A NEGLECTED RESOLUTION OF RUSSELL'S PARADOX OF PROPOSITIONS

GABRIEL UZQUIANO

University of Southern California and University of St Andrews

Abstract.   Bertrand Russell offered an influential paradox of propositions in Appendix B of *The Principles of Mathematics*, but there is little agreement as to what to conclude from it. We suggest that Russell's paradox is best regarded as a limitative result on propositional granularity. Some propositions are, on pain of contradiction, unable to discriminate between classes with different members: whatever they predicate of one, they predicate of the other. When accepted, this remarkable fact should cast some doubt upon some of the uses to which modern descendants of Russell's paradox of propositions have been put in recent literature.

Bertrand Russell offered an influential paradox of propositions in Appendix B of *The Principles of Mathematics*, but there is still little agreement as to what exactly should be concluded from it. The paradox has been turned into an argument for the conclusion that there is no set of all propositions Menzel (2012), and, more recently, into an argument that, on pain of contradiction, some propositions are not members of any classes Deutsch (2014). Moreover, closely related arguments have traditionally been thought to establish that there is no set of all worlds Davies (1981); that there is no set of all truths Grim (1984); that there are no maximal states of affairs Bringsjord (1985); that there is no set of all possible states of affairs Chihara (1998); and, more recently, that something is amiss in standard accounts of propositions McGee & Rayo (2000).

The purpose of this article is twofold. One is to argue that the thesis that there is no set of all propositions is ineffectual as a response to Russell's paradox of propositions. The other is to suggest that Russell's paradox is best regarded as a limitative result on propositional granularity. On pain of contradiction, some propositions must conflate classes with different members: whatever they predicate of one class, they predicate of another class with different members. This interpretation, in turn, should cast doubt upon some of the uses to which a family of modern descendants of Russell's paradox of propositions have been put in the literature.

§1. Russell's paradox of propositions   Russell summarizes the argument in Appendix B of (Russell, 1903, .p. 527):

GABRIEL UZQUIANO

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

Let us make Russell's premises completely explicit:

(1) If m is a class of propositions, there is a proposition of the form *every proposition in m is true*.
(2) If m and n are different classes of propositions, then the proposition *every proposition in m is true* is different from the proposition *every proposition in n is true*.
(3) There is a class w of all and only propositions of the form *every proposition in m is true*, for some class of propositions m to which the proposition does not belong.

Russell's paradox of propositions is the observation that a contradiction follows from premises (1), (2), and (3).

Russell's argument suggests a reductio of the assumption that there is a set of all propositions. Menzel (2012) has pointed out how, given minimal set-theoretic assumptions, a contradiction immediately follows from (4), (5), and (6):

(4) There is a set u of all propositions.
(5) If s is a subset of u, then there is a proposition p_s associated with s, e.g., the proposition *every proposition in s is true*.
(6) If s and t are different subsets of u, then the proposition p_s is different from the proposition p_t.

The argument merely involves an appeal to the axiom of separation of Zermelo-Fraenkel set theory with choice and ur-elements (ZFCU):

> By separation, given (4), we form $r = \{p \in u : \exists s \subseteq u(p = p_s \wedge p \notin s)\}$.[1] Since $r \subseteq u$, by (5), there is a proposition p_r associated with r. Now, on the one hand, if $p_r \notin r$, then p_r is not an element of a set to which it is associated, namely r, whence $p_r \in r$. On the other hand, if $p_r \in r$, then, by definition of r, there is some $s \subseteq u$ such that $p_r = p_s$ and $p_r \notin s$. Since $p_r \in r$, $r \neq s$, and, by (6), $p_r \neq p_s$. So, $p_r \in r$ iff $p_r \notin r$.

The conclusion that there is no set of all propositions appears to be inescapable.[2]

---

[1] This observation corresponds to premise (3) in Russell's original argument.
[2] A less economical argument invokes Cantor's basic result that there is no on-one function from the power set of a set a into a: if there is a set u of all propositions,

But this cannot the end of the matter. Even if there is no set of all propositions, we can still speak of all propositions in the plural. We can ask whether, in line with Russell's first premise, given a plurality of propositions, there is a proposition associated with them.[3] And the question still arises whether, in line with Russell's second premise, different pluralities of propositions have different propositions associated with them. There is even a plural variant of Russell's third premise according to which there is a plurality of all and only those propositions associated with a plurality of propositions of which they are not one. So, we have all the ingredients we need for a plural version of Russell's paradox of propositions. This is, in fact, the thrust of McGee & Rayo (2000), which concludes that something is amiss with standard accounts of propositions.[4]

Now, Deutsch (2014) has recently offered a regimentation of Russell's paradox of propositions in a single-sorted extension of Morse-Kelley class theory (MK) from which he concludes that propositions are members of classes. In the remainder of the paper, we explain how to move from a single- to a two-sorted extension of MK, where we develop a different diagnosis of the problem. We operate in a two-sorted framework largely for ease of exposition, since we have in mind a plural interpretation of the formalism: in the case of MK, we read quantification over classes as plural quantification over sets, and we take a set to be a member of a class of sets if and only if it is one of them.[5] To speak of the class of all sets is to speak of *them* in the plural. We write, as usual, that a class is *proper* if and only if it fails to form a set.

We need to attend to one last item of business before we begin our discussion. One may initially take the claim that there is no set of all propositions to be in tension with the iterative conception of set on which sets are "formed" in stages of a procession of cumulative stages from an initial domain of ur-elements. Since propositions are presumably ur-elements, they are all available for collection at the very first stage of the iteration. So, one may reasonably expect the formation of a set of all propositions—and, further, a set of all ur-elements—to occur at the very first stage of the hierarchy alongside the formation of the empty set.

Some expositions of ZFCU simply take it for granted that there is a set of all ur-elements.[6] There is, however, no technical difficulty associated with the development

---

then, by the power set axiom, there is a set $\mathcal{P}u$ of all subsets of u. But (5) and (6), if true, will generate a one-one function from $\mathcal{P}u$ to u, contradicting Cantor's observation.

[3] It is sometimes convenient to use the singular locution 'plurality' in order to refer to some objects in the plural. The fact that this locution appears to yield grammatically singular expressions such as 'the plurality of all propositions' should not be taken to indicate that the expression refers in the singular to some setlike object to which all propositions belong. All uses of the locution 'plurality' are to be officially eliminated in favor of plural idioms.

[4] Notice, incidentally, that, in order to deal with the plural counterpart of Russell's paradox of propositions, the proposal outlined in Deutsch (2014), for example, would require one to deny that every proposition can enter into the *one of* relation.

[5] Cartwright (2001) and Uzquiano (2003) look at the plural interpretation of impredicative theories of classes in more detail.

[6] Jech (2013), for example, supplements the language of ZFC with a new constant symbol for a set of all ur-elements.

of "impure" set theory without such a set.[7] Indeed, some developments of impure set theory explicitly add the requirement that there is no set of all ur-elements for the sake of mathematical convenience.[8] Still, there may appear to be a tension between the conclusion of Russell's paradox of propositions and the iterative conception. To the extent to which the iterative conception provides the primary motivation for most of the axioms of ZFCU, this tension should be a source of some concern.

There are two main options at this point. One is to develop an alternative to ZFCU designed to accommodate the existence of a set of all ur-elements, and the other is to opt for a more liberal interpretation of the iterative conception on which ur-elements are no longer required to form a set. Menzel (2014) has recently explored the first course of action in great detail. The proposed modification of ZFCU is intended to accommodate the existence of "wide sets", which, while formed at low stages of the cumulative hierarchy, are not "mathematically determinable", since they are too numerous to be measured by any cardinal number. The axioms of replacement and power set, in particular, are appropriately restricted in order to accommodate them.[9] In Menzel's framework, the existence of a set of all propositions is an immediate consequence of the axiom of separation, and we must look elsewhere for an appropriate response to the set-theoretic formulation of Russell's paradox. In fact, much of our discussion carries over to this case as the suggestion that we view the inconsistency of (4), (5), and (6) as a reductio of (6), namely, the claim that different sets of propositions have different propositions associated with them.

One may alternatively confine the iterative conception of set to the thought that sets are formed in stages of a cumulative hierarchy and argue that the existence of a set of all ur-elements is, in fact, not mandated by it.[10] Zermelo (1930) described the cumulative hierarchy of models of ZFCU, which he called "normal domains", which are completely determined by (i) the "width" of the basis of ur-elements and (ii) the "height" given by the supremum of ordinal numbers represented in the model. Sets are formed in ranks of a cumulative hierarchy in each normal domain, but there is some evidence that the existence of a set of ur-elements is not required by the construction.[11] There are, for example, normal domains in which the "height" of the cumulative hierarchy is matched by its "width". One way to accommodate

---

[7] In Barwise (1975), KPU is Kripke-Platek set theory with ur-elements and KPU$^+$ is the theory augmented with the axiom that there is a set of all ur-elements.

[8] Two axiomatizations of set theory augmented with the requirement that there is a proper class of ur-elements are given by Barwise & Moss (1996) and Friedman (2004).

[9] Replacement is restricted to the claim that, very roughly, the range $F[a]$ of an operation F on a set a is a set if either a is mathematically determinable or if $F[a]$ is bounded by a rank. The axiom of power set is similarly stated as the claim that for every set a, there is a set of all and only "mathematically determinable" subsets of a. Menzel's proposal requires two more axioms intended to enforce the existence of a partition of the universe into ranks and the thought that only mathematically determinable sets admit an increase in cardinality.

[10] Indeed, Forster (2008) argues that there is nothing the thought that sets are formed in stages that rules out for example, the iterative formation of a set and its complement in a two way cumulative process that is captured by Church-Oswald models in which every set has a complement.

[11] Kanamori (2004) has some discussion of this feature of Zermelo's conception of a "normal domain".

a proper class of ur-elements is to make sure only "mathematically determinable" sets—which can, in fact, be measured by some cardinal number—are formed at a given stage. More formally, if κ is a strongly inaccessible cardinal and U is the basis of ur-elements, set $\mathcal{P}^{<\kappa}(X)$ to be the set of all subsets of X of cardinality strictly less than κ. We now define:

$$U_0 = U;$$
$$U_{\alpha+1} = U_\alpha \cup \mathcal{P}^{<\kappa}(U_\alpha);$$
$$U_\lambda = \bigcup_{\alpha<\lambda} U_\alpha, \qquad\qquad \text{for } \lambda \text{ a limit ordinal.}$$

If we set $M = \langle U_\kappa, \in \restriction U_\kappa \rangle$ and we let *Set* denote $U_\kappa - U$, we obtain a model of ZFCU in which there is no set of all ur-elements. In fact, M is a model in which no rank of the form U_α $(\alpha < \kappa)$ forms a set. But there is still a clear sense in which sets are formed in stages of a cumulative hierarchy.

§2. Russell's paradox in an impredicative theory of classes   We set out to regiment Russell's paradox of propositions in a suitable extension of Morse-Kelley class theory (MK). As presented in the Appendix to Kelley (1975), Morse-Kelley set theory is formulated in a single-sorted language with uppercase variables $X, Y, Z, ...$ and two primitive non-logical symbols: a two-place predicate, $\in$, and a class abstract operator, $\{ : \}$, which combines a variable X and a formula $A(X)$ into a singular term $\{X : A(X)\}$. Following Kelley, we introduce a one-place predicate MX, read: "X is a *member*," to abbreviate: $\exists Y\ X \in Y$. And we may now stipulate that classes are governed by an Axiom of Extent and a Classification Schema:

$$\forall X \forall Y (X = Y \leftrightarrow \forall Z(Z \in X \leftrightarrow Z \in Y)) \qquad\qquad (\textit{Extent})$$

$$\forall X(X \in \{Y : A(Y)\} \leftrightarrow (MX \wedge A(X))) \qquad\qquad (\textit{Classification})$$

The first axiom tells us that classes are extensional, while the second tells us that, given a condition $A(X)$, there is a class of all and only *members* such that $A(X)$. In addition to these axioms, Kelley's axiomatization has counterparts of axioms of ZF such as union, power set, replacement, infinity, and foundation as well as an axiom of global choice according to which there is a functional relation F selecting exactly one element from each non-empty set, e.g., if x is non-empty, then $F(x) \in x$.

In what follows, we make two adjustments to Kelley's axiomatization. First, we let the uppercase variables range over propositions as well as classes. So, we expand the language with primitive predicate, ΠX, read: "X is a proposition", and a primitive functional symbol $\pi(C)$, which combines with a class term C to yield a singular term $\pi(C)$, read: "the proposition that every proposition in C is true."[12] We explicitly lay down an auxiliary assumption to make sure that π is, in fact, an operation that maps a class of propositions into a proposition:

$$\forall X\ \Pi\pi(X) \qquad\qquad\qquad (\Pi 1)$$

---

[12] There is nothing special about the interpretation Russell had in mind. We could have interpreted $\pi(C)$ to read: "the proposition that C is consistent"; "the proposition that C is self-identical"; etc.

Since, like Russell, we conceive of propositions as *ur*-elements, we explicitly record the auxiliary assumption that propositions are *members* by means of another axiom:

$$\forall X(\Pi X \to MX) \qquad\qquad (\Pi2)$$

This will, in fact, be the crucial difference between our approach and the solution advocated by Deutsch (2014), which, in the face of paradox, rejects ($\Pi2$). To bring out the cost we associate with this view, we now explain how to move from the single-sorted formulation of our extension of MK to a two-sorted formulation in which reference to classes is analyzed as plural reference to sets and propositions. The two-sorted formulation of the framework is similar the one given in Fraenkel et al. (1973), but it is important to be clear on what exactly is involved in the transition to a two-sorted formulation.

   We now follow Mendelson (1997) and augment the language of MKU with lower-case variables $x, y, z, ...$, which are stipulated to range over *members*. In other words, we take a formula of the form $\forall x A(x)$ to abbreviate: $\forall X(MX \to A(X))$, and we take $\exists x A(x)$ to abbreviate: $\exists X(MX \land A(X))$.[13] Finally, we take $\{x : A(x)\}$ to abbreviate: $\{X : MX \land A(X)\}$. Given Kelley's axioms of extent and classification, the stipulation immediately yields the following two important consequences:

$$\forall X\forall Y(X = Y \leftrightarrow \forall z(z \in X \leftrightarrow z \in Y)) \qquad (\textit{Class Extensionality})$$

$$\forall x(x \in \{y : A(y)\} \leftrightarrow A(x)) \qquad (\textit{Class Abstraction})$$

By existential generalization, *Class Abstraction* entails an impredicative form of class comprehension commonly associated with MK:

$$\exists X\forall x(x \in X \leftrightarrow A(x)), \qquad (\textit{Impredicative Class Comprehension})$$

where X does not occur freely in $A(x)$. Notice that we allow for class variables to appear bound in A. This is, in fact, the key difference between MK and other predicative theories of classes such as von Neumann-Bernays-Gödel class theory (NBG). Unfortunately, we will not be able to make do with predicative instances of class comprehension, if we want to reproduce Russell's derivation of the contradiction.

   Let me emphasize that the move to a two-sorted formulation of the extension of MK is not motivated by purely technical considerations, but rather by the fact that it invites a plural interpretation on which lowercase variables are taken to range over sets and uppercase variables are interpreted as plural variables over sets. While an atomic formula of the form $x \in y$ is read: "x is an element of y", an atomic formula of the form $x \in Y$ is read: "x is one of Y". On this interpretation, to claim that some propositions are never members of classes is to claim that some propositions never stand in the *one of* relation to some propositions. And this is exactly why we find the move made by Deutsch (2014) very costly.

   We are finally in a position to regiment each of Russell's premises. The first premise is the claim that there is, for every class of propositions X, a proposition $\pi(X)$ to the effect that every proposition in X is true. In symbols:

$$\forall X(\forall x(x \in X \to \Pi x) \to \exists x\ x = \pi(X)) \qquad (R1)$$

Notice that the choice of regimentation is not completely innocent. Given ($\Pi1$) and ($\Pi2$) above, once π is introduced as a functional symbol in a classical framework,

---

[13] In each case, X should be the first variable not occurring in $A(x)$.

then the formula $\exists x \ x = \pi(X)$ automatically becomes a theorem. No matter, by itself, $(R1)$ appears to be a harmless addition to the axioms of the present extension of MK: to satisfy $(R1)$ in a model of our extension of MK, we need only posit the existence of two propositions: $\top, \bot$, read: "the True" and "the False". The extension of the predicate Π will be given by the set: $\{\top, \bot\}$, and the denotation of $\pi(C)$ will generally correspond to the truth value of the proposition that every proposition in C is true.

The identification of materially equivalent propositions would hardly be faithful to Russell's intentions. The thrust of his second premise is to make sure that different propositions are associated by π to different classes of propositions:

$$\forall X \forall Y (\forall x (x \in X \to \Pi x) \wedge \forall x (x \in Y \to \Pi x) \to (X \neq Y \to \pi(X) \neq \pi(Y))) \quad (R2)$$

But as Russell explicitly observed, the premise would quickly be falsified if one were to identify materially equivalent propositions:

> In order to deal with this contradiction, it is desirable to reopen the question of the identity of equivalent propositional functions and of the nature of the logical product of two propositions. If m be a class of propositions, their logical product is the proposition "every m is true", which I shall denote $\wedge' m$. If we now consider the logical product of the class of propositions composed of m together with $\wedge' m$, this is equivalent to "Every m is true and every m is true", i.e., to "every m is true", i.e., to $\wedge' m$. [...] Thus if we identify equivalent propositional functions ($\wedge' m$ being a propositional function of m), the proof of the above contradiction fails, since every proposition of the form $\wedge' m$ is the logical product of a class of which it is a member and of a class of which it is not a member.

There is not much solace in Russell's observation:

> But such an escape is in reality impracticable, for it is quite self-evident that equivalent propositional functions are not identical. Who will maintain, for example, that "x is an even prime other than 2" is identical with "x is one of Charles II's wise deeds or foolish sayings"? Yet those are equivalent if a well-known epitaph is to be credited.

But we do better if we opt for necessary equivalence as a more reasonable standard for the individuation of propositions. Consider, for example, a coarse-grained account of propositions on which each proposition amounts to a class of possible worlds in the tradition of David Lewis and Robert Stalnaker. They have a ready response to Russell's paradox of propositions: $(R2)$, in particular, is simply false: given two different propositions p and q, $\{p, \neg p\}$ and $\{q, \neg q\}$ are classes with different members. And yet, the proposition *every proposition in $\{p, \neg p\}$ is true* is necessarily equivalent to the proposition that *every proposition in $\{q, \neg q\}$ is true*.

Russell's second premise requires a fine-grained account of propositions on which one is allowed to distinguish two propositions on the grounds that they involve different predications: whatever a proposition predicates of a given class, it cannot be identical to a proposition that predicates something of a class with different members. To the extent to which the propositions involve different predications,

they must be different.[14] There is no denying that this is a prima facie very attractive thought, and its great appeal is part of the reason why Russell's paradox of propositions raises a serious difficulty.

The final premise Russell requires follows directly from the instance of *Class Abstraction* corresponding to $\{x : \exists Y(\forall y(y \in Y \to \Pi y) \wedge x = \pi(Y) \wedge x \notin Y))\}$ by existential generalization:

$$\exists X \forall x(x \in X \leftrightarrow \exists Y(\forall y(y \in Y \to \Pi y) \wedge x = \pi(Y) \wedge x \notin Y)) \qquad (R3)$$

A contradiction follows directly from $(R1)$, $(R2)$, and $(R3)$ by Russell's argument. By $(R3)$, there is a class R to which a given x belongs iff $x = \pi(Y)$ for some class Y to which x does not belong. By $(R1)$, there is a proposition r identical to $\pi(R)$. Since all propositions are members, they lie in the range of the lowercase variables. Now, if $\pi(R) \notin R$, then $\pi(R)$ satisfies the condition associated to R and $\pi(R) \in R$. But if $\pi(R) \in R$, then given our definition of R, there is *some* class of propositions Y such that $\pi(R)$ is $\pi(Y)$ and $\pi(R) \notin Y$. Since $\pi(R) \in R$, $R \neq Y$, which contradicts $(R2)$.

Just to be clear, given the choice of regimentation, $(R1)$ follows from $(\Pi 1)$ and $(\Pi 2)$ by classical logic. And $(R3)$ is just an instance of impredicative class comprehension. So, maybe $(R2)$ should be subjected to closer scrutiny.

§3. What is the moral of Russell's paradox of propositions?   One may hope to bring Cantor's theorem to bear to the question of whether there is a set of all propositions. Cantor's theorem is often glossed as the claim that a set has strictly more subsets than elements. But if there is a set u of all propositions, (5) and (6) will induce a one-one function from the power set of u into u, which appears to contradict Cantor's theorem. One drawback of this argument, when compared to the argument given at the outset, is the implicit appeal to the power set axiom, which plays no role in the original reductio. Still, one may wonder whether an appropriate generalization of Cantor's theorem for classes can shed some light on the class-theoretic version of the problem.

Cantor's theorem does not speak to the question of whether a class has more subclasses than members, but Bernays (1942) explained how to prove a generalization of Cantor's theorem for classes. But while Cantor's theorem is concerned with an explicit cardinality comparison between a given set a and its powerset $\mathcal{P}a$, one difficulty at this point is that, unlike sets, classes are never members, which means that there is never opportunity to collect all subclasses of a class A into another class. No matter, one may still hope to generalize Cantor's basic result as the claim that, given some class A, there is no "function F from A onto all subclasses of A." But this raises another challenge, since it is not obvious how to make sense of a "class-valued" function, which maps members of a class A to subclasses of A.[15]

---

[14] Beall (2000) makes a similar point in response to Grim's argument for the conclusion that there is no set of all truths as developed in Grim (1984).

[15] Some definitions:

  - A *relational class* R is a class of ordered pairs.
  - A *functional class* F is a relational class such that for all x, y, z, $\langle x, y \rangle \in F$ and $\langle x, z \rangle \in F$, then $y = z$. We write $F(x) = y$ to abbreviate: $\langle x, y \rangle \in f$.

In the context of a predicative theory of classes, Bernays (1942) explained how to simulate a "class-valued" function from members of a given class A to subclasses of A by means of a binary relational class on A.[16] In particular, we can take a binary relational class R on A to map a member a of A to the class $\{x \in A : \langle a, x \rangle \in R\}$, which is itself a subclass of A. We can even write $R(a) = B$ to abbreviate: $\forall x(\langle a, x \rangle \in R \leftrightarrow x \in B)$. Bernays' generalization of Cantor's theorem for classes is merely the observation that no binary relational class on A simulates a *class-valued function* from members of A *onto* subclasses of A: given a binary relational class R on A, there is, on pain of contradiction, a subclass D of A such that *no* member d of A is such that $R(d) = D$.

THEOREM 3.1. (Bernays' Theorem) (NBG) *No binary relational class R on a class A simulates a class-valued function from members of A onto subclasses of A.*

Proof. Let R be a binary relation on A, and consider the class $D = \{x \in a : \langle x, x \rangle \notin R\}$, which exists by predicative comprehension.[17] Now, there is no member d of A such that $R(d) = D$. Otherwise, if a member d of A is such that $R(d) = D$, then $\langle d, d \rangle \in R$ iff $d \in D$ iff $\langle d, d \rangle \notin R$, which leads to contradiction. □

To bring Bernays' result to bear on the question of whether $(R1)$ and $(R2)$ are satisfiable, we need to move beyond the resources of predicative comprehension. First, let me mention a suggestive result: No operation π from subclasses of a given class A to members of A gives us a one-one map from subclasses of A to members of A:

COROLLARY 3.2. (MK) *If π is an operation from subclasses of a given class A to members of A, then there are two different subclasses B and C of A such that $\pi(B) = \pi(C)$.*

Proof. Define R to be a relational class on A such that for all $x, y \in A$, $\langle x, y \rangle \in R$ if and only if either (i) there is *exactly one* class B of members of A such that $x = \pi(B)$ and $y \in B$ or else (ii) $y \in A$. R is supposed to simulate a class-valued function that maps a member x of A to the *unique* subclass B of A such that $\pi(B) = x$, *if such a class exists*, and to A, otherwise. The existence of R is sanctioned by impredicative class comprehension. Now, by Bernays' theorem, there is some class B of members of A such that no member b of A is such that $R(b) \neq B$. It follows that $\pi(B) = \pi(C)$ for some other class C of members of A. Otherwise, since $\pi(B) \in A$, we would have $R(\pi(B)) = B$. □

Apart from the use of impredicative comprehension, it may be helpful to note that the argument gives no clue as to the identity of the subclasses B and C of A for which $\pi(B) = \pi(C)$. We can do better if we adapt Russell's original argument into an alternative proof of Corollary 3.2.

Proof. Set R to be the class $\{x : \exists Y \subseteq A(x = \pi(Y) \wedge x \notin Y)\}$, which exists by impredicative class comprehension. Given $(\Pi1)$ and $(\Pi2)$, $\pi(R)$ is a member, which

---

[16] R is *relational class R on A* if, and only if, R is a relational class and $R \subseteq A \times A$, where $A \times A$ is the class of ordered pairs of members of A.

[17] The axiom schema of predicative comprehension of NBG generates the existence of a class of exactly the sets that satisfy an open formula $A(x)$ of the language of NBG in which no bound class variables occur.

lies in the range of the lowercase variables. Now, observe that $\pi(R) \in R$, since, if $\pi(R) \notin R$, then, by definition of R, $\pi(R) \in R$. It follows that there is some $Y \subseteq R$ such that $\pi(R) = \pi(Y)$, even though $\pi(R) \notin Y$. Since $\pi(R) \in R$, we have that $R \neq Y$.   □

Russell's argument constitutes an improvement over the original proof by providing a partial specification of a counterexample to the claim that π is a one-one operation from subclasses of A to members of A: one class, namely, R, but not the other, is explicitly defined in terms of π. So, the question naturally arises whether there is a more informative argument whereby one is a position to explcitly define in terms of π two different subclasses B and C of A for which $\pi(B) = \pi(C)$.

This is the class-theoretic version of the question of whether there are "constructive" proofs of the Cantorian proposition that there is no one-one function from $\mathcal{P}a$ into a given set a, i.e., proofs which, given a function f from $\mathcal{P}a$ to a, specify two different subsets b and c of a for which $f(b) = f(c)$. This question has, in fact, been discussed by Boolos (1997) and Kanamori (2004).[18] Here is a first pass at a "constructive" version of the Cantorian result.

THEOREM 3.3. *Given a function $f : \mathcal{P}a \to a$, there are two different subsets b and c of a definable from f such that $f(b) = f(c)$.*

Proof. Define a functional class F from the ordinals into a by transfinite recursion on the ordinals:

$$F(\alpha) = f(\{F(\beta) : \beta < \alpha\}).^{19}$$

By Hartogs' theorem, there is a least ordinal β such that $F(\beta) = F(\gamma)$ for some $\beta < \gamma$.[20] We set $b = \{F(\alpha) : \alpha < \beta\}$ and $c = \{F(\alpha) : \alpha < \gamma\}$. Then:

- $f(b) = f(c)$. (This is because, by choice of b and c, $f(b) = F(\beta) = F(\gamma) = f(c)$.)
- $b \neq c$. (This is because $f(c) \notin b$ but $f(c) \in c$. For recall that $f(c) = F(\beta)$, for $\beta < \gamma$, but $f(c) \neq F(\alpha)$, for $\alpha < \beta$.)

□

While quite simple, the proof just now given is not without shortcomings. One drawback of the argument is that it appears to rely on the prior development of the theory of ordinals, which makes extensive appeal to the axiom of replacement. But the other, more important problem, is that the proof cannot be generalized for the case of classes. In the set-theoretic case, we rely on the assumption that the range of F is a set in order to be able to exploit Hartogs' theorem to guarantee the existence of some ordinal β such that $F(\beta) = F(\gamma)$ for some $\gamma < \beta$. But in

---

[18] See, for example, Corollary 3 in Kanamori (2004).

[19] The picture is one in which F maps ordinals into elements of a as follows:

- $F(0) = f(\emptyset)$
- $F(1) = f(\{F(0)\}) = f(\{f(\emptyset)\})$
- $F(2) = f(\{F(0), F(1)\}) = f(\{f(\emptyset), f(\{f(\{\emptyset\})\})\})$
- $\cdots$

[20] Hartogs' theorem states, roughly, that if a is a set, some ordinal α is such that no one-one function exists from α to a.

the class-theoretic case in which we are given an operation π from subclasses of a given class A to members of A, there is no reason to take the range of F to be a set, which blocks the application of Hartogs' theorem.

Boolos (1997) and Kanamori (2004) have independently employed a different style of argument for the same conclusion, one which makes no use of replacement and can be generalized to the case of classes. As presented in Kanamori (2004), the argument makes crucial use of a lemma, which he attributes to Zermelo, corresponding, roughly, to the choiceless part of Zermelo's proof of the well-ordering theorem.[21] In what follows, we state a class-theoretic version of the lemma in the framework of MK.[22]

THEOREM 3.4. (MK) *If π is an operation from subclasses of A to members of A, there is a unique subclass M of A and a strict well-ordering R on M such that:*

(i) For all $x \in M$, $\pi(R_x) = x$.
(ii) $\pi(M) = M$.

Proof. Call a subclass M of A a π-*chain* iff there is a strict well-ordering R on M such that for each $x \in M$, $\pi(R_x) = x$. Quite generally, π-chains take the following form:

$$\emptyset, \quad \{\pi(\emptyset)\}, \quad \{\pi(\emptyset), \pi(\{\pi(\emptyset)\})\}, \quad \{\pi(\emptyset), \pi(\{\pi(\emptyset)\}), \pi(\{\pi(\emptyset), \pi(\{\pi(\emptyset)\})\})\}, \quad ...$$

where each of them comes with a unique witnessing strict well-ordering. Following Kanamori (2004), uniqueness follows from the following observation:

---

[21] This is Theorem 1 in Kanamori (2004).
[22] Some definitions:

- The domain of a relational class R, $dom(R)$, is the class $\{x : \exists y \langle x, y \rangle \in R\}$.
- The range of a relational class R, $ran(R)$, is $\{x : \exists y \langle y, x \rangle \in R\}$.
- The field of a relational class R, $field(R)$, is $dom(R) \cup ran(R)$, which exists by predicative comprehension.
- A relational class R is a *strict well-ordering* iff:

    — R is irreflexive on $field(R)$, e.g., for all $x \in field(R)$, $\langle x, x \rangle \notin R$,
    — R is connected on $field(R)$, e.g., for all $x, y \in field(R)$, $\langle x, y \rangle \in R$ or $\langle y, x \rangle \in R$.
    — R is anti-symmetric on $field(R)$, e.g., for all $x, y \in field(R)$, if $\langle x, y \rangle \in R$ and $\langle y, x \rangle \in R$, then $x = y$.
    — R is transitive on $field(R)$, e.g., for all $x, y, z \in field(R)$, if $\langle x, y \rangle \in R$ and $\langle y, z \rangle \in R$, then $\langle x, z \rangle \in R$.
    — If $X \subseteq field(R)$ and $X \neq \emptyset$, then X has an R-minimal member x, e.g., for all $y \in X$, $\langle x, y \rangle \in R$.

- A relational class R is a *strict well-ordering on A* iff R is a strict well-ordering and $R \subseteq A \times A$.
- X is a *initial segment* of R iff $X \subseteq field(R)$ and for all $x \in field(R)$, for all $y \in X$, if $\langle x, y \rangle \in R$, then $x \in X$.
- R_a is the *strict initial segment of R determined by a*, which we define as the largest initial segment of R not containing a, e.g., $\bigcup\{X : X$ is an initial segment of $R \wedge a \notin X\}$.
- If $\langle M, R \rangle$ and $\langle N, S \rangle$ are well-orderings, then $\langle M, R \rangle$ is *similar* to $\langle N, S \rangle$ iff there is a one-one functional class F from M onto N such that for all $x, y \in M$ xRy iff $F(x)SF(y)$.

- If M_1 and $M2$ are two π-chains with corresponding well-orderings R_1 and R_2 on them, then one of $\langle M_1, R_1\rangle$ and $\langle M_2, R_2\rangle$ is an initial segment of the other.

By the comparability of well-orderings, we can assume without loss of generality that $\langle M_1, R_1\rangle$ is similar to an initial segment of $\langle M_2, R_2\rangle$. Let I be a one-one functional relation from M_1 onto some subclass of M_2 where for all $x, y \in M_1$, xR_1y iff $I(x)R_2I(y)$. We argue that I is, in fact, the identity map. For if not, let m be the R_1-minimal member of M for which $m \neq I(m)$. By R_1-minimality of m, we know that $R_{1m} = R_{2I(m)}$. But now:

$$I(m) = \pi(R_{2I(m)}) = \pi(R_{1m}) = m.$$

Contradiction. In particular, if $M_1 = M_2$, then $R_1 = R_2$.

Another consequence of the preceding observation is that the union of two π-chains is itself a π-chain, whose witnessing strict well-ordering is given by the union of the strict well-orderings corresponding to the initial π-chains. Now, let M be the union of all π-chains, i.e., $\bigcup\{N : N$ is a π-chain$\}$, which exists by impredicative comprehension. The union of witnessing well-orderings R for all π-chains will give us a strict well-ordering of M satisfying condition (i) in the statement of the theorem. To check that M also verifies condition (ii), notice that if $\pi(M) \notin M$, then $M \cup \{\pi(M)\}$ would generate another π-chain, contradicting the definition of M as the union of all π-chains. $\qquad\square$

We are finally in a position to state a generalization of Theorem 3.3.

COROLLARY 3.5. (MK) *If π is an operation from subclasses of A to members of A, then there are two subclasses B and C of A definable from π such that $B \neq C$ and $\pi(B) = \pi(C)$.*

Proof. Given π, set M and R as in Theorem 3.4. and let $B = R_{\pi(M)}$ and $C = M$. We have the following:

- $B \neq C$, since $\pi(M) \in C$ but $\pi(M) \notin B$, and
- $\pi(B) = \pi(C)$, since $\pi(R_{\pi(M)}) = \pi(M)$.

$\qquad\square$

The argument, in fact, proves a non-obvious strengthening of the corollary: if π is an operation from subclasses of A to members of A, then there are two *well-orderable* subclasses B and C of A definable from π such that $B \subset C$ and $\pi(B) = \pi(C)$.[23]

The main interest of Corollary 3.5. is that, in the presence of $\Pi1$ and $\Pi2$, it provides a recipe for the construction of an explicit counterexample to ($R2$). Although we had originally flagged ($R2$) for closer scrutiny, we had not yet been in a position to provide a completely explicit illustration of its failure.

Two comments are in order. Notice, first, that the formulation of the counterexample is perfectly general and does not rely on any specific interpretation of π. In particular, we can let $\pi(C)$ denote a simple predication of the form "C is F," e.g., C is consistent. In this respect, $\pi(C)$ resembles a singular proposition like *Beth is a philosopher* and *Scott is a philosopher*. Now, the last two propositions involve two simple predications of two different individuals: one is directly concerned with

---

[23] Call a class B *well-orderable* iff B is the field of some strict well-ordering R.

Beth whereas the other is directly concerned with Scott. Since they involve different predications, we find it almost irresistible to conclude that they are different propositions. On the plural interpretation of the formalism, $\pi(C)$ is not exactly a singular proposition, since, when read "C is consistent", it involves a collective predication of some objects in the plural. However, it is still very difficult to resist the pull of the claim that, when concerned with two classes C and D with different members, *C is consistent* and *D is consistent* involve two different simple predications, which should result in two different propositions. And yet, corollary 3.5. now enables us to construct two classes C and D with different members such that the proposition that C is consistent turns out to be identical to the proposition that D is consistent. So, a simple predication of consistency of C turns out to be tantamount to a simple predication of consistency of D.[24]

The formulation of an explicit counterexample to $(R2)$ does not, by itself, tell us how pervasive the failure of π to map subclasses of A in a one-one fashion into A must be. A glance at the set-theoretic case may be of some help. It follows from the Zermelo-König inequality, which is a theorem of ZFC, that the cofinality of the cardinality of the power set of a is strictly larger than the cardinality of a: if κ is the cardinality of a, then the union of κ sets of cardinality less than 2^κ will still have cardinality strictly less than 2^κ. So, if f is a function from the power set of a onto a, then there is at least one member b of a to which f maps as many subsets of a as there are subsets of a altogether. Given an appropriate form of global choice for classes, a class-theoretic generalization of the Zermelo-König inequality would seem to suggest that, given an operation π from subclasses of A to members of A, one member a of A must be assigned by π to as many subclasses of A as there are subclasses of A altogether.

§4. The significance of Russell's paradox of propositions Russell's paradox of propositions has had many modern descendants. For a taste, consider the argument for the claim that there is no set of all truths given in Grim (1984). As Menzel (2012) observes, the argument is parallel to Russell's paradox of propositions when we substitute "class of truths" for "class of propositions". Grim concludes that there is no set of all truths, but the problem remains when we make allowance for proper classes of propositions:

(7) If T is a class of truths, there is a truth associated with T.[25]

(8) If T and Υ are different classes of truths, then the truth associated with T is different from the truth associated with Υ.

In symbols, let T be a predicate read: "is a truth", and let $\tau(X)$ be a singular term for the proposition associated with X. Given auxiliary assumptions to the effect

---

[24] While certainly difficult to fathom, the present response to Russell's paradox of propositions is perhaps not unprecedented. Theorem 1.3 in Zalta (1999) involves a similar claim in the framework of Zalta's theory of abstract objects: for each property F, there is at least a pair of different abstract objects such that the singular propositions that predicate F of each of them turn out to be identical. Thanks to Christopher Menzel for this reference.

[25] If t is a given truth, then we may consider $t \in$ T, if $t \in$ T, or $t \notin$ T, otherwise.

that all truths are *members* and thus lie in the range of the lowercase variables of the language, we can regiment the two principles as follows:

$$\forall X(\forall x(x \in X \to \mathrm{T}x) \to \exists x\ \tau(X)) \qquad (G1)$$

$$\forall X\forall Y(\forall x(x \in X \to \mathrm{T}x) \wedge \forall x(x \in Y \to \mathrm{T}x) \to (X \neq Y \to \tau(X) \neq \tau(Y))) \quad (G2)$$

But we know that $(G2)$ is unattainable. This is not unlike the response given by Beall (2000), except that we need not rely on the identification of necessarily identical propositions. $(G2)$ is admittedly incompatible with a coarse-grained account of propositions, but the larger point is that $(G2)$ is incoherent: some truths conflate classes of truths with different members; whatever they predicate of one, they must, on pain of contradiction, predicate of the others.

Modern descendants of Russell's paradox have been thought to raise a serious threat for possible worlds accounts of modality, which are generally committed to instances of the following schema:

($\diamond$) It is possible that p iff p is true at some possible world w.

Notice that ($\diamond$) is meant to remain neutral as to what exactly is the nature of possible worlds. Possible worlds have been conceived as maximal consistent classes of propositions Adams (1974), total consistent propositions (Poscript to Prior (1977)), total states of affairs Plantinga (1992) Plantinga (1976), ways the world might be Stalnaker (1976), possible states of the world Kripke (1980), or spatio-temporally disconnected concrete universes Lewis (1986).

The problem requires minimal assumptions. First, we make explicit two common, albeit not universal, presuppositions:

(9) Every proposition p has a negation: $\neg p$.[26]
(10) If p is a proposition and w is a world, then either p is true at w or $\neg p$ is true at w.

Now, call a class of propositions M *maximal* if and only if for every proposition p, M contains either p or the negation of p as a member. And call a class of propositions M *compossible* if, and only if, it is possible that every proposition in M is true. A *maximal compossible* class of propositions is a maximal class of propositions such that it is possible that every one of them is true.

The key observation is that ($\diamond$), given minimal assumptions, appears to commit one to the existence of maximal compossible classes of propositions. Whatever possible worlds may be, if propositions are true at them, then they determine maximal compossible classes of propositions—whether or not one proceeds to define a possible world as a maximal compossible class of propositions:

(11) For each possible world w, there is maximal compossible class of propositions, i.e., the class of propositions that are true at w.

But the existence of maximal compossible classes of propositions appears to be directly threatened by a variant of Russell's paradox of propositions. The paradox, in particular, arises from the inconsistency of the following apparently plausible principles:

---

[26] Adams (1981), for example, takes exception to this claim.

(*C*1) There is a maximal compossible class of propositions M.

(*C*2) For each class C of propositions in M, there is a proposition p_C associated with C, e.g., the proposition that C is a class of propositions.

(*C*3) If C and D are different classes of propositions in M, then the propositions p_C, p_D, $\neg p_C$, and $\neg p_D$ are pairwise different.

(*C*4) There is a class of propositions to which a proposition p belongs if and only if p is neither the proposition p_C associated with a class C of propositions to which p does not belong nor the negation $\neg p_C$, of the proposition p_C associated with a class C of propositions to which p does not belong.

The argument should be familiar by now: given a maximal compossible class M of propositions, by (*C*1), there is, by (*C*4) a subclass R of M to which a proposition p belongs iff p is not p_C for some subclass C of M to which p does not belong or $\neg p_D$ for some subclass D of M to which p does not belong. Now, consider p_R, which exists by (*C*2). If $p_R \notin R$, then, since p_R is different from $\neg p_R$, by definition of R, $p_R \in R$. Contradiction. It follows that $p_R \in R$. So, either $p_R = p_C$ for some class C such that $p_R \notin C$ or $p_R = \neg p_D$ for some class D such that $p_R \notin D$. Either way, since $p_R \in R$, neither $p_R \notin C$ nor $p_R \notin D$, which means that $R \neq C$ or $R \neq D$, which contradicts (*C*3).

   Notice that (*C*3) is no more plausible than (*R*2) was in Russell's original argument. So, by rejecting (*C*3), we seem to block the derivation of the contradiction. Whatever concerns one may have with ($\diamond$), the fact that a contradiction follows from utterly plausible premises should not be one of them.

§5. A desperate move?   The suggestion that, on pain of contradiction, *some* propositions conflate classes with different members difficult to fathom. How could a proposition, in addition, conflate such a vast range of classes of propositions as required by an appropriate generalization of the Zermelo-König inequality? It may be fine to adopt some coarse-grained account of propositions or another, but it seems difficult to accept that the choice is handed down to us by logic alone. Desperate times call for desperate measures. So, one may be tempted by the thesis that whatever propositions may be, they are not objects. For if propositions are not objects, then they cannot be members of classes. Nor can they enter into the *one of* relation to other classes. So, Russell's paradox of propositions never arises.

   The devil is in the details. If propositions are not objects, what exactly are they? In a broadly Fregean framework, one may enforce a distinction between objects, first-level concepts under which objects fall, second-level concepts under which first-level concepts fall, etc. The Fregean hiearchy is supposed to align with a hiearchy of linguistic expressions: names stand for objects, first-level predicates stand for first-level concepts, etc. One proposal at this point might be to identify propositions with the values of zero-place predicate variables. Propositions would, on this view, belong to the same ontological category as Fregean first-level concepts, and they should not be confused with objects. This, I should hasten to add, is *not* Frege's own view, who seemed unfazed by Russell's paradox on the grounds that he could sidestep it by appeal to the distinction between sense and reference.[27]

_____

[27] It is very much an open question whether he should have in fact been much more concerned than he appeared to be. Klement (2001) discusses both Frege's calm reaction

One problem with the suggestion is that it is exposed by a variation on Russell'
paradox framed in terms of predication. In order to regiment the paradox, we may
adopt a Fregean formal language similar to the one described by Bell (1995).[28] For
ease of exposition, we may reserve uppercase letters $X, Y, Z, ...$ for first-level con-
cepts and underline uppercase letters for second-level concepts $\underline{X}, \underline{Y}, \underline{Z}, ...$. Notice
that even if propositions are not objects, they still fall under second-level concepts.
We may, in addition, introduce an identity symbol in order to express a second-
level relation of identity between first-level concepts. Finally, we supplement the
language with a functional term, %, which combines with a second-level predicate
in order to yield a first-level concept term, $\%\underline{X}$, which stands for a zero-place first-
level concept, e.g., the proposition that every first-level concept falling under $\underline{X}$ is
self-identical. The following two principles are inconsistent:

$$\forall \underline{X} \ \exists X \ X = \%\underline{X}) \tag{F1}$$

$$\forall \underline{X} \forall \underline{Y}(\underline{X} \neq \underline{Y} \to \%\underline{X} \neq \%\underline{Y}) \tag{F2}$$

$(F1)$ states that every second-level concept is correlated with a proposition uniquely
associated with it, e.g., the proposition that all its instances are self-identical.
In contrast, $(F2)$ states that different second-level concepts give rise to different
propositions uniquely associated with them. The derivation of the contradiction
should now be familiar.

Impredicative comprehension for second-level concepts yields:

$$\exists \underline{X} \forall X(\underline{X}X \leftrightarrow \exists \underline{Y} \exists Y(Y = \%\underline{Y} \wedge \neg \underline{Y}Y)$$

Let $\underline{R}$ be such a second-level concept. By $(F1)$, $\%\underline{R}$ is a proposition
uniquely associated to it. If $\neg\underline{R}\%\underline{R}$, then the proposition satisfies
the condition associated with the second-level concept and $\underline{R}\%\underline{R}$.
But if $\underline{R}\%\underline{R}$, then there is some second-level concept $\underline{Y}$ such that
$\%\underline{R} = \%\underline{Y}$ but $\neg\underline{Y}\%\underline{R}$, which contradicts $(F2)$.

Even if propositions belong to the same ontological category as first-level con-
cepts, we should not expect them to be able to discriminate between different
second-level concepts. This conclusion is not much more comfortable than the
suggestion that if propositions are objects, then not all of them can discriminate
classes with different members.

The Fregean framework may suggest perhaps a different resolution of the para-
dox. The problem arises only because we take the operation % to assign to each
second-level concepts a zero-place first-level concept. But maybe we could avoid the
difficulty if we make sure that % does not map $n+1$th-level concepts into lower level
zero-place concepts. Maybe propositions are layered in different levels of the Fregean
hierarchy as zero-place concepts of different levels. One way in which the proposal
may be implemented is by means of the inchoate requirement that propositions
concerned with items of level n in the Fregean hierarchy be themselves items of level

---

to the paradox and the question of whether similar difficulties arise for his account of
      sense.

[28]   Bell (1995) takes the framework to be regimented in a many-sorted first-order language,
      but this is not compulsory for present purposes. Instead, one may take the intended
      interpretation of the framework to be given by the Fregean hierarchy of concepts.

at least n in the Fregean hierarchy. While the general thought may seem enough to dispel the risk of paradox, it comes at a high cost: all propositions occur at one level or another of the Fregean hierarchy, which means that no proposition appears to be concerned with all propositions. We seem deprived from the resources to state genuinely general propositions concerned with all the propositions there are.

§6. Conclusion   There is an intense debate in the philosophy of language as to whether propositions are fine or coarse-grained. Even after we discard the identification of materially equivalent propositions as obviously inadequate, we are still left with a wide range of live answers to this question. At one end of the spectrum, we find a coarse-grained view of propositions as classes of possible worlds, one which identifies necessarily equivalent propositions. At the other end, we find a Russellian conception of propositions as complex objects made out of simpler constituents, one on which propositions are, one and all, fine-grained: in particular, when a proposition predicates something of a class, then it contains the very class as one of the constituents of the proposition, and we have every reason to distinguish propositions with different constituents.

We have suggested that Russell's paradox of propositions is best viewed as a constraint on propositional granularity: no matter how fine-grained one may like propositions to be, they cannot all be as fine-grained as to be able to discriminate between any two classes. This is, by itself, hardly a vindication of the identification of necessarily equivalent propositions, but it appears to put pressure on the Russellian conception of propositions to the extent to which one is no longer in a position to hold that propositions, one and all, are fine grained. However one seeks to develop a fine-grained view of propositions, one must make room for the identification of propositions predicating one and the same thing of two different classes.This observation provides a neglected response to a family of modern descendants of Russell's paradox. One may, to be sure, be surprised to learn that the fact that *some* propositions are coarse-grained is handed down to us by logic alone, but once we come to terms with it, we have the resources to respond to a wide variety of related problems discussed in recent literature.

§7. Acknowledgments   I'm grateful to Andrew Bacon, John Hawthorne, Jeff Russell, and an anonymous referee for helpful comments and discussion. Special thanks are due to Christopher Menzel for many detailed comments, which lead to substantive improvements over previous drafts of the paper.

BIBLIOGRAPHY

Adams, R. M. (1974). Theories of actuality. *Noûs* **8**(3), 211–231.

Adams, R. M. (1981). Actualism and thisness. *Synthese* **49**(1), 3–41.

Barwise, J. (1975). *Admissible Sets and Structures: An Approach to Definability Theory.* Springer-Verlag.

Barwise, J., & Moss, L. (1996, July). *Vicious Circles: On the Mathematics of Non-Wellfounded Phenomena.* Center for the Study of Language and Information Publications.

Beall, J. (2000). A neglected response to the Grim result. *Analysis* **60**(1), 38–41.

18   GABRIEL UZQUIANO

Bell, J. (1995, March). Type reducing correspondences and well-orderings: Frege's and Zermelo's constructions re-examined. *Journal of Symbolic Logic* **60**(01), 209–221.

Bernays, P. (1942). A system of axiomatic set theory: Part IV. general set theory. *The Journal of Symbolic Logic* **7**(4), 133–145.

Boolos, G. (1997). Constructing Cantorian counterexamples. *Journal of Philosophical Logic* **26**(3), pp. 237–239.

Bringsjord, S. (1985). Are there set theoretic possible worlds? *Analysis* **45**(1), 64.

Cartwright, R. L. (2001). A question about sets. In *Fact and Value*, pp. 29–46. MIT Press.

Chihara, C. (1998). *The Worlds of Possibility: Modal Realism and the Semantics of Modal Logic.* Oxford: Clarendon Press.

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

Deutsch, H. (2014, January). Resolution of some paradoxes of propositions. *Analysis* **74**(1), 26–34.

Forster, T. (2008, June). The iterative conception of set. *The Review of Symbolic Logic* **1**(01), 97–110.

Fraenkel, A., Bar-Hillel, Y., & Levy, A. (1973). *Foundations of Set Theory.* Studies in Logic and the Foundations of Mathematics. Elsevier Science.

Friedman, H. (2004). Faithful representation in set theory with atoms.

Grim, P. (1984). There is no set of all truths. *Analysis* **44**(4), 206–208.

Jech, T. J. (2013, April). *The Axiom of Choice.* Courier Dover Publications.

Kanamori, A. (2004). Zermelo and set theory. *Bulletin of Symbolic Logic*, 487–553.

Kelley, J. L. (1975). *General Topology.* Springer Science & Business Media.

Klement, K. C. (2001). Russell's paradox in appendix B of the Principles of Mathematics: Was Frege's response adequate? *History and Philosophy of Logic* **22**(1), 13–28.

Kripke, S. A. (1980). *Naming and Necessity.* Harvard University Press.

Lewis, D. K. (1986). *On the Plurality of Worlds.* Blackwell Publishers.

McGee, V., & Rayo, A. (2000). A puzzle about de rebus belief. *Analysis* **60**(4), 297–299.

Mendelson, E. (1997). *Introduction to Mathematical Logic, Fourth Edition.* Discrete Mathematics and Its Applications. Taylor & Francis.

Menzel, C. (2012). Sets and worlds again. *Analysis* **72**(2), 304–309.

Menzel, C. (2014). Wide sets, ZFCU, and the iterative conception. *Journal of Philosophy* **111**(2), 57–83.

Plantinga, A. (1976). Actualism and possible worlds. *Theoria* **42**(1-3), 139–160.

Plantinga, A. (1992). *The Nature of Necessity.* Clarendon Press.

Prior, A. N. (1977). *Worlds, Times, and Selves.* Duckworth.

Russell, B. (1903). *The Principles of Mathematics.* W. W. Norton & Company.

Stalnaker, R. C. (1976). Possible worlds. *Noûs* **10**(1), 65–75.

Uzquiano, G. (2003). Plural quantification and classes. *Philosophia Mathematica* **11**(1), 67–81.

Zalta, E. (1999, December). Natural numbers and natural cardinals as abstract objects: A partial reconstruction of Frege's Grundgesetze in object theory. *Journal of philosophical logic* **28**(6), 617–658.

Zermelo, E. (1930). Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre. *Fundamenta Mathematicae* **16**, 29–47.