalta 1995, 2006, Bueno & Zalta 2005, and Nodelman & Zalta 2014.

[46]The object-theoretic analysis is that the axioms and theorems of mathematical theories are not simply true, but have a true reading and a false reading. They are false as bald

that encode set-theoretic properties encode consistent properties. If the set-theoretic principles are inconsistent, say because they include principles that lead to the problematic maximal sets of propositions, then the abstract objects that encode set-theoretic properties encode inconsistent properties. The bottom line is that the object-theoretic analysis doesn't endorse any objects that *exemplify* the properties of the problematic maximal sets of propositions.

Thus, if the object-theoretic analysis of set theory is correct, there is no longer any philosophical motivation for taking the axioms of set theory as basic, or accepting sets as *sui generis* objects of our ontology (as opposed to abstract objects that encode their set-theoretic properties). Such an analysis is sufficient for theoretical needs, and avoids mathematical objects as philosophically unarticulated primitives in our ontology. Notably, we don't need sets of propositions to reconstruct possible worlds or situations. Maximal, consistent sets of propositions are unnecessary given that we have possible worlds (in the object-theoretic sense defined in Section 2) that are are maximal and consistent. Arbitrary sets of propositions are also unnecessary, given that we have situations (also defined in Section 2) that are governed by the principle: for any condition φ on propositions p, there is a situation that encodes just the properties F of the form $[\lambda y\ p]$ constructed out of propositions satisfying φ (Zalta 1993, Theorem 1). So not only does object theory allow us to theorize about worlds and propositions without sets, but gives us a set-free background ontology of abstract objects within which sets can be identified.

6 Conclusion

We have examined a number of paradoxes from the perspective of object theory. We have argued that object theory provides a framework in which (a) restrictions that avoid object-theoretic versions of the classical Russell paradox are philosophically justified and (b) worlds and propositions are robustly characterized without invoking sets. We have argued further that defining worlds as sets of propositions is the source of the

statements of fact, but true when read as prefixed by the *In the theory T* operator. Thus, to take an example, $\varnothing_{\text{ZF}}$ (the null set of ZF) is identified as the abstract object that encodes exactly the properties F such that *in ZF*, $F\varnothing_{\text{ZF}}$. Consequently, it becomes a theorem that $\varnothing_{\text{ZF}}$ encodes a property F if and only if *in ZF*, $F\varnothing_{\text{ZF}}$.

PoM paradoxes and, hence, no similar paradoxes arise in object theory. Moreover, we have shown that, even if set theory is added to object theory, the falsity of at least one premise of each of the *PoM* paradoxes falls out as a theorem, further highlighting the volatility of set-theoretic definitions of worlds. Finally, regarding Kaplan's "paradox", while no comparable problem arises for possible world semantics when propositions are defined as sets of worlds, we have argued that such definitions obscure the logical falsity of principle **K** and, as a consequence, have led to a (mis)perception of paradox where none actually exists.

The pernicious effects of hitching one's modal wagon to set theory should thus be clear. Without invoking sets, object theory offers an integrated account of worlds and propositions that assigns them clear existence and identity conditions without any threat of paradox. In the end, worlds and propositions can be set free.

Bibliography

Adams, R., 1974, "Theories of Actuality", *Noûs*, 8(3): 211–231.

———, 1981, "Actualism and Thisness", *Synthese*, 49: 3–41.

Anderson, C. A., 2009, "The Lesson of Kaplan's Paradox about Possible World Semantics", in J. Almog and P. Leonardi (eds.), *The Philosophy of David Kaplan*, New York: Oxford University Press, 85–92.

Boolos, G., 1971, "The Iterative Conception of Set", *The Journal of Philosophy*, 68(8): 215–231.

———, 1987, "The Consistency of Frege's *Foundations of Arithmetic*", in *On Being and Saying*, J. Thomson (ed.), Cambridge, MA: MIT Press, 3–20; reprinted in Boolos 1998, 183–201.

———, 1998, *Logic, Logic, and Logic*, Cambridge, MA: Harvard University Press.

Bringsjord, S., 1985, "Are There Set Theoretic Possible Worlds?", *Analysis*, 45(1): 64.

Bueno, O., and E. Zalta, 2005, "A Nominalist's Dilemma and its Solution", *Philosophia Mathematica*, 13: 294–307.

Chihara, C., 1998, *The Worlds of Possibility*, Oxford: Clarendon Press.

Clark, R., 1978, "Not Every Object of Thought Has Being: A Paradox in Naive Predication Theory", *Noûs*, 12: 181–188.

Davies, M., 1981, *Meaning, Quantification, Necessity: Themes in Philosophical Logic*, London: Routledge & Kegan Paul.

Frege, G., 1891, "Function and Concept", in P. Geach and M. Black (eds. and trans.), *Translations from the Philosophical Writings of Gottlob Frege*, Oxford: Blackwell, 1980 (3rd edition), 21–41.

Grim, P., 1986, "On Sets and Worlds: A Reply to Menzel", *Analysis*, 46(4): 186–191.

———, 1991, *The Incomplete Universe*, Cambridge, MA: The MIT Press.

Kanamori, A., 1996, "The Mathematical Development of Set Theory from Cantor to Cohen", *The Bulletin of Symbolic Logic*, 2(1): 1–71.

———, 2004, "Zermelo and Set Theory", *The Bulletin of Symbolic Logic*, 10(4): 487–553.

Kaplan, D., 1995, "A Problem in Possible World Semantics", in W. Sinnott-Armstrong, D. Raffman, and N. Asher (eds.), *Modality, Morality and Belief: Essays in Honor of Ruth Barcan Marcus*, Cambridge: Cambridge University Press, 41–52.

Lewis, D., 1970, "General Semantics", *Synthese*, 22(1–2): 18–67.

———, 1986, *On the Plurality of Worlds*, Oxford: Basil Blackwell.

Linsky, B., and E. Zalta, 1995, "Naturalized Platonism vs. Platonized Naturalism", *The Journal of Philosophy*, 92(10): 525–555.

———, 2006, "What is Neologicism?", *The Bulletin of Symbolic Logic*, 12(1): 60–99.

Lycan, W., and S. Shapiro, 1986, "Actuality and Essence", *Midwest Studies in Philosophy*, 11: 343–377.

McMichael, A., and E. Zalta, 1980, "An Alternative Theory of Nonexistent Objects", *Journal of Philosophical Logic*, 9: 297–313.

Menzel, C., 1986, "On Set Theoretic Possible Worlds", *Analysis*, 46(2): 68–72.

———, 2008, "Actualism", *The Stanford Encyclopedia of Philosophy*, Winter 2008 Edition, E. N. Zalta (ed.), URL = `http://plato.stanford.edu/archives/win2008/entries/actualism/`.

———, 2012, "Sets and Worlds Again", *Analysis*, 72(2): 304–309.

———, 2014, "Wide Sets, ZFCU, and the Iterative Conception", *The Journal of Philosophy*, 111(2): 57–83.

Menzel, C., and E. Zalta, 2014, "The Fundamental Theorem of World Theory", *Journal of Philosophical Logic*, 43/2: 333–363, doi: 10.1007/s10992-012-9265-z.

Montague, R., 1974, *Formal Philosophy*, New Haven: Yale University Press.

Nodelman, U., and E. Zalta, 2014, "Foundations for Mathematical Structuralism," *Mind*, 123/489: 39–78.

Plantinga, A., 1974, *The Nature of Necessity*, Oxford: Clarendon Press.

Pollock, J., 1984, *The Foundations of Philosophical Semantics*, Princeton: Princeton University Press.

Rapaport, W., 1978, "Meinongian Theories and a Russellian Paradox", *Noûs*, 12: 153–180.

Russell, B., 1903, *The Principles of Mathematics*, New York: W. W. Norton & Co.

Sider, T., 2002, "The Ersatz Pluriverse", *The Journal of Philosophy*, 99(6): 279–315.

Stalnaker, R., 1976, "Possible Worlds", *Noûs*, 10: 65–75.

———, 1984, *Inquiry*, Cambridge, MA: Bradford/MIT Press.

———, 2003, *Ways a World Might Be: Metaphysical and Anti-Metaphysical Essays*, Oxford: Clarendon Press.

Stalnaker, R. and D. Oderberg, 2009, "Modality and Possible Worlds", in J. Kim, E. Sosa, and G. Rosenkrantz (eds.), *A Companion to Metaphysics*, second edition, Oxford: Blackwell, 45–55.

Zalta, E., 1983, *Abstract Objects: An Introduction to Axiomatic Metaphysics*, Dordrecht: D. Reidel.

———, 1993, "Twenty-Five Basic Theorems in Situation and World Theory", *Journal of Philosophical Logic*, 22: 385–428.

———, 1999, "Natural Cardinals and Natural Numbers as Abstract Objects: A Partial Reconstruction of Frege's *Grundgesetze* in Object Theory", *Journal of Philosophical Logic*, 28(6): 619–660.

———, 2000a, "The Road Between Pretense Theory and Object Theory", in *Empty Names, Fiction, and the Puzzles of Non-Existence*, A. Everett and T. Hofweber (eds.), Stanford: CSLI Publications, 117–147.

———, 2000b, "Neo-Logicism? An Ontological Reduction of Mathematics to Metaphysics", *Erkenntnis*, 53 (1–2): 219–265.

———, 2006, "Essence and Modality", *Mind*, 115 (459): 659–693.

Zermelo, E., 1908, "Untersuchungen über die Grundlagen der Mengenlehre I", *Mathematische Annalen*, 65: 261–281. Translated as "Investigations in the Foundations of Set Theory I" in J. van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931*, Cambridge, MA: Harvard University Press, 199–215.

———, 1930, "Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre", *Fundamenta Mathematicae*, 16: 29–47.