

## CAN MODALITIES SAVE NAIVE SET THEORY?

PETER FRITZ

IFIKK, University of Oslo

HARVEY LEDERMAN

Department of Philosophy, Princeton University

TIANKAI LIU

Department of Mathematics, University of Utah  
and

DANA SCOTT

Visiting Scholar, University of California, Berkeley

*To the memory of Prof. Grigori Mints, Stanford University*

*Born: June 7, 1939, St. Petersburg, Russia*

*Died: May 29, 2014, Palo Alto, California*

**§1. Introduction.** Considering only pure sets, the naive set comprehension principle says, for any condition, that there is a set containing all and only the sets satisfying this condition. In first-order logic, this can be formulated as the following schematic principle, where  $\varphi$  may be any formula in which  $y$  does not occur freely:

$$\exists y \forall x (x \in y \leftrightarrow \varphi). \quad (\text{Comp})$$

Russell's paradox shows that the instance obtained by letting  $\varphi$  be  $x \notin x$  is inconsistent in classical logic. One response to the paradox is to restrict naive set comprehension by ruling out this and other problematic instances: only for each of some *special conditions* is it claimed there is a set containing all and only the sets satisfying the condition. Many well-known set theories can be understood as instances of this generic response, differing in how they understand *special*. For example, the axiom schema of separation in Zermelo-Fraenkel set theory (ZF) restricts set comprehension to conditions which contain, as a conjunct, the condition of being a member of some given set:

$$\exists y \forall x (x \in y \leftrightarrow \varphi \wedge x \in z). \quad (\text{Sep})$$

Similarly, in Quine's New Foundations (NF) set comprehension is restricted to conditions which are stratified, where  $\varphi$  is stratified just in case there is a mapping  $f$  from individual variables to natural numbers such that for each subformula of  $\varphi$  of the form  $x \in y$ ,

---

Received: June 8, 2015.

2010 *Mathematics Subject Classification*: 03B45.

*Keywords and phrases*: set theory, naive comprehension, modal logic.

$f(y) = f(x) + 1$  and for each subformula of  $\varphi$  of the form  $x = y$ ,  $f(x) = f(y)$ . Both of these responses block Russell's paradox by ruling out the condition  $x \notin x$ .

Must every restriction of naive comprehension take the form of simply ruling out certain instances? In this article, we suggest and explore a different approach. As we have seen, standard set comprehension axioms restrict attention to some *special conditions*: for each of these special conditions, they provide for the existence of a set containing all and only the sets which satisfy it. Instead of restricting the conditions one is allowed to consider, we propose restricting the way in which the sets in question satisfy a given condition: for every condition, our comprehension axiom will assert the existence of a set containing all and only the sets satisfying that condition *in a special way*. Although we will suggest some more specific ways of understanding *in a special way* below, we will investigate the fruitfulness of this idea abstractly, considering whether there is any way for the qualification *in a special way* to behave according to which the correspondingly restricted version of set comprehension yields an interesting theory of sets. Consequently, we enrich the first-order language of pure set theory by an uninterpreted unary sentential operator  $\square$  which formalizes *in a special way*. Using this operator, the new comprehension principle can be stated as follows:

$$\exists y \forall x (x \in y \leftrightarrow \square \varphi). \quad (\text{Comp}\square)$$

Since the instance for  $\varphi$  being  $x \notin x$  is not ruled out, how might Russell's paradox be blocked on this approach? The set claimed to exist by this instance—call it the Russell set—must contain all and only the sets which satisfy, in a special way, the condition of not containing themselves, and so in particular, the Russell set must contain itself if and only if it satisfies, in a special way, the condition of not containing itself. *Prima facie*, this can be verified if the Russell set satisfies the condition of not containing itself without satisfying it in a special way.

Standard set theories such as ZFC are extremely well-developed; why should we set aside these theories and investigate an unfamiliar principle such as (Comp $\square$ )? Our main motivation is exploratory: we wish to see what results from this simple and natural alternative to more standard comprehension principles. The fact that one response to the set-theoretic paradoxes is known to lead to a rich mathematical theory should not prevent us from attempting to understand the consequences of other conceptually simple, mathematically natural replies.

A second motivation for investigating alternative restrictions of naive set comprehension is that separation requires a relatively large number of additional axioms to yield a strong set theory. What if (Comp $\square$ ), perhaps together with an axiom of extensionality, i.e.,

$$\forall x (x \in y \leftrightarrow x \in z) \rightarrow y = z \quad (\text{Ext})$$

gives rise to a useful set theory? Such an axiomatic system would constitute a considerable gain in simplicity over ZFC. (Simplicity is of course not enough: the axioms of NF, which are just extensionality and set comprehension restricted to stratified conditions, are considerably simpler than the axioms of ZF, but most set theorists still choose to work in the latter.)

A third reason for restricting set comprehension as in (Comp $\square$ ) is that this restriction fits certain views in the philosophy of mathematics and logic, on suitable ways of understanding the qualification "in a special way". One example is fictionalism, which will be discussed below. For another example, we can understand *in a special way* as *determinately*. To motivate this idea, consider an analogy to philosophical discussions of truth. Those who wish to preserve classical logic in the face of the liar paradox sometimes

introduce an operator to be read “determinately”, which distinguishes paradox-free claims from claims which are paradoxical. Equivalences such as  $\varphi \leftrightarrow T^{\ulcorner \varphi \urcorner}$  hold (where  $T$  is a truth predicate, and  $\ulcorner \varphi \urcorner$  is the Gödel number of  $\varphi$ ) if it is determinate whether  $\varphi$  holds, but not necessarily otherwise.<sup>1</sup> Given the well-known parallels between the liar paradox and Russell’s paradox, one might wonder whether a modification of this approach can be extended to the set-theoretic paradoxes. On this approach, comprehension might be restricted as in (Comp $\Box$ ), requiring only that for every condition  $\varphi$ , there is a set containing all and only the sets which are *determinately*  $\varphi$ . As described in the previous paragraph, this strategy might block a version of Russell’s paradox: if the Russell set does not belong to itself, it may nevertheless not be determinate that it does not belong to itself.

In the modal logics used below,  $\Box\top$  is equivalent to  $\top$  (the tautological constant), and so the corresponding instance of (Comp $\Box$ ) asserts the existence of a universal set. A final reason for exploring this comprehension principle is therefore an interest in set theories with a universal set; see Forster (1995) for motivations for admitting a universal set, and for an overview of existing approaches to such set theories.

In this article, we approach set theories based on (Comp $\Box$ ) from an abstract perspective, considering different principles governing  $\Box$  in different quantified modal logics, asking whether (Comp $\Box$ ) is consistent in them, and if so, whether any interesting set theory emerges from it. In keeping with our second motivation for exploring such systems, of potentially finding axiom systems which are simpler than the standard axioms of ZFC, we will focus on evaluating the strength of modalized comprehension on its own, or in conjunction with (Ext). It would also be interesting to consider the prospects of developing modal set theories based on these as well as additional axioms, but such considerations will mostly be beyond the scope of this article.

§2 formally introduces the quantified modal logics which we will use to investigate (Comp $\Box$ ). The relevant logics will be standard free extensions of arbitrary normal modal logics. Standard classes of Kripke models will be introduced, and later used in an instrumental capacity to prove consistency and other underivability results.

§3 considers (Comp $\Box$ ) in the strong modal logic **S5**, showing that (Comp $\Box$ ) is consistent in **S5** and so *a fortiori* in all weaker modal logics. Unfortunately, while the principle is consistent in these modal logics, the set theory it gives rise to is very weak. Indeed, we show that in **S5**, the nonmodal consequences of (Comp $\Box$ ) are precisely characterized by the extensional theory which states, for any finite number of elements, that there is the set containing them, and the set containing every other element. Moreover, we show that this result extends to any theory obtained by adding further nonmodal principles to **S5** + (Comp $\Box$ ). Thus there seems to be little hope of restricting naive comprehension using a single modal operator and obtaining a strong set theory while at the same time preserving the conceptual simplicity of a theory with (modal) comprehension and extensionality alone.

From the results just sketched, it is natural to conclude that (Comp $\Box$ ) is too weak. How could it be strengthened in a way which preserves the intuitive motivation with which we started? One idea can be motivated by our earlier example of a theory based on interpreting the operator  $\Box$  as “determinately”. On this interpretation, for every condition, (Comp $\Box$ ) asserts the existence of a set of all and only the sets which determinately satisfy the condition. But we may alter this principle so that for every condition, it states the existence of a set such that for all sets, *determinately*, the former contains the latter if and only if the latter *determinately* satisfies a given condition. Although we have articulated this idea

<sup>1</sup> For something along these lines, see Bacon (2015).

using a particular interpretation of  $\Box$ , it can of course be stated fully abstractly, as:

$$\exists y \forall x \Box (x \in y \leftrightarrow \Box \varphi). \quad (\Box \text{Comp} \Box)$$

$(\Box \text{Comp} \Box)$  turns out to be far stronger than  $(\text{Comp} \Box)$ ; indeed, as we show in §4.1, it overshoots the mark, as it is inconsistent in the relatively weak modal logic **KT**, and so *a fortiori* in its extensions. In response to this result, one could suggest using two different primitive modal operators in  $(\Box \text{Comp} \Box)$ , but since this drastically increases the space of available options, we won't consider it here. Instead, we explore the more restricted option of replacing the second occurrence of  $\Box$  in  $(\Box \text{Comp} \Box)$  by a string of modal operators and negations. In §4.2, we show exhaustively that such variants are inconsistent in **S4**.

The results obtained so far leave open the consistency of  $(\Box \text{Comp} \Box)$  in modal logics weaker than **KT**. Such logics do not prove the axiom  $(T)$ :  $\Box \varphi \rightarrow \varphi$ . This might seem nonnegotiable, given our earlier schematic reading of the modality as satisfying a condition *in a special way*: whatever satisfies a condition in a special way satisfies it *simpliciter*. But there are ways of reading  $\Box$  on which it is natural to give up  $(T)$ . One such reading arises naturally from the position of fictionalism in the philosophy of mathematics.<sup>2</sup> One of the fundamental questions in the philosophy of mathematics—some might say, *the* fundamental question—concerns the existence of mathematical objects. Are there any sets, for example? An important proposal is that mathematical objects do not in fact exist, although they exist according to a literally false, but nevertheless useful theory. This is a “fictionalist” approach to the philosophy of mathematics not because it absurdly claims that proofs from mathematical axioms are somehow merely fictionally correct, but because it claims that the things such as sets which we take to exist when we are writing and reading mathematics do not really exist. Reading  $\Box$  as “in the fiction (of there being sets)”,  $\Box \varphi \rightarrow \varphi$  is naturally rejected; fictionalists hold that in the fiction, there are sets, but in fact, there are none.

One axis of variation among versions of fictionalism concerns what we take to be the relevant fiction. At least on some versions of fictionalism, it's up to us—it depends on which fiction we choose. We may therefore choose to work in a fiction specified by the axioms of some standard set theory. But fictionalists may also explore more adventurous fictions; in particular, it is an intriguing idea to let the fiction specify what sets there are by making reference to the fiction itself. Therefore, one claim we may choose to be part of the fiction is that for any condition, there is a set containing all and only the sets satisfying the condition in the fiction. If we read  $\Box$  as “in the fiction”, the relevant part of the fictionalist theory is captured by a schema consisting of necessitations of universal closures of  $(\text{Comp} \Box)$ , as follows:

$$\Box \forall z_1 \dots \forall z_n \exists y \forall x (x \in y \leftrightarrow \Box \varphi(x, z_1, \dots, z_n)).$$

Here,  $\varphi(x, z_1, \dots, z_n)$  indicates that  $x, z_1, \dots, z_n$  are the free variables of  $\varphi$ . Of course, an analogous fictionalist theory can be specified along the lines of  $(\Box \text{Comp} \Box)$ . This kind of fictionalist would thus not endorse the truth of  $(\text{Comp} \Box)$  or  $(\Box \text{Comp} \Box)$ , but rather the truth of these principles *in the fiction*. Nevertheless, the availability of the fictionalist interpretation motivates the study of modal set theories based on  $(\text{Comp} \Box)$  and  $(\Box \text{Comp} \Box)$ , since these theories can then be understood as candidates for the *fictions* at issue in the relevant version of fictionalism. We are not aware of any proposal of this kind in the literature, but

<sup>2</sup> See, e.g., Field (1980, 1989), Rosen (2001), and Yablo (2001). Balaguer (2008) provides a helpful overview, with many more references.

it seems to us to be sufficiently congenial to fictionalism to be motivated by the general fictionalist project in the philosophy of mathematics.

§4.3 considers  $(\Box\text{Comp}\Box)$  in the context of modal logics which do not prove  $(T)$ . We first observe that this principle, as well as  $(\text{Comp}\Box)$ , are trivially consistent if the background modal logic does not prove the axiom  $(D)$ :  $\Box\phi \rightarrow \Diamond\phi$ . Reading  $\Box$  as “according to the fiction”, we may read  $\Diamond$  as “consistent with the fiction”; thus  $(D)$  expresses the very natural idea that what holds according to the fiction is consistent with the fiction. We then show that  $(\Box\text{Comp}\Box)$  is also inconsistent if **KD** is extended by one of the well-known axioms (4), (B) and (5).

§4 covers most of the standard normal modal logics and shows that among them, the principle is only consistent in very weak modal logics, and trivial if consistent. §5 therefore returns to the original principle  $(\text{Comp}\Box)$ , and considers its prospects in normal modal logics which are not included in **S5**. §5.1 shows that it is inconsistent in any proper extension of **S5**. §5.2 considers the logic **KDDc**, axiomatized by  $(D)$  and its converse  $(Dc)$ :  $\Diamond\phi \rightarrow \Box\phi$ . On the fictionalist interpretation of the modal operators, this expresses the natural constraint that the fiction be complete: what is consistent with the fiction must hold according to the fiction. We show that  $(\text{Comp}\Box)$  is consistent in **KDDc**; the models used in this proof suggest that the resultant set theories are significantly stronger than the set theory **S5** +  $(\text{Comp}\Box)$ . Indeed, the resulting theory can be consistently extended by principles in the presence of which  $(\Box\text{Comp}\Box)$  follows from  $(\text{Comp}\Box)$ , and so the former principle is also nontrivially consistent in **KDDc**.

We owe our investigation of these questions to the late Grigori “Grisha” Mints. In October of 2009 at Stanford University, Mints asked Scott whether a naive set theory could be consistent in modal logic. At that time Scott could not answer the consistency question, and neither could Mints, though they both agreed that a set theory based on  $(\text{Comp}\Box)$  would probably be very weak. In November 2014, Scott received a notice from Carnegie Mellon that there would be a philosophy seminar on a naive set theory by Lederman (see Field, Lederman, & Øgaard (2017)). Scott wrote him for his article and said, “By the way, there is this question of Grisha Mints, and I wonder if you have an opinion?” Lederman sent back a sketch of a proof of inconsistency for a slightly strengthened version of  $(\Box\text{Comp}\Box)$ , which did not quite work out, but the exchange became the basis for §4.1 and §4.2. In the first draft of the article, Scott and Lederman left open the consistency of  $(\text{Comp}\Box)$ , although they observed that it was not inconsistent by the analogue of the Russell set alone. Scott and Lederman tried out several model possibilities for the consistency of that principle, without success. In March of 2015 Liu approached them with a related model, which after a small correction gave a consistency proof; a few days later, Fritz approached them with essentially the same model. Fritz then provided the results of §5 on his own. Fritz and Lederman jointly wrote the introduction, and Fritz wrote the remainder of the article.

Before delving into the formal details of this article, we will situate the principles  $(\text{Comp}\Box)$  and  $(\Box\text{Comp}\Box)$  in the existing literature. Modalized comprehension principles have been studied in a number of different settings.<sup>3</sup> One is intensional higher-order logic

<sup>3</sup> There are also a number of ways in which modal logic has been used in connection with set theory without considering modalized comprehension principles. E.g., in provability logic, propositional modal logics can be used to study abstract features of a complex predicate expressing “it is provable in ZFC that” in certain first-order theories; see, e.g., Solovay (1976) and Boolos (1995). Propositional modal logics have also been used to characterize certain properties of forcing, roughly interpreting  $\Box$  as expressing “it is true in all forcing extensions that”; see, e.g., Hamkins

(see, e.g., Gallin (1975, p. 77) or Zalta (1988, p. 22)), where a syntactic distinction between types allows for an unrestricted comprehension principle. Such discussions usually work with models with constant (first- and higher-order) domains; for discussions of comprehension principles appropriate for variable domains of all types, see Williamson (2013, chaps. 6.3–6.4) and Fritz & Goodman (2016, §5).

Another form of modal comprehension principles occurs in modal set theories which are obtained by modalizing common set theories. Such a system for metaphysical necessity is presented in Fine (1981); systems for epistemic modalities were developed by several authors in the 1980s—see the contributions by Myhill, Goodman and Ščedrov to Shapiro (1985), or the references in Goodman (1990).

In the lecture at which Mints posed his original question (see Scott (2010)), Scott presented his Modal ZF, which uses the following modalization of the axiom schema of separation:

$$\exists y \Box \forall x (x \in y \leftrightarrow x \in u \wedge \varphi). \quad (\text{MZF Comp})$$

The modal comprehension principles mentioned in the preceding three paragraphs differ fundamentally from the naive principles in that they are modalizations of comprehension principles which do not give rise to the Russell paradox, either through employing type distinctions or through restrictions on the formulas with which instances may be obtained. Modalizing naive comprehension in order to make it consistent has been less widespread, but several such strands can be identified in the literature. The first uses modality to make the iterative conception of set explicit by reformulating comprehension to say that at some stage, there is a set defined by a given condition, using a possibility operator to formalize “at some stage”. Pioneered by Parsons (1983), such principles have been recently investigated in Studd (2013) and Linnebo (2013); see also Hellman (1989) and Linnebo (2010).

The second strand, and the closest to our own work, goes back to Fitch (1966, 1967a).<sup>4</sup> Fitch works in a language extended by a term-forming operator, which creates a term from a variable and a formula, and which we may write as  $\{.\}$ . He considers the principle:

$$x \in \{x : \varphi\} \leftrightarrow \Box \varphi. \quad (\text{Abst}\Box)$$

Fitch works in a strengthening of a quantified modal logic based on the propositional modal logic **KD**; all instances of the Barcan formula, (T) and (4) are true in the system, but one cannot necessitate instances of them (one can, however, necessitate instances of (Abst $\Box$ )). (A concrete example of this failure of necessitation is given in Fitch (1967a, pp. 102–3) (cf. Fitch (1967, p.107)).) Fitch proves the consistency of his system using techniques similar to those used later by Gilmore (1967) and Kripke (1975), but which Fitch had developed as early as Fitch (1942) (cf. Fitch (1948, 1963)). It is easy to see that both (Comp $\Box$ ) and ( $\Box$ Comp $\Box$ ) are derivable from (Abst $\Box$ ) by standard quantifier rules and necessitation, so Fitch establishes the consistency of these two principles in his extension

---

(2003), Hamkins & Woodin (2005), and Hamkins & Löwe (2008). For a different connection between modal logic and forcing, see Smullyan & Fitting (1996, Part III) and Fitting (2003). For yet another set-theoretic modality, see Blass (1990). Such combinations of modal logic and set theory are less closely related to the topic of this article than the variant modal set comprehension principles discussed in the following. (Thanks to Reviewer #4 for pushing us to mention the connections between modality and set theory discussed in this note.)

<sup>4</sup> The original technical article Fitch (1967a) contained an error, pointed out in a review by Rundle (1969); a correction appeared as Fitch (1970).

of **KD**. In §5.2 we prove a different, more natural strengthening of **KD** consistent using a different construction. (Our logic is incommensurable with Fitch's, so while both results imply the consistency of the comprehension principles in **KD**, neither result implies the other.) Despite these commonalities, Fitch's system differs from the ones investigated below in a number of crucial details. One difference is that we investigate  $(\text{Comp}\Box)$  and  $(\Box\text{Comp}\Box)$  separately, rather than  $(\text{Abst}\Box)$ , from which both of these principles follow. Another difference is that Fitch assumes the necessity of membership ( $x \in y \rightarrow \Box(x \in y)$ ); we return to this in §3. For further discussion of Fitch's work, and the history of his comparative neglect by other authors in the field, see Cantini (2009, §4.2).

A third strand starts with Aczel & Feferman (1980), who save the naive comprehension principle from inconsistency by replacing its material biconditional with a primitive binary intensional operator. Feferman (1984, §12) shows that if we abbreviate  $\varphi \equiv \top$  (where  $\equiv$  is their binary intensional operator) by  $\Box\varphi$ , we obtain from his models a consistency proof for a modal logic which validates  $(\text{Comp}\Box)$ , along with the Barcan formula, the modal schemes  $(K)$ ,  $(T)$  and  $(4)$ , and a number of other strong principles. Like Fitch, Aczel and Feferman consider a principle which features set-abstracts, and not merely the quantifiers as in more standard comprehension principles. Unlike Fitch's system, however, the logic in Feferman's system is not a normal modal logic; as Feferman observes (1984, p. 100) there are false instances of  $\Box(\varphi \vee \neg\varphi)$  in his model construction. The principle  $\neg\Box\varphi \rightarrow \Box\neg\Box\varphi$  also has false instances; the logic is thus substantially weaker than the one we show to be consistent with  $(\text{Comp}\Box)$ . With that said, however, in Feferman's model, as in Fitch's (but not in ours), the necessity of membership is maintained.

A fourth strand is the following comprehension principle, proposed by Krajíček (1987):

$$\exists y\forall x((\Box x \in y \leftrightarrow \Box\varphi) \wedge (\Box\neg x \in y \leftrightarrow \Box\neg\varphi)). \quad (\text{MCA})$$

Krajíček proves that this principle is inconsistent in **S5**, and it seems still to be an open problem whether it is consistent in the relatively weak modal logic **KT** (see Krajíček (1988) and Kaye (1993)). This fourth strand of the literature is also close to our own project. Indeed, a reasonable response to the results of this article might be to investigate (MCA) with renewed energy. Once the principles considered here have been shown unlikely to yield an interesting set theory, (MCA) might be thought the most promising remaining modalized comprehension principle.

**§2. Logics and models.** In this section, we set out the logics and models with which we will investigate our modalized set comprehension principles. Let  $\mathcal{L}_0^\Box$  be a language of propositional modal logic, based on a countably infinite set of proposition letters, from which formulas are constructed using the Boolean operators  $\neg$  and  $\wedge$  and a unary modal operator  $\Box$ . A set of formulas in this language is a *normal modal logic* if it contains all propositional tautologies and the distributivity axiom  $(K) = \Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$ , and is closed under the rules of modus ponens, uniform substitution and necessitation. The normal modal logic *axiomatized* by formulas  $A_1, \dots, A_n$ , written **KA** $_1 \dots \mathbf{A}_n$ , is the smallest normal modal logic containing  $A_1, \dots, A_n$ .

Let  $\mathcal{L}_1(\in)$  be a first-order language based on a countably infinite set of individual variables, in which atomic predications are formed using two binary relation symbols  $=$  and  $\in$ , and complex formulas are constructed using the Boolean operators  $\neg$  and  $\wedge$ , and universal quantifiers  $\forall x$ . Let  $\mathcal{L}_1^\Box(\in)$  be the extension of  $\mathcal{L}_1(\in)$  by the unary sentential operator  $\Box$ . Other common symbols, such as  $\vee$ ,  $\rightarrow$ ,  $\top$ ,  $\perp$ ,  $\exists$ ,  $\diamond$ ,  $\neq$  and  $\notin$  will be used as metalinguistic abbreviations in the usual way, and  $Ex$  will be used to abbreviate  $\exists y(y = x)$ ,

where  $y$  is the first variable distinct from  $x$  in some given order of the variables. Finite conjunctions and disjunctions will sometimes be written using  $\bigwedge$  and  $\bigvee$ , respectively, as in  $\bigwedge_{i < n} \varphi_i$  and  $\bigvee_{i < n} \varphi_i$ . Where  $n = 0$ , they are understood to stand for  $\top$  and  $\perp$ , respectively. Let  $\mathcal{S}_1(\epsilon)$  and  $\mathcal{S}_1^\square(\epsilon)$  be the sets of sentences, i.e., closed formulas, of these languages.

For any normal modal logic  $\Lambda$ , let  $Q\Lambda$  be the set of  $\mathcal{L}_1^\square$ -formulas derivable in the following axiomatic calculus, adapted from Hughes & Cresswell (1996, chaps. 16 & 17):

- ( $\Lambda$ ) any substitution instance of a theorem of  $\Lambda$ .
- ( $\forall 1E$ )  $(\forall x\varphi \wedge Ey) \rightarrow \varphi[y/x]$ .
- ( $\forall \rightarrow$ )  $\forall x(\varphi \rightarrow \psi) \rightarrow (\forall x\varphi \rightarrow \forall x\psi)$ .
- ( $VQ$ )  $\varphi \leftrightarrow \forall x\varphi$ , provided  $x$  is not free in  $\varphi$ .
- ( $UE$ )  $\forall xEx$ .
- ( $I1$ )  $x = x$ .
- ( $I2$ )  $x = y \rightarrow (\varphi \rightarrow \psi)$ , where  $\varphi$  and  $\psi$  differ only in that  $\varphi$  has free  $x$  in places where  $\psi$  has free  $y$ .
- ( $LNI$ )  $x \neq y \rightarrow \square(x \neq y)$ .
- ( $MP$ ) From  $\varphi$  and  $\varphi \rightarrow \psi$ , derive  $\psi$ .
- ( $N$ ) From  $\varphi$ , derive  $\square\varphi$ .
- ( $UG$ ) From  $\varphi$ , derive  $\forall x\varphi$ .
- ( $UGL\forall^n$ ) From  $\varphi_1 \rightarrow \square(\varphi_2 \rightarrow \dots \rightarrow \square(\varphi_n \rightarrow \square\psi) \dots)$ , derive  $\varphi_1 \rightarrow \square(\varphi_2 \rightarrow \dots \rightarrow \square(\varphi_n \rightarrow \square\forall x\psi) \dots)$ , where  $x$  is not free in  $\varphi_1, \dots, \varphi_n$ .

For any further axioms or axiom schemas  $A_1, \dots, A_n$ , let  $Q\Lambda + A_1 + \dots + A_n$  be the set of formulas derivable in the axiomatic calculus obtained by adding  $A_1, \dots, A_n$  to the above. If  $X$  is a logic (set of formulas) and  $\varphi$  a formula, we write  $X \vdash \varphi$  for  $\varphi \in X$ .

The axioms and rules of  $Q\Lambda$  may look unnatural and overly complicated. The reason for choosing this system is that (i) it includes classical first-order logic, in the sense that all *sentences* of standard first-order logic are derivable, (ii) it smoothly combines with arbitrary normal modal logics and (iii) it does not prove the following two schematic principles, known respectively as the Barcan formula and its converse:

- ( $BF$ )  $\forall x\square\varphi \rightarrow \square\forall x\varphi$ ,
- ( $CBF$ )  $\square\forall x\varphi \rightarrow \forall x\square\varphi$ .

These two principles have been at the heart of debates in the metaphysics of necessity and possibility. If the modal operators  $\square$  and  $\diamond$  are read as “necessarily” and “possibly”, respectively, the Barcan formula entails that if it is possible that something exists, then there is something which is possibly identical to it. Given that necessarily, everything is self-identical, an instance of the converse Barcan formula allows us to show that everything exists necessarily. Needless to say, both of these results are highly controversial. Similarly, in the present setting, there is no obvious reason why the claim that every set satisfies a condition in a special way should be equivalent to the claim that in a special way every set satisfies that condition. Of course, it is interesting whether such a principle can be consistently added to a given theory in  $\mathcal{L}_1^\square(\epsilon)$ , and we will return to this question at various points below.

Another aspect of  $Q\Lambda$  worth mentioning is the axiom ( $LNI$ ) and the fact that an analogous principle for  $=$  is derivable (see the proof of Lemma 3.2). These assumptions are natural on the metaphysical interpretation of modal operators just mentioned; in the present setting, they are merely included for simplicity, as they lead to a natural logic for



identity which corresponds to a standard treatment of identity in Kripke models. Although this won't be explicitly discussed in the following, it is clear that a number of the following results do not depend on the identity axioms. In particular, none of the derivations showing that particular modalized comprehension principles are inconsistent in a given modal logic features the identity symbol; *a fortiori* these results do not depend on laws governing identity.

One useful feature of  $Q\Lambda$  is that it admits the rule of substitution of equivalents: if  $\varphi \leftrightarrow \psi$  is provable, then so is  $\chi \leftrightarrow \chi'$ , where  $\chi'$  differs from  $\chi$  only in having  $\psi$  in some places where  $\chi$  has  $\varphi$ ; see Hughes & Cresswell (1996) for further discussion. Appeals to this fact, as well as other elementary features of  $Q\Lambda$  such as its including classical propositional logic, will not be explicitly recorded in derivations and other arguments in the following.

Turning to model theory, a Kripke frame (for  $\mathcal{L}_0^\square$ ) is a pair  $\langle W, R \rangle$  consisting of a set  $W$ , the "possible worlds", and a binary "accessibility" relation  $R \subseteq W^2$ ; validity of  $\mathcal{L}_0^\square$ -formulas relative to a class of Kripke frames is defined as usual. In the following, we will often exploit well-known results to the effect that an  $\mathcal{L}_0^\square$ -formula is a theorem of some normal modal logic just in case it is valid on some class of frames. These results are summarized in the following table. Each line states the name of an axiom, the axiom itself, the name of a condition on the accessibility relation of a frame, and a statement of this condition in first-order logic, with quantifiers ranging over the worlds of the relevant frame and initial universal quantifiers omitted:

|       |  |               |                                      |
|-------|--|---------------|--------------------------------------|
| $T$   | $\Box p \rightarrow p$                   | reflexivity   | $wRw$                                |
| $B$   | $p \rightarrow \Box \Diamond p$          | symmetry      | $wRv \rightarrow vRw$                |
| $4$   | $\Box p \rightarrow \Box \Box p$         | transitivity  | $(wRv \wedge vRu) \rightarrow wRu$   |
| $5$   | $\Diamond p \rightarrow \Box \Diamond p$ | euclideaness  | $(wRv \wedge wRu) \rightarrow vRu$   |
| $D$   | $\Box p \rightarrow \Diamond p$          | seriality     | $\exists v(wRv)$                     |
| $D_c$ | $\Diamond p \rightarrow \Box p$          | functionality | $(wRv \wedge wRu) \rightarrow v = u$ |

For any axioms  $A_1, \dots, A_n$  taken from this list,  $\mathbf{KA}_1 \dots \mathbf{A}_n$  can be shown to be the set of  $\mathcal{L}_0^\square$ -formulas valid on the class of Kripke frames whose accessibility relations satisfy all of the conditions corresponding to  $A_1, \dots, A_n$  in this table. Proofs of these results can be found in any standard introduction to modal logics, e.g., Hughes & Cresswell (1996).

A Kripke model (for  $\mathcal{L}_1^\square(\in)$ ) is a structure  $\langle W, R, D, V \rangle$  such that  $\langle W, R \rangle$  is a Kripke frame,  $D$  is a function mapping each world to a set, and  $V$  maps each world to a binary relation on  $\bigcup_{w \in W} D(w)$ . Truth of a formula is defined relative to a world  $w$  and an assignment function  $a$ , with the crucial clauses as follows:

- $M, w, a \models x = y$  iff  $a(x) = a(y)$ ,
- $M, w, a \models x \in y$  iff  $\langle a(x), a(y) \rangle \in V(w)$ ,
- $M, w, a \models \Box \varphi$  iff for all  $v \in W$ , if  $wRv$  then  $M, v, a \models \varphi$ ,
- $M, w, a \models \forall x \varphi$  iff for all  $o \in D(w)$ ,  $M, w, a[o/x] \models \varphi$ ,

where  $a[o/x]$  maps  $x$  to  $o$  and every other variable  $y$  to  $a(y)$ . Validity on a model is defined as truth in all worlds on all assignment functions, and validity on a class of frames as validity on all models based on a frame in the class. A routine induction shows that if  $\Lambda$  is valid on a class of Kripke frames  $C$ , then  $Q\Lambda$  is valid on the class of Kripke models based on a frame in  $C$ , and that this extends to adding  $(BF)/(CBF)$  if the models in the class are all decreasing/increasing (in the sense that if  $wRv$  then  $D(w) \supseteq D(v)/D(w) \subseteq D(v)$ ). (Note that this observation depends on our model-theoretic treatment of identity; the analogous claim does not hold for a variant which adds to the truth-conditions for

identity the requirement that  $a(x) \in D(w)$ .) To show the consistency of, say,  $Q\Lambda + (BF) + (Comp\Box)$  it will therefore suffice to construct a nonempty class of decreasing Kripke models based on frames validating  $\Lambda$  on which  $(Comp\Box)$  is valid.

**§3.  $(Comp\Box)$  in  $S5$ .** Recall the basic modalized comprehension principle:

$$\exists y \forall x (x \in y \leftrightarrow \Box \varphi). \tag{Comp\Box}$$

We first consider  $(Comp\Box)$  in the strong modal logic  $S5$ , the normal modal logic axiomatized by the following two axioms:

- (T)  $\Box p \rightarrow p,$
- (5)  $\Diamond p \rightarrow \Box \Diamond p.$

We show that  $QS5 + (Comp\Box)$  is consistent. In fact, this will fall out of a theorem which is stronger in several ways. First, the theorem will precisely characterize the nonmodal fragment of  $QS5 + (Comp\Box)$  as the first-order theory axiomatized by the following two axiom schemas:

- (F)  $\exists y \forall x (x \in y \leftrightarrow \bigvee_{i \leq n} x = z_i),$
- (CF)  $\exists y \forall x (x \in y \leftrightarrow \bigwedge_{i \leq n} x \neq z_i).$

Informally, this theory can be summed up by saying that for any finite number of sets, there is the set containing them and the set not containing them. Second, the characterization of the nonmodal fragment is preserved under adding any further nonmodal assumptions: for any set of nonmodal sentences  $\Gamma$ , the nonmodal fragment of  $QS5 + (Comp\Box) + \Gamma$  is  $(F) + (CF) + \Gamma$ . Third, the theorem will prove that this holds for all normal modal logics included in  $S5$  which include the weak modal logic  $KD$ , axiomatized by the single axiom:<sup>5</sup>

$$(D) \quad \Box p \rightarrow \Diamond p.$$

**THEOREM 3.1.** *Let  $\Lambda$  be a normal modal logic such that  $KD \subseteq \Lambda \subseteq S5$ ,  $\Gamma \subseteq S_1(\in)$ , and  $\varphi \in S_1(\in)$ . Then*

$$Q\Lambda + (Comp\Box) + \Gamma \vdash \varphi \text{ iff } (F) + (CF) + \Gamma \vdash \varphi.$$

We split up the proof into several lemmas. First, we establish the right-to-left direction by showing that the modal theory proves  $(F)$  and  $(CF)$ :

**LEMMA 3.2.** *All instances of  $(F)$  and  $(CF)$  are provable in  $QKD + (Comp\Box)$ .*

*Proof.* The necessity of identity is derivable using the axioms and rules of  $Q\Lambda$ ; see Hughes & Cresswell (1996, p. 313):

$$(LI) \quad x = y \rightarrow \Box x = y.$$

Using  $(D)$ ,  $(LI)$ , and  $(LNI)$ , the following two are easily derived:

---

<sup>5</sup> An earlier version of the article only demonstrated the consistency of  $QS5 + (Comp\Box)$ , using a model construction along the lines carried out below. We owe the suggestion for the first strengthening, to show that the nonmodal fragment of this logic is precisely  $(F) + (CF)$ , as well as a sketch of the proof, to Reviewer #6. The reviewer also noted that this result holds when the axiom of extensionality is added; this suggestion led to the second strengthening.

- (1)  $x = y \leftrightarrow \Box x = y,$
- (2)  $\Diamond x = y \leftrightarrow \Box x = y.$

Note that the following is a theorem of **KD**:

$$(3) \quad \bigwedge_{i \leq n} (\Diamond p_i \leftrightarrow \Box p_i) \rightarrow (\Box \bigvee_{i \leq n} p_i \leftrightarrow \bigvee_{i \leq n} \Box p_i).$$

Instantiating with  $x = z_i$  for  $p_i$  and appealing to (1) and (2), we obtain:

$$(4) \quad \Box \bigvee_{i \leq n} x = z_i \leftrightarrow \bigvee_{i \leq n} \Box x = z_i.$$

The following is an instance of (Comp $\Box$ ):

$$(5) \quad \exists y \forall x (x \in y \leftrightarrow \Box \bigvee_{i \leq n} x = z_i).$$

(F) follows from (4) and (5). The proof of (CF) is analogous. □

For the converse direction, consider any  $\varphi \in \mathcal{S}_1(\in)$  such that  $\varphi \notin (F) + (CF) + \Gamma$ . By the completeness of first-order logic, there is a model of  $(F) + (CF) + \Gamma$  in which  $\varphi$  is false. Let  $A$  be its domain and  $B \subseteq A^2$  be the relation used to interpret  $\in$ . We show that  $\varphi \notin Q\Lambda + (\text{Comp}\Box) + \Gamma$ , for any given normal modal logic  $\Lambda$  included in **S5**, by constructing a Kripke model  $M = \langle W, R, D, V \rangle$  based on a frame validating **S5** which validates (Comp $\Box$ ) and  $\Gamma$  such that  $D(w) = A$  and  $V(w) = B$  for some  $w \in W$ .

The idea behind the model construction is to let the interpretation of  $\in$  vary among worlds to such an extent that the only witnesses required for the validity of (Comp $\Box$ ) correspond to the finite and cofinite sets—which can be specified using  $=$  alone. This can be done by using permutations of  $A$  as worlds, interpreting  $\in$  accordingly. As we show, it is not even necessary to include all permutations. More precisely, for any permutation  $\pi$  of  $A$ , let the set of elements of  $A$  not mapped to themselves by  $\pi$  be called the *support* of  $\pi$ ; let  $W$  be the set of permutations of  $A$  with finite support. Let  $R = W^2$ —thereby ensuring that the underlying frame validates **S5**—and  $D$  be such that  $D(\pi) = A$  for all  $\pi \in W$ . Recalling that  $B$  interprets the membership relation in the model with which we began, for all  $\pi \in W$ , let

$$V(\pi) = \pi(B) = \{ \langle \pi(o_1), \pi(o_2) \rangle : \langle o_1, o_2 \rangle \in B \}.$$

To prove that  $M$  validates (Comp $\Box$ ), define the extension of a formula with a distinguished variable relative to a world and assignment:

$$\llbracket \varphi(x) \rrbracket_{M, \pi, a} = \{ o \in A : M, \pi, a[o/x] \models \varphi \}.$$

We show how to apply a permutation of  $A$  to worlds and assignment functions, and prove that the extension of a formula  $\varphi$  is invariant if we apply to the world and assignment function any permutation of  $A$  which maps all parameters of  $\varphi$  to themselves. It will then follow that the extension of a formula is determined entirely by its parameters, in the sense that it either contains all other elements of  $A$  or none of them; consequently, it expresses a set which is finite or cofinite in  $A$ . Along the way, we establish that all worlds agree on the sentences they validate: given that one world, the identity permutation, validates  $\Gamma$ , it follows that  $\Gamma$  is valid on the model. The following makes this strategy precise.

For any permutations  $\pi$  and  $\rho$ , let  $\pi\rho$  be the composition of  $\pi$  and  $\rho$ ; for  $X \subseteq W$ , let  $\pi X = \{ \pi\rho : \rho \in X \}$ ; for  $O \subseteq A$ , let  $\pi(O) = \{ \pi(o) : o \in O \}$ . For any assignment function  $a$ , let  $\pi(a)$  be the assignment function such that  $\pi(a)(z) = \pi(a(z))$ . In the following lemmas, unless noted otherwise,  $\pi$  and  $\rho$  are arbitrary members of  $W$ ,  $a$  is an arbitrary assignment function, and  $\varphi$  an arbitrary formula.

LEMMA 3.3.  $M, \pi, a \models \varphi$  iff  $M, \rho\pi, \rho(a) \models \varphi$ .

*Proof.* By induction on the complexity of  $\varphi$ ; only the case for  $\in$  is interesting:

$M, \pi, a \models x \in y$  iff,

$\langle a(x), a(y) \rangle \in \pi(B)$  iff,

$\langle \rho(a)(x), \rho(a)(y) \rangle \in \rho\pi(B)$  iff,

$M, \rho\pi, \rho(a) \models x \in y$ . □

Since by construction all elements of  $\Gamma$  are true at the identity permutation (on any assignment), this lemma already establishes that  $\Gamma$  is valid on  $M$ .

LEMMA 3.4.  $\rho(\llbracket \varphi(x) \rrbracket_{M,\pi,a}) = \llbracket \varphi(x) \rrbracket_{M,\rho\pi,\rho(a)}$ .

*Proof.*  $\rho(\llbracket \varphi(x) \rrbracket_{M,\pi,a})$

$= \{ \rho(o) : o \in A \text{ and } M, \pi, a[o/x] \models \varphi \}$

$= \{ o \in A : M, \pi, a[\rho^{-1}(o)/x] \models \varphi \}$

$= \{ o \in A : M, \rho\pi, \rho(a[\rho^{-1}(o)/x]) \models \varphi \}$  (by the previous lemma)

$= \{ o \in A : M, \rho\pi, \rho(a)[o/x] \models \varphi \}$

$= \llbracket \varphi(x) \rrbracket_{M,\rho\pi,\rho(a)}$ . □

LEMMA 3.5. If  $\rho(a(z)) = a(z)$  for all variables  $z$  free in  $\varphi$ ,  $\rho(\llbracket \Box\varphi(x) \rrbracket_{M,\pi,a}) = \llbracket \Box\varphi(x) \rrbracket_{M,\pi,a}$ .

*Proof.*  $\rho(\llbracket \Box\varphi(x) \rrbracket_{M,\pi,a})$

$= \llbracket \Box\varphi(x) \rrbracket_{M,\rho\pi,\rho(a)}$  (by the previous lemma)

$= \bigcap_{\sigma \in W} \llbracket \varphi(x) \rrbracket_{M,\sigma,\rho(a)}$

$= \bigcap_{\sigma \in W} \llbracket \varphi(x) \rrbracket_{M,\sigma,a}$  (since  $\rho(a(z)) = a(z)$  for all variables  $z$  free in  $\varphi$ )

$= \llbracket \Box\varphi(x) \rrbracket_{M,\pi,a}$ . □

LEMMA 3.6. If  $O \subseteq A$  is finite and  $O' \subseteq A$  is such that  $\pi(O') = O'$  for all  $\pi \in W$  such that  $\pi(o) = o$  for all  $o \in O$ , then  $O'$  is finite or cofinite in  $A$ .

*Proof.* Assume for the sake of contradiction that  $O'$  is infinite and coinfinite in  $A$ . Then there are  $o_1, o_2 \in A \setminus O$  such that  $o_1 \in O'$  and  $o_2 \notin O'$ . Now consider the transposition  $(o_1 o_2)$  which switches  $o_1$  and  $o_2$ .  $(o_1 o_2) \in W$ , but  $(o_1 o_2)(O') \neq O'$ . But this contradicts the assumption. □

LEMMA 3.7. (Comp $\Box$ ) is valid in the model  $M$ .

*Proof.* Consider any  $\varphi$  in which  $y$  is not free, any  $\pi \in W$  and any assignment  $a$ . It suffices to show that  $M, \pi, a \models \exists y \forall x (x \in y \leftrightarrow \Box\varphi)$ . Note that by the preceding two lemmas,  $X = \llbracket \Box\varphi(x) \rrbracket_{M,\pi,a}$  is finite or cofinite in  $A$ . Since  $B$  is the relation interpreting  $\in$  in a model of  $(F) + (CF)$ , there is an  $o \in A$  such that for all  $o' \in A$ ,  $\langle o', o \rangle \in B$  iff  $o' \in \pi^{-1}(X)$ . Hence for all  $o' \in A$ ,  $\langle o', \pi(o) \rangle \in \pi(B)$  iff  $o' \in X$ , so  $\pi(o)$  witnesses the existential claim. □

This completes the proof of Theorem 3.1. The construction presented here restricts the worlds to permutations *with finite support* only to demonstrate concretely that the relevant models need only contain countably many worlds. One could arrive at the same conclusion more abstractly by including all permutations, and then appealing to an analog of the Löwenheim-Skolem theorem such as the one established in Kripke (1959, p. 7).

$(F) + (CF)$  is a very weak set theory: For every infinite set  $A$ , there are many functions  $f$  from  $A$  to its powerset whose range contains all finite and cofinite subsets of  $A$ . For

any such  $A$  and  $f$ ,  $A$  and  $B = \{\langle o_1, o_2 \rangle : o_1 \in f(o_2)\}$  form a model of  $(F) + (CF)$ . Theorem 3.1 therefore shows that  $(\text{Comp}\Box)$  is a weak comprehension principle on its own. But it does more: it also shows that considering only the nonmodal fragment, the deductive strength of  $(\text{Comp}\Box)$  is precisely captured by  $(F) + (CF)$ . This suggests that if we were to try to enrich  $(\text{Comp}\Box)$  by supplementary axioms to obtain an interesting set theory, we might need to do so using modal axioms: it seems *prima facie* unlikely that  $(F) + (CF)$  could play a useful role in axiomatizing a stronger extensional set theory with a universal set. Any further modal axioms which are to strengthen the resulting set theory must not already be validated by the models constructed above. It is therefore of interest to note some natural modal principles which they validate. Since the models constructed above have constant domains, they validate  $(BF)$  and  $(CBF)$ . The next proposition records three further validities.

PROPOSITION 3.8. *The following are valid on any model  $M$  as above:*

$$\begin{aligned} (\text{Comp}\Diamond) \quad & \exists y \forall x (x \in y \leftrightarrow \Diamond \varphi(x)), \\ (\text{Mem}) \quad & \Diamond x \in y, \\ (\text{Non}) \quad & \Diamond x \notin y. \end{aligned}$$

*Proof.* The proof of  $(\text{Comp}\Diamond)$  is analogous to that of  $(\text{Comp}\Box)$ . For  $(\text{Mem})$ , consider any  $\pi \in W$  and assignment function  $a$ . Let  $o \in A$  such that  $\langle o', o \rangle \in B$  for all  $o' \in A$ , and  $\rho = (a(y), o)$ . Then  $\langle a(x), a(y) \rangle \in \rho(B)$ , and so  $M, \rho, a \models x \in y$ , as required. The proof of  $(\text{Non})$  is similar.  $\square$

Since all of these principles are already valid on the models constructed above, they cannot be used to strengthen  $QS5 + (\text{Comp}\Box)$ .<sup>6</sup> Two natural principles which are not valid on these models are the rigidity of membership, and the rigidity of nonmembership:

$$\begin{aligned} x \in y \rightarrow \Box x \in y. & \quad (\text{Rig}\in) \\ x \notin y \rightarrow \Box x \notin y. & \quad (\text{Rig}\notin) \end{aligned}$$

We will show that adding these leads to inconsistency. In fact, in the presence of the following axiom (which is provable in  $S5$ ), adding the former is sufficient for inconsistency:

$$(B) \quad p \rightarrow \Box \Diamond p,$$

for in the presence of  $(B)$ ,  $\text{Rig}\notin$  can be derived from  $\text{Rig}\in$ . However,  $(\text{Rig}\in)$  by itself does not lead to inconsistency in  $QKD + (\text{Comp}\Box)$ ; this follows from the consistency results in Fitch (1967a).

PROPOSITION 3.9.

1.  $QKD + (\text{Comp}\Box) + (\text{Rig}\in) + (\text{Rig}\notin)$  is inconsistent.
2.  $QKDB + (\text{Comp}\Box) + (\text{Rig}\in)$  is inconsistent.

<sup>6</sup> Another modal principle valid on the models discussed here is the homogeneity schema discussed in Fine (1978). (Thanks to Reviewer #6 for this observation.) Calling a formula *de dicto* if it contains no subformula of the form  $\Box \varphi$  where  $\varphi$  has a free variable, Fine shows that for every formula  $\varphi$ , there is a *de dicto* formula  $\psi$  such that  $\varphi \leftrightarrow \psi$  is provable in the logic which results from adding the homogeneity schema to  $QS5$ . So, even forcing every formula to be equivalent to a *de dicto* formula would not help to take us beyond  $(F) + (CF)$ .

*Proof.*

(i)

- (1)  $y \in y \rightarrow \Box y \in y$  (Rig $\in$ )
- (2)  $\Box y \in y \rightarrow \neg \Box y \notin y$  (D)
- (3)  $\neg(y \in y \wedge \Box y \notin y)$  (1), (2)
- (4)  $y \notin y \rightarrow \Box y \notin y$  (Rig $\notin$ )
- (5)  $\neg(y \notin y \wedge \neg \Box y \notin y)$  (4)
- (6)  $\neg(y \in y \leftrightarrow \Box y \notin y)$  (3), (5)
- (7)  $(\forall x(x \in y \leftrightarrow \Box x \notin x) \wedge Ey) \rightarrow (y \in y \leftrightarrow \Box y \notin y)$  ( $\forall 1E$ )
- (8)  $Ey \rightarrow \neg \forall x(x \in y \leftrightarrow \Box x \notin x)$  (6), (7)
- (9)  $\forall y(Ey \rightarrow \neg \forall x(x \in y \leftrightarrow \Box x \notin x))$  (8), (UG)
- (10)  $\forall y Ey$  (UE)
- (11)  $\forall y \neg \forall x(x \in y \leftrightarrow \Box x \notin x)$  (9), (10), ( $\forall \rightarrow$ )
- (12)  $\neg \exists y \forall x(x \in y \leftrightarrow \Box x \notin x)$  (11)

(ii) In place of step (4) we insert the following derivation:

- (4a)  $\Diamond y \notin y \rightarrow y \notin y$  (Rig $\notin$ )
- (4b)  $\Box \Diamond y \notin y \rightarrow \Box y \notin y$  (4a), (K), (N)
- (4c)  $y \notin y \rightarrow \Box \Diamond y \notin y$  (B)
- (4d)  $y \notin y \rightarrow \Box y \notin y$  (4b, c)

The rest of the proof is exactly as above. □

**§4. ( $\Box\text{Comp}\Box$ ).** **S5** is a very strong modal logic; most common modal logics are included in it. Since the results of the previous section indicate that even in such a strong background logic, ( $\text{Comp}\Box$ ) only leads to a relatively weak set theory, it is natural to strengthen it by replacing its material biconditional by a strict biconditional:

$$\exists y \forall x \Box(x \in y \leftrightarrow \Box \varphi). \tag{\Box\text{Comp}\Box}$$

This principle was the subject of Mints’s original question; it turns out to be inconsistent in the relatively weak modal logic **KT**, axiomatized by the single axiom (*T*). Had he seen the weakness of ( $\text{Comp}\Box$ ) and our first contradiction using ( $\Box\text{Comp}\Box$ ) below, we suspect Mints himself would have considered variants of ( $\Box\text{Comp}\Box$ ) in which the second modal operator is replaced by some other modality  $*$ , i.e., any string of modal operators and negations:

$$\exists y \forall x \Box(x \in y \leftrightarrow * \varphi). \tag{\Box\text{Comp}*}$$

Accordingly, we will also consider these principles systematically below.

**4.1. Inconsistency in **KT**.** The inconsistency of ( $\Box\text{Comp}\Box$ ) in **KT** follows from the fact that the negation of the following principle, for  $*$  =  $\Box$ , is derivable in **QKT**:

$$\exists y \forall x \Box(x \in y \leftrightarrow *x \notin x). \tag{\Box\text{Russell}*}$$

The argument for this will be factored into a propositional and a quantified part; the latter part is most conveniently formulated more generally for arbitrary modalities, so as to be of further use below.<sup>7</sup>

<sup>7</sup> We are grateful to Reviewer #5 and especially to Reviewer #6 for suggestions which greatly simplified the proofs of the inconsistency of the different instances of ( $\Box\text{Comp}*$ ) considered here.

LEMMA 4.1.  $\mathbf{KT} \vdash \neg\Box(p \leftrightarrow \Box\neg p)$ .

*Proof.* By a standard completeness result, it suffices to show that this formula is valid on reflexive Kripke frames. Assume for contradiction that there is a Kripke model  $M$  based on a frame  $\langle W, R \rangle$ , where  $R$  is reflexive, and a world  $w \in W$  such that  $M, w \vDash \Box(p \leftrightarrow \Box\neg p)$ . Then  $M, v \vDash p \leftrightarrow \Box\neg p$  for all  $v \in W$  such that  $wRv$ . We distinguish two cases:

Case 1.  $M, w \vDash \Diamond p$ . There must be a  $v \in W$  such that  $wRv$  and  $M, v \vDash p$ . So  $M, v \vDash \Box\neg p$ , contradicting the reflexivity of  $R$ .  $\zeta$ .

Case 2.  $M, w \not\vDash \Diamond p$ . Then  $M, v \vDash \neg p$  for all  $v \in W$  such that  $wRv$ . Thus  $M, w \vDash \Box\neg p$ , so by reflexivity of  $R$ ,  $M, w \vDash \neg p$ . By reflexivity of  $R$  again,  $M, w \vDash p \leftrightarrow \Box\neg p$ , and so  $M, w \vDash p$ .  $\zeta$ . □

LEMMA 4.2. For any normal modal logic  $\Lambda$ , if  $\Lambda \vdash \neg\Box(p \leftrightarrow *\neg p)$ , then  $Q\Lambda \vdash \neg(\Box\text{Russell}^*)$ .

*Proof.*

- (1)  $\neg\Box(y \in y \leftrightarrow *y \notin y)$  (Λ)
- (2)  $(\forall x\Box(x \in y \leftrightarrow *x \notin x) \wedge Ey) \rightarrow \Box(y \in y \leftrightarrow *y \notin y)$  (∀1E)
- (3)  $Ey \rightarrow \neg\forall x\Box(x \in y \leftrightarrow *x \notin x)$  (1), (2)
- (4)  $\forall y(Ey \rightarrow \neg\forall x\Box(x \in y \leftrightarrow *x \notin x))$  (3), (UG)
- (5)  $\forall yEy$  (UE)
- (6)  $\forall y\neg\forall x\Box(x \in y \leftrightarrow *x \notin x)$  (4), (5), (∀→)
- (7)  $\neg\exists y\forall x\Box(x \in y \leftrightarrow *x \notin x)$  (6) □

COROLLARY 4.3.  $Q\mathbf{KT} \vdash \neg(\Box\text{Russell}\Box)$ .

**4.2. Inconsistency of variants in S4.** We consider  $(\Box\text{Comp}^*)$  in the more restrictive setting of  $\mathbf{S4}$ , the normal modal logic axiomatized by  $(T)$  and the following axiom:

$$(4) \Box p \rightarrow \Box\Box p.$$

Say that two modalities  $*$  and  $\dagger$  are equivalent in a normal modal logic  $\Lambda$  if  $\Lambda \vdash *p \leftrightarrow \dagger p$ . Up to equivalence in  $\mathbf{S4}$ , there are fourteen modalities; see, e.g., Hughes & Cresswell (1996, p. 55). A useful way of generating them is as follows: let the *dual* of a modality  $*$  be  $\neg*\neg$ , and the *inner negation* of  $*$  be  $*\neg$ . Up to equivalence in  $\mathbf{S4}$ , every modality can be generated from  $-$  (the empty sequence),  $\Box$ ,  $\Box\Diamond$  and  $\Box\Diamond\Box$  by taking duals and inner negations.

We show that in  $QS4$ ,  $\Box\text{Comp}^*$  is inconsistent for every modality  $*$ . We start with the four basic modalities:

LEMMA 4.4.  $\mathbf{S4} \vdash \neg\Box(p \leftrightarrow *\neg p)$  for every modality  $*$   $\in \{-, \Box, \Box\Diamond, \Box\Diamond\Box\}$ .

*Proof.* The case of  $-$  is immediate, and the case of  $\Box$  follows from Lemma 4.1. The remaining two cases can be established by similar arguments using the completeness of  $\mathbf{S4}$  with respect to reflexive and transitive frames. We illustrate this using the case of  $\Box\Diamond$ , and leave the case of  $\Box\Diamond\Box$  to the reader.

Assume for contradiction that there is a Kripke model  $M$  based on a frame  $\langle W, R \rangle$  where  $R$  is reflexive and transitive, and a world  $w$  such that  $M, w \vDash \Box(p \leftrightarrow \Box\Diamond\neg p)$ . Then  $M, v \vDash p \leftrightarrow \Box\Diamond\neg p$  for all  $v \in W$  such that  $wRv$ . We distinguish two cases:

Case 1.  $M, w \vDash \Box p$ . Then  $M, w \vDash p$  by reflexivity of  $R$ ; for the same reason  $M, w \vDash p \leftrightarrow \Box\Diamond\neg p$ , hence  $M, w \vDash \Box\Diamond\neg p$  and finally  $M, w \vDash \Diamond\neg p$ .  $\zeta$ .

Case 2.  $M, w \not\models \Box p$ . There is a  $v \in W$  such that  $wRv$  and  $M, v \models \neg p$ . So  $M, v \models \neg \Box \Diamond \neg p$ . So there is a  $u \in W$  such that  $vRu$  and  $M, u \models \Box p$ . By the transitivity of  $R$ ,  $wRu$ , and so  $M, u \models p \leftrightarrow \Box \Diamond \neg p$ . Applying the reasoning of case 1 to  $u$  instead of  $w$ , a contradiction follows.  $\zeta$ .  $\square$

The next lemma extends this result to the duals of the basic modalities:

LEMMA 4.5. *For any normal modal logic  $\Lambda$  and modality  $*$ , if  $\Lambda \vdash \neg \Box(p \leftrightarrow * \neg p)$ , then  $\Lambda \vdash \neg \Box(p \leftrightarrow \neg * \neg \neg p)$ .*

*Proof.* Assume  $\Lambda \vdash \neg \Box(p \leftrightarrow * \neg p)$ . By classical logic,  $p \leftrightarrow * \neg p$  and  $\neg p \leftrightarrow \neg * \neg p$  are provably equivalent in  $\Lambda$ , and so replaceable preserving theoremhood in  $\Lambda$ . Thus  $\Lambda \vdash \neg \Box(\neg p \leftrightarrow \neg * \neg p)$ . Substituting  $\neg p$  for  $p$ , it follows that  $\Lambda \vdash \neg \Box(\neg \neg p \leftrightarrow \neg * \neg \neg p)$ . Finally, replacing  $\neg \neg p$  by  $p$  produces  $\Lambda \vdash \neg \Box(p \leftrightarrow \neg * \neg \neg p)$ .  $\square$

COROLLARY 4.6. *For every modality  $*$ , QS4 proves the negation of an instance of  $\Box \text{Comp}^*$ .*

*Proof.* If  $*$  is a member or a dual of a member of  $\{-, \Box, \Box \Diamond, \Box \Diamond \Box\}$ , it follows from Lemmas 4.2, 4.4, and 4.5 that  $\text{QS4} \vdash \neg(\Box \text{Russell}^*)$ . If not, there is a modality  $\dagger$  such that  $*$  is equivalent, in **S4**, to  $\dagger \neg$ . As just noted,  $\text{QS4} \vdash \neg(\Box \text{Russell}^\dagger)$ , i.e.,  $\text{QS4} \vdash \neg \exists y \forall x \Box(x \in y \leftrightarrow \dagger \neg x \in x)$ , whence  $\text{QS4} \vdash \neg \exists y \forall x \Box(x \in y \leftrightarrow *x \in x)$ . The claim follows from this with the fact that  $\exists y \forall x \Box(x \in y \leftrightarrow *x \in x)$  is the instance of  $\Box \text{Comp}^*$  for  $\varphi = x \in x$ .  $\square$

The present results leave open the nontrivial consistency of  $(\Box \text{Comp}^*)$  in normal modal logics  $\Lambda$  weaker than **S4**. However, in such a setting, we may well be faced with infinitely many distinct modalities, up to equivalence in  $\Lambda$ , which makes a general assessment of the situation difficult. We therefore return to  $(\Box \text{Comp} \Box)$  for the remainder of this section.

**4.3. Extensions of KD.** Since  $(\Box \text{Comp} \Box)$  is inconsistent in **KT**, let us consider normal modal logics weaker than, or incomparable to, **KT**. We first show that the principle is trivially consistent in logics which fail to prove the axiom  $(D)$ . We then consider three natural extensions of **KD**, namely **KD4**, **KD5**, and **KDB**. We show that  $(\Box \text{Comp} \Box)$  is inconsistent in all of them. The question of the consistency of the principle in **KD** itself will be considered later, and answered in the affirmative.

The consistency of  $(\Box \text{Comp} \Box)$  in normal modal logics not containing  $(D)$  extends to  $(\text{Comp} \Box)$ , and these results are conveniently established together:

PROPOSITION 4.7. *For any normal modal logic  $\Lambda$  such that  $(D) \notin \Lambda$ ,  $Q\Lambda + (\text{Comp} \Box) + (\Box \text{Comp} \Box)$  is consistent.*

*Proof.* If  $(D) \notin \Lambda$ , then  $\Lambda \subseteq \mathbf{Ver}$ , the normal modal logic axiomatized by the axiom  $\Box p$  (see Hughes & Cresswell (1996, p. 67)). Let  $M = \langle W, R, D, V \rangle$  be a model such that  $W$  is a singleton  $\{w\}$ ,  $R$  is the empty relation,  $D(w)$  is nonempty, and  $V(w) = D(w)^2$ . Since  $\Box p$  is valid on the underlying frame, so is  $\Lambda$ .  $(\Box \text{Comp} \Box)$  is trivially validated, and the validity of  $(\text{Comp} \Box)$  follows from the existence of a universal set—indeed, every element is a universal set.  $\square$

From the model construction in the proof, it is clear that the result could be strengthened by adding various further principles, such as  $(BF)$  and  $(CBF)$ , and that the resulting theories are uninteresting.

The inconsistency of  $(\Box \text{Comp} \Box)$  in three standard extensions of **KD** will be based on the following lemma:



LEMMA 4.8.

- KD4**  $\vdash \neg\Box(p \leftrightarrow \Box\neg p)$ .  
**KDB**  $\vdash \neg\Box(p \leftrightarrow \Box\Diamond\neg p)$ .  
**KD5**  $\vdash \Box\neg\Box(p \leftrightarrow \Box\neg p)$ .

*Proof.* These can be established model-theoretically using the corresponding classes of frames, satisfying seriality and, respectively, transitivity, symmetry and euclideaness. The first two cases are similar to the proofs of Lemmas 4.1 and 4.4. For the third case, consider any Kripke model  $M$  based on a serial and euclidean frame  $\langle W, R \rangle$ , and  $w \in W$ . Consider any  $v \in W$  such that  $wRv$ . By euclideaness, for any  $x, y \in W$  such that  $xRy, yRy$ . So  $uRu$  for all  $u \in W$  accessible from  $v$  via a finite sequence of worlds. As in the proof of Lemma 4.1, it follows that  $M, v \models \neg\Box(p \leftrightarrow \Box\neg p)$ , and so  $M, w \models \Box\neg\Box(p \leftrightarrow \Box\neg p)$ .  $\square$

Using Lemma 4.2, the derivability of the negation of the relevant instance of  $(\Box\text{Comp}\Box)$  follows immediately in the first two cases. The third case requires a slight modification of the argument:

LEMMA 4.9. *For any normal modal logic  $\Lambda$  containing  $(D)$ , if  $\Lambda \vdash \Box\neg\Box(p \leftrightarrow \Box\neg p)$ , then  $Q\Lambda \vdash \neg\Box(\Box\text{Russell}\Box)$ .*

*Proof.* Note that  $(\Box\neg p \wedge \Box(q \rightarrow p)) \rightarrow \Box\neg q$  is valid on all Kripke frames, and so provable in **K**. By uniform substitution, **K**  $\vdash (\Box\neg\Box p \wedge \Box(q \rightarrow \Box p)) \rightarrow \Box\neg q$ .

- |     |  |  |
|-----|--|--|
| (1) | $\Box\neg\Box(y \in y \leftrightarrow \Box y \notin y)$  | $(\Lambda)$                                      |
| (2) | $\Box((\forall x\Box(x \in y \leftrightarrow \Box x \notin x) \wedge Ey) \rightarrow \Box(y \in y \leftrightarrow \Box y \notin y))$ | $(\forall 1E), (N)$                              |
| (3) | $(\Box\neg\Box p \wedge \Box(q \rightarrow \Box p)) \rightarrow \Box\neg q$  | established above                                |
| (4) | $\Box\neg(\forall x\Box(x \in y \leftrightarrow \Box x \notin x) \wedge Ey)$   | $(1) - (3)$                                      |
| (5) | $\Box\forall y(Ey \rightarrow \neg\forall x\Box(x \in y \leftrightarrow \Box x \notin x))$   | $(4), (UGL\forall^0)$                            |
| (6) | $\Box\forall yEy$  | $(UE), (N)$                                      |
| (7) | $\Box\forall y\neg\forall x\Box(x \in y \leftrightarrow \Box x \notin x)$  | $(5), (6), (\forall\rightarrow), (N), (\Lambda)$ |
| (8) | $\neg\Box\exists y\forall x\Box(x \in y \leftrightarrow \Box x \notin x)$  | $(7), (D)$ $\square$                             |

Summarizing the results just established, we obtain:

COROLLARY 4.10. *For each  $\Lambda \in \{\mathbf{KD4}, \mathbf{KDB}, \mathbf{KD5}\}$ ,  $Q\Lambda + (\Box\text{Comp}\Box)$  is inconsistent.*

**§5. (Comp $\Box$ ) beyond S5.** In strengthening  $(\text{Comp}\Box)$  to  $(\Box\text{Comp}\Box)$ , we seem to have overshoot our mark; the principle is inconsistent even in the weak normal modal logic **KT**. This directs attention back to  $(\text{Comp}\Box)$ . For although we have seen that in **S5**,  $(\text{Comp}\Box)$  is too weak to yield an interesting theory, this leaves open the possibility that  $(\text{Comp}\Box)$  could yield an interesting theory in logics not included in **S5**, whether (a) in logics stronger than **S5**, or (b) in logics incomparable with **S5**. In this section we consider these two possibilities in order. We first show that  $(\text{Comp}\Box)$  is in fact inconsistent in every strengthening of **S5**. We then consider one example of a natural normal logic incomparable to **S5**, showing that  $(\text{Comp}\Box)$  is consistent in this logic, and that the resulting theory exhibits at least *prima facie* interesting behavior.

**5.1. Inconsistency in proper extensions of S5.** To prove that  $(\text{Comp}\Box)$  is inconsistent in every proper extension of **S5**, we will make use of a theorem due to Scroggs (1951). This theorem entails that for any normal modal logic  $\Lambda$  such that **S5**  $\subsetneq \Lambda$ , there is an  $n < \omega$  such that  $\Lambda \vdash (\text{Alt}_n)$ , where  $(\text{Alt}_n)$  is the following formula:

$$(\text{Alt}_n) \quad \forall_{i \leq n} \Box \left( \bigwedge_{j < i} P_j \rightarrow p_i \right).$$

This formula can be seen as requiring a world to be able to access at most  $n$  worlds, in the following sense: for any Kripke frame  $\langle W, R \rangle$  and  $w \in W$ ,  $M, w \models (\text{Alt}_n)$  for every model  $M$  based on  $\langle W, R \rangle$  just in case there are at most  $n$  worlds  $v \in W$  such that  $wRv$ .

Indeed, we will show more generally that for all  $n < \omega$ , the inconsistency can be derived in  $\text{QKTA}_{\text{Alt}_n} + (\text{Comp}\Box)$ . From this, we will then be able to derive the inconsistency of  $(\text{Comp}\Box)$  in any proper extension of **S5** via Scroggs's theorem.

**PROPOSITION 5.1.** *For all  $n < \omega$ ,  $\text{QKTA}_{\text{Alt}_n} + (\text{Comp}\Box)$  is inconsistent.*

*Proof.* We will rely on the following principles of elementary quantification theory; their derivability in **QK** is easy to demonstrate model-theoretically via the completeness result in Hughes & Cresswell (1996, chap. 16):

- ( $\forall 1'$ )  $\forall y (\forall x \vartheta \rightarrow \vartheta[y/x])$ .
- (*Perm*)  $\forall y_0 \dots \forall y_m \vartheta \rightarrow \forall y_{\pi(0)} \dots \forall y_{\pi(m)} \vartheta$ ,  
where  $\pi$  is a permutation of  $\{0, \dots, m\}$ .
- (*Chain*)  $\forall y_0 \dots \forall y_m \vartheta_0 \rightarrow (\forall y_0 \dots \forall y_m \vartheta_1 \rightarrow \dots \forall y_0 \dots \forall y_m \vartheta_k)$ ,  
where  $\vartheta_0 \rightarrow (\vartheta_1 \rightarrow \dots \vartheta_k)$  is a substitution instance of a tautology.
- (*Iter*)  $\bigwedge_{i \leq m} \forall y_0 \dots \forall y_{i-1} \exists y_i \vartheta_i \rightarrow \exists y_0 \dots \exists y_m \bigwedge_{i \leq m} \vartheta_i$ ,  
where for all  $i < j \leq m$ ,  $y_j$  is not free in  $\vartheta_i$ .

In the last principle,  $\forall y_0 \dots \forall y_{0-1}$  signifies an empty string.

Define the following formulas, for every  $m < \omega$  and variable  $z$ :

$$\begin{aligned} \varphi_m(z) &:= \bigwedge_{i < m} y_i \notin y_i \rightarrow z \notin z, \\ \psi_m &:= x \in y_m \leftrightarrow \Box \varphi_m(x). \end{aligned}$$

We first show that for all  $m < \omega$ ,  $\text{QKT} \vdash \forall y_0 \dots \forall y_m (\bigwedge_{i \leq m} \forall x \psi_i \rightarrow \bigwedge_{i \leq m} y_i \notin y_i)$ , by induction on  $m$ .

The base case of  $m = 0$  is established by the following derivation in **QKT**:

- (1)  $\forall y_0 (\forall x \psi_0 \rightarrow (y_0 \in y_0 \leftrightarrow \Box \varphi_0(y_0)))$  ( $\forall 1'$ ,  $\vartheta = \psi_0$ )
- (2)  $\Box \varphi_0(y_0) \rightarrow y_0 \notin y_0$  (*T*)
- (3)  $(y_0 \in y_0 \leftrightarrow \Box \varphi_0(y_0)) \rightarrow y_0 \notin y_0$  (2)
- (4)  $\forall y_0 ((y_0 \in y_0 \leftrightarrow \Box \varphi_0(y_0)) \rightarrow y_0 \notin y_0)$  (3), (*UG*)
- (5)  $\forall y_0 (\forall x \psi_0 \rightarrow y_0 \notin y_0)$  (1), (4), (*Chain*)

For the induction step, assume as the induction hypothesis that the induction claim holds for some  $m < \omega$ ; we show that any derivation witnessing this can be continued to witness that the induction claim holds for  $m + 1$ :

- (1)  $\forall y_0 \dots \forall y_m (\bigwedge_{i \leq m} \forall x \psi_i \rightarrow \bigwedge_{i \leq m} y_i \notin y_i)$  **IH**
- (2)  $\forall y_0 \dots \forall y_{m+1} (\bigwedge_{i \leq m} \forall x \psi_i \rightarrow \bigwedge_{i \leq m} y_i \notin y_i)$  (1), (*UG*), (*Perm*)
- (3)  $\forall y_{m+1} (\forall x \psi_{m+1} \rightarrow (y_{m+1} \in y_{m+1} \leftrightarrow \Box \varphi_{m+1}(y_{m+1})))$  ( $\forall 1'$ ,  $\vartheta = \psi_{m+1}$ )
- (4)  $\forall y_0 \dots \forall y_{m+1} (\forall x \psi_{m+1} \rightarrow (y_{m+1} \in y_{m+1} \leftrightarrow \Box \varphi_{m+1}(y_{m+1})))$  (2), (*UG*)
- (5)  $\forall y_0 \dots \forall y_{m+1} \left( \bigwedge_{i \leq m+1} \forall x \psi_{m+1} \rightarrow \left( \bigwedge_{i \leq m} y_i \notin y_i \wedge \right. \right.$   
 $\left. \left. (y_{m+1} \in y_{m+1} \leftrightarrow \Box \varphi_{m+1}(y_{m+1})) \right) \right)$  (3), (4), (*Chain*)
- (6)  $\Box \varphi_{m+1}(y_{m+1}) \rightarrow (\bigwedge_{i < m+1} y_i \notin y_i \rightarrow y_{m+1} \notin y_{m+1})$  (*T*)

- (7)  $(\bigwedge_{i \leq m} y_i \notin y_i \wedge (y_{m+1} \in y_{m+1} \leftrightarrow \Box \varphi_{m+1}(y_{m+1}))) \rightarrow y_{m+1} \notin y_{m+1}$  (6)
- (8)  $\forall y_0 \dots \forall y_{m+1} ((\bigwedge_{i \leq m} y_i \notin y_i \wedge (y_{m+1} \in y_{m+1} \leftrightarrow \Box \varphi_{m+1}(y_{m+1}))) \rightarrow y_{m+1} \notin y_{m+1})$  (7), (UG)
- (9)  $\forall y_0 \dots \forall y_{m+1} (\bigwedge_{i \leq m+1} \forall x \psi_{m+1} \rightarrow y_{m+1} \notin y_{m+1})$  (5), (8), (Chain)

Since the induction claim holds for all  $m < \omega$ , it holds in particular for  $n$ . We show how to continue any derivation witnessing this to produce a derivation of an inconsistency in  $QKTA_{\text{alt}_n} + (\text{Comp}\Box)$ . Here, e.g.,  $(2_{i \leq n})$  abbreviates a sequence of lines  $(2_0), \dots, (2_n)$ .

- (1)  $\forall y_0 \dots \forall y_n (\bigwedge_{i \leq n} \forall x \psi_i \rightarrow \bigwedge_{i \leq n} y_i \notin y_i)$  established above
- (2 <sub>$i \leq n$</sub> )  $\forall y_i (\forall x \psi_i \rightarrow (y_i \in y_i \leftrightarrow \Box \varphi_i(y_i)))$  ( $\forall 1', \vartheta = \psi_i$ )
- (3 <sub>$i \leq n$</sub> )  $\forall y_0 \dots \forall y_n (\forall x \psi_i \rightarrow (y_i \in y_i \leftrightarrow \Box \varphi_i(y_i)))$  (2 <sub>$i$</sub> ), (UG), (Perm)
- (4)  $\forall y_0 \dots \forall y_n (\bigwedge_{i \leq n} \forall x \psi_i \rightarrow (\bigwedge_{i \leq n} y_i \notin y_i \rightarrow \bigwedge_{i \leq n} \neg \Box \varphi_i(y_i)))$  (3 <sub>$i \leq n$</sub> ), (Chain)
- (5)  $\forall y_0 \dots \forall y_n (\bigwedge_{i \leq n} \forall x \psi_i \rightarrow \bigwedge_{i \leq n} \neg \Box \varphi_i(y_i))$  (1), (4) (Chain)
- (6)  $\bigvee_{i \leq n} \Box (\bigwedge_{j < i} y_j \notin y_j \rightarrow y_i \notin y_i)$  (Alt<sub>n</sub>)
- (7)  $\bigvee_{i \leq n} \Box \varphi_i(y_i)$  =(6)
- (8)  $\forall y_0 \dots \forall y_n \bigvee_{i \leq n} \Box \varphi_i(y_i)$  (7), (UG)
- (9)  $\forall y_0 \dots \forall y_n \neg \bigwedge_{i \leq n} \forall x \psi_i$  (5), (8), (Chain)
- (10)  $\neg \exists y_0 \dots \exists y_n \bigwedge_{i \leq n} \forall x (x \in y_i \leftrightarrow \Box \varphi_i(x))$  (9)
- (11 <sub>$i \leq n$</sub> )  $\exists y_i \forall x (x \in y_i \leftrightarrow \Box \varphi_i(x))$  (Comp $\Box$ )
- (12)  $\bigwedge_{i \leq n} \forall y_0 \dots \forall y_{i-1} \exists y_i \forall x (x \in y_i \leftrightarrow \Box \varphi_i(x))$  (11 <sub>$i \leq n$</sub> ), (UG)
- (13)  $\exists y_0 \dots \exists y_n \bigwedge_{i \leq n} \forall x (x \in y_i \leftrightarrow \Box \varphi_i(x))$  (12), (Iter)
- (14)  $\perp$  (10), (13) □

With Scroggs’s theorem, we obtain:

**COROLLARY 5.2.** *For any normal modal logic  $\Lambda$  such that  $\mathbf{S5} \subsetneq \Lambda$ ,  $Q\Lambda + (\text{Comp}\Box)$  is inconsistent.*

**5.2. KDDc.** We now turn to a normal modal logic incomparable in strength with **S5**: **KDDc**, the normal modal logic axiomatized by the axioms (D) and the converse of (D):

$$(D_c) \ \Diamond p \rightarrow \Box p.$$

The system **KDDc** is typical in allowing the distribution of  $\Box$  across all Boolean operators: for example in this system  $\neg \Box \varphi$  is equivalent to  $\Box \neg \varphi$  and  $\Box(\varphi \vee \psi)$  is equivalent to  $\Box \varphi \vee \Box \psi$ . This makes the theory particularly interesting as a fictionalist theory: although not everything which is true in the fiction is true, the fiction is complete in the sense that  $\Box \varphi \vee \Box \neg \varphi$  is a theorem schema. If we add (BF) and (CBF), this property of movement extends to the quantifiers as well:  $\exists x \Box \varphi$  is now equivalent to  $\Box \exists x \varphi$ . Using this further property, together with one application of necessitation, every instance of  $(\Box \text{Comp}\Box)$  is derivable from a corresponding instance of  $(\text{Comp}\Box)$ . Whereas in logics which prove (T),  $(\Box \text{Comp}\Box)$  is at least as strong as  $(\text{Comp}\Box)$ , in  $QKDDc + (\text{BF}) + (\text{CBF})$ , the situation is reversed, so that the consistency of  $QKDDc + (\Box \text{Comp}\Box) + (\text{BF}) + (\text{CBF})$  is demonstrated by:

**THEOREM 5.3.**  *$QKDDc + (\text{Comp}\Box) + (\text{BF}) + (\text{CBF})$  is consistent.*

We split up the proof into several lemmas. Consistency will be established by constructing a series of Kripke models; although none of them has a world which verifies every

theorem of  $QKDDc + (Comp\Box)$ , every such theorem is true in some world of one of the models, and this suffices for consistency.

For each  $n < \omega$ , let  $M_n = \langle W_n, R_n, D_n, V_n \rangle$ , with the components defined as follows: Let  $W_n = \{0, \dots, n\}$ ,  $R_n = \{\langle w + 1, w \rangle : w < n\}$  and  $D_n(w) = \mathbb{N}$  for all  $w \leq n$ . To define  $V_n$ , let  $X$  be the set of pairs  $\langle \varphi, \bar{s} \rangle$  such that for some  $m < \omega$ ,  $\varphi$  is an  $\mathcal{L}_1^\Box(\in)$ -formula whose free variables are all contained in  $\{x_0, \dots, x_m\}$  and  $\bar{s}$  is a sequence of  $m$  elements of  $\mathbb{N}$ ; informally,  $X$  can be understood as the set of conditions on a single variable definable in  $\mathcal{L}_1^\Box(\in)$  with parameters from  $\mathbb{N}$ . The idea behind the definition of  $V_n$  is to let it agree with  $V_w$ , for any  $w < n$ , on all  $v \leq w$ , with  $V_{n+1}(n + 1)$  interpreting  $\in$  such that for each condition in  $X$ , there is an element which witnesses the existence of a set containing exactly the elements satisfying this condition in  $M_n, n$ . We will make this precise in a general way which allows us to use the present model construction for some additional consistency arguments below. Let  $f$  be a surjective partial function from  $\mathbb{N}$  to  $X$ , and  $Z \subseteq \mathbb{N} \times (\mathbb{N} \setminus \text{dom}(f))$ ; informally, the elements in the domain of  $f$  will serve as witnesses for  $(Comp\Box)$  at all worlds, and suitable choices of  $Z$  will be used later to verify or falsify various principles to be shown independent from  $QKDDc + (Comp\Box)$ . For any  $\langle \varphi, \bar{s} \rangle \in X$  and  $r \in \mathbb{N}$ , we will write  $M_n, w \models \langle \varphi, \bar{s} \rangle[r]$  to state that  $M_n, w, a \models \varphi$ , where  $a(x_0) = r$  and  $\langle a(x_1), \dots, a(x_m) \rangle = \bar{s}$ . (Note that this is well-defined: the choice of  $a$  does not matter as long as it satisfies the stated conditions, since there are no free variables in  $\varphi$  apart from  $x_0, \dots, x_m$ .)  $V_n$  is now defined recursively. The base case is:

$$V_0(0) = D^2.$$

For the recursion step, let:

$$\begin{aligned} V_{n+1}(w) &= V_n(w) \text{ for all } w \leq n, \\ V_{n+1}(n + 1) &= \{ \langle r, t \rangle : t \in \text{dom}(f) \text{ and } M_n, n \models f(t)[r] \} \cup Z. \end{aligned}$$

It will be useful to work with a variant calculus for  $QKDDc$ . Recall that for any normal modal logic  $\Lambda$ ,  $Q\Lambda$  is defined as the set of  $\mathcal{L}_1^\Box(\in)$ -formulas derivable in an axiomatic calculus which includes an axiom schema  $(\Lambda)$  whose instances are the substitution instances of  $\Lambda$ -theorems. Call this calculus  $CQ\Lambda$ . Letting  $\Phi$  be the set of proposition letters of  $\mathcal{L}_0^\Box$ , we can understand the substitution operations involved in the schema  $(\Lambda)$  as the functions  $\sigma : \Phi \rightarrow \mathcal{L}_1^\Box$ . Extending this to complex  $\mathcal{L}_0^\Box$ -formulas in the obvious way, the axiom schema can be stated more carefully as follows:

$$(\Lambda) \sigma(\varphi) \quad \text{where } \sigma : \Phi \rightarrow \mathcal{L}_1^\Box \text{ and } \varphi \in \Lambda.$$

Our first lemma tells us that in case  $\Lambda = \mathbf{KA}_1 \dots \mathbf{A}_n$ , this axiom schema can be restricted to the substitution instances of the axioms of  $\mathbf{KA}_1 \dots \mathbf{A}_n$ . Letting  $\text{Taut}$  be the set of tautologies, this axiom schema can be formulated more carefully as follows:

$$(\mathbf{KA}_1 \dots \mathbf{A}_n)_R \sigma(\varphi) \quad \text{where } \sigma : \Phi \rightarrow \mathcal{L}_1^\Box \text{ and } \varphi \in \text{Taut} \cup \{(K), A_1, \dots, A_n\}.$$

Let  $C_R Q\mathbf{KA}_1 \dots \mathbf{A}_n$  be the axiomatic calculus which consists of the axiom schemas and rules of  $CQ\mathbf{KA}_1 \dots \mathbf{A}_n$ , except for restricting the axiom schema  $(\mathbf{KA}_1 \dots \mathbf{A}_1)$ —consisting of the substitution instances of the theorems of  $\mathbf{KA}_1 \dots \mathbf{A}_1$ —to  $(\mathbf{KA}_1 \dots \mathbf{A}_1)_R$ .

**LEMMA 5.4.** *For any  $\mathcal{L}_0^\Box$ -formulas  $A_1 \dots A_n$ , each instance of the axiom schema  $(\mathbf{KA}_1 \dots \mathbf{A}_n)$  is derivable in  $C_R Q\mathbf{KA}_1 \dots \mathbf{A}_n$ .*

*Proof.* We proceed by induction on the complexity of proofs in the propositional calculus used to define  $\mathbf{KA}_1 \dots \mathbf{A}_n$ , showing that for each  $\varphi$  derivable in this calculus,  $\sigma(\varphi)$

is derivable in  $C_R QKA_1 \dots A_n$  for all  $\sigma : \Phi \rightarrow \mathcal{L}_1^\square$ . The base case is immediate: if  $\varphi$  is an axiom of the calculus used to define  $KA_1 \dots A_n$ , then  $\sigma(\varphi)$  is an instance of  $(KA_1 \dots A_nR)$ .

For the rule of modus ponens, assume that  $\varphi$  is obtained by applying modus ponens to  $\psi$  and  $\psi \rightarrow \varphi$ . By the induction hypothesis,  $\sigma(\psi)$  and  $\sigma(\psi \rightarrow \varphi)$  are derivable in  $C_R QKA_1 \dots A_n$ . Since the latter calculus also includes the rule of modus ponens and  $\sigma(\psi \rightarrow \varphi) = \sigma(\psi) \rightarrow \sigma(\varphi)$ ,  $\sigma(\varphi)$  is derivable in it as well. The induction step for necessitation is similarly immediate.

For the rule of uniform substitution, assume that  $\varphi$  is obtained by applying uniform substitution to  $\psi$ . Thus  $\varphi = \tau(\psi)$  for some  $\tau : \Phi \rightarrow \mathcal{L}_0^\square$ . By the induction hypothesis, for every  $\sigma : \Phi \rightarrow \mathcal{L}_1^\square$ ,  $\sigma(\psi)$  is derivable in  $C_R QKA_1 \dots A_n$ . Thus for every  $\sigma : \Phi \rightarrow \mathcal{L}_1^\square$ ,  $\sigma \circ \tau(\psi) = \sigma(\varphi)$  is derivable in  $C_R QKA_1 \dots A_n$ , as required.  $\square$

Let an  $\mathcal{L}_1^\square$ -formula  $\varphi$  be valid in a world  $w$  of a model  $M$ , written  $M, w \models \varphi$  if  $M, w, a \models \varphi$  for all assignment functions  $a$ . For the variant calculus of  $QKDDc$ , we can show that aside from the axiom  $(D)$  and the rule of necessitation, all other axioms are valid in all worlds of the models constructed above, and all other rules preserve validity in a world in them:

LEMMA 5.5. *Let  $w \leq n < \omega$ . All axioms of  $C_R QKDDc$  apart from  $(D)$  are valid in  $M_n, w$ , and all rules of  $C_R QKDDc$  apart from  $(N)$  preserve truth in  $M_n, w$ .*

*Proof.* Routine; see Hughes & Cresswell (1996, pp. 123 & 295) for  $(D_c)$  and the remaining principles, respectively.  $\square$

Likewise,  $(Comp\Box)$  is valid in all worlds of these models:

LEMMA 5.6. *For all  $n < \omega$ , every instance of  $(Comp\Box)$  is valid in  $M_n, w$  for all  $w \leq n$ .*

*Proof.* We proceed by induction on  $n$ .

For the base case, let  $n = 0$ , and so  $w = 0$  as well. Since  $M_0, 0 \models \Box\varphi$  for all  $\varphi$ , it suffices for there to be some  $t \in \mathbb{N}$  such that  $\langle r, t \rangle \in V_0(0)$  for all  $r \in \mathbb{N}$ . By construction of  $V_0(0)$ , this is in fact the case for all  $t \in \mathbb{N}$ .

For the induction step, assume as the induction hypothesis that for all  $w \leq n$ , every instance of  $(Comp\Box)$  is valid in  $M_n, w$ . It follows that every instance of  $(Comp\Box)$  is valid in  $M_{n+1}, w$ , for all  $w \leq n$ : generally, for every formula  $\varphi$  and  $w \leq n$ ,  $M_n, w \models \varphi$  iff  $M_{n+1}, w \models \varphi$ . (This is routine to establish based on the observation that  $M_n$  and  $M_{n+1}$  are isomorphic when restricted to the worlds accessible via finite paths from any such  $w$ .) So consider an instance of  $(Comp\Box)$ ; without loss of generality, we assume this to be  $\exists y \forall x_0 (x_0 \in y \leftrightarrow \Box\varphi)$ , where the free variables in  $\varphi$  are contained in  $\{x_0, \dots, x_m\}$ . It suffices to show that this formula is valid in  $M_{n+1}, n+1$ .

Let  $a$  be any assignment function. Since  $f$  is surjective on  $X$ , there is a  $t \in \mathbb{N}$  such that  $f(t) = \langle \varphi, \langle a(x_1), \dots, a(x_m) \rangle \rangle$ . We show that  $t$  witnesses the existential claim, i.e., that for all  $r \in \mathbb{N}$ ,  $M_{n+1}, n+1, a[t/y][r/x_0] \models x_0 \in y \leftrightarrow \Box\varphi$ :

$$\begin{aligned} M_{n+1}, n+1, a[t/y][r/x_0] \models x_0 \in y \text{ iff,} \\ \langle r, t \rangle \in V_{n+1}(n+1) \text{ iff,} \\ M_n, n \models f(t)[r] \text{ iff,} \\ M_n, n \models \langle \varphi, \langle a(x_1), \dots, a(x_m) \rangle \rangle[r] \text{ iff,} \\ M_n, n, a[r/x_0] \models \varphi \text{ iff,} \\ M_{n+1}, n, a[r/x_0] \models \varphi \text{ iff,} \\ M_{n+1}, n+1, a[t/y][r/x_0] \models \Box\varphi. \end{aligned}$$

Thus all instances of  $(\text{Comp}\Box)$  are valid in  $M_{n+1}$ ,  $n + 1$ . Combining this with the earlier observation, it follows that all such instances are valid in  $M_{n+1}$ ,  $w$  for all  $w \leq n + 1$ .  $\square$

While substitution instances of  $(D)$  may not be valid in  $M_n$ ,  $0$ , they are valid in all other worlds:

LEMMA 5.7. *If  $0 < w \leq n < \omega$  then every substitution instance of  $(D)$  is valid in  $M_n$ ,  $w$ .*

*Proof.* Routine, using the fact that  $wR_n w - 1$ .  $\square$

For the following lemma, let  $C_R QKDDc + (\text{Comp}\Box)$  be the calculus obtained from  $C_R QKDDc$  by adding the axiom schema  $(\text{Comp}\Box)$ . Let a  $\mathcal{L}_1^\Box$ -formula  $\varphi$  be *derivable in  $C_R QKDDc + (\text{Comp}\Box)$  using  $n$  applications of necessitation  $(N)$*  if there is a proof of  $\varphi$  in  $C_R QKDDc + (\text{Comp}\Box)$  such that for all except at most  $n$  lines of the proof, the relevant formula is an axiom of the calculus or obtainable from formulas occurring earlier in the proof using a rule other than  $(N)$ .

LEMMA 5.8. *If  $n < w \leq m < \omega$  and  $\varphi$  is a  $\mathcal{L}_1^\Box$ -formula derivable in  $C_R QKDDc + (\text{Comp}\Box)$  using  $n$  applications of the rule of necessitation  $(N)$  then  $\varphi$  is valid in  $M_m$ ,  $w$ .*

*Proof.* By induction on  $n$ . For brevity, we will take derivability and similar notions to be tacitly relativized to  $C_R QKDDc + (\text{Comp}\Box)$ .

For the base case of  $n = 0$ , consider any  $\mathcal{L}_1^\Box$ -formula  $\varphi$  derivable using 0 applications of  $(N)$ . A subinduction on the length of a proof witnessing this shows that every line in it is valid in  $M_w$ ,  $m$ ; the relevant steps are immediate by Lemmas 5.5, 5.6, and 5.7 (note that  $n = 0 < w$ ).

For the induction step, assume that the claim holds for  $n$ . Let  $\varphi$  be derivable using  $n + 1$  applications of  $(N)$ , and  $n + 1 < w \leq m < \omega$ . If  $\varphi$  is derivable using  $n$  applications of  $(N)$ , the claim is immediate by the induction hypothesis. If not, consider a proof of  $\varphi$  and let  $k$  be the number of the first line which is neither an axiom nor obtainable from a formula occurring earlier in the proof using any rule other than  $(N)$ . Then all lines before the  $k$ th line are derivable using  $n$  applications of  $(N)$  and so by the induction hypothesis is valid in  $M_m$ ,  $w$ . In particular, the line from which the  $k$ th line is obtained by  $(N)$  is derivable using  $n$  applications of  $(N)$ , and so by the induction hypothesis is valid in  $M_m$ ,  $w - 1$ ; thus the  $k$ th line is valid in  $M_m$ ,  $w$ . As in the base case, a subinduction using the same lemmas now shows that every line after the  $k$ th line is valid in  $M_m$ ,  $w$ .  $\square$

The consistency of  $QKDDc + (\text{Comp}\Box)$  is straightforward to derive from this lemma:

*Proof of Theorem 5.3.* Assume for contradiction that  $QKDDc + (\text{Comp}\Box) \vdash \perp$ . By Lemma 5.4,  $\perp$  is derivable in  $C_R QKDDc + (\text{Comp}\Box)$ . As every derivation is of finite length, there is some  $n < \omega$  such that  $\perp$  is derivable in  $C_R QKDDc + (\text{Comp}\Box)$  using  $n$  applications of  $(N)$ . By Lemma 5.8, it follows that  $\perp$  is valid in  $M_{n+1}$ ,  $n + 1$ .  $\not\perp$ .

As the models used in this proof employ constant domains,  $(BF)$  and  $(CBF)$  are also valid at every world.  $\square$

In contrast to the models used to show the consistency of the modal set theories considered earlier in this article (see §3 and §4.3), the models used to demonstrate the consistency of  $QKDDc + (\text{Comp}\Box)$  do not have an intuitively trivial structure. Without considering further set-theoretic axioms, it is therefore the most promising system considered in this article. Yet, it also suffers from serious limitations. In particular, the theory does not prove

natural principles about the closure of sets under Boolean operations, as we will now demonstrate. Consider the following three principles:

- (Union)  $\forall x \forall y \exists z \forall w (w \in z \leftrightarrow (w \in x \vee w \in y))$ ,
- (Con)  $\forall x \forall y \exists z \forall w (w \in z \leftrightarrow (w \in x \wedge w \in y))$ ,
- (Neg)  $\forall x \exists y \forall z (z \in y \leftrightarrow z \notin x)$ .

The first of these is most familiar, as the axiom of union; the second is usually obtained as an instance of separation; and the third is a common principle in set theories with a universal set. None of them is proven by  $QKDDc + (Comp\Box)$ , nor are their negations:

PROPOSITION 5.9.  $QKDDc + (Comp\Box)$  proves neither (Union), (Con), (Neg), nor any of their negations.

*Proof.* We give the details for (Neg) and its negation; the other cases are similar. Consider the parameters  $f$  and  $Z$  in the construction of the models used in the proof of Theorem 5.3. Note that the definition of  $M_0$  is independent of these parameters. Let  $\Sigma$  be the set of  $S \subseteq \mathbb{N}$  such that for some  $\langle \varphi, \bar{s} \rangle \in X$ ,  $S = \{r \in \mathbb{N} : M_0, 0 \models \langle \varphi, \bar{s} \rangle[r]\}$ .  $\Sigma$  is closed under complement: if  $S \in \Sigma$ , witnessed by  $\langle \varphi, \bar{s} \rangle$ , then  $\mathbb{N} \setminus S \in \Sigma$  is witnessed by  $\langle \neg\varphi, \bar{s} \rangle$ . Thus on the one hand, choosing  $f$  to be a bijection from  $\mathbb{N}$  to  $X$  guarantees that  $M_1, 1 \models (Neg)$ . On the other hand, choosing  $f$  such that  $\mathbb{N} \setminus \text{dom}(f)$  is a singleton  $\{t\}$  and  $Z = \{\langle r, t \rangle : r \in S\}$  for some  $S \notin \Sigma$  (which must exist for cardinality reasons) guarantees that  $M_1, 1 \models \neg(Neg)$ . □

**§6. Conclusion.** We began our investigation with the modal principle  $(Comp\Box)$ . The most commonly studied modal logics are sublogics of **S5**. In all such logics the principle  $(Comp\Box)$  is consistent, but on its own, it leads to a weak set theory. In response to this observation, we considered strengthening the comprehension principle, to  $(\Box Comp\Box)$ . This alternative principle turned out to be inconsistent in all extensions of the modal system **T**, as well as in a number of standard extensions of **KD**. These results led us to examine  $(Comp\Box)$  in proper extensions of **S5**, where we showed it to be inconsistent. We then considered it and  $(\Box Comp\Box)$  in the system **KDDc**, where we showed them to be consistent, although in certain ways also limited in strength.

Overall, we have seen that while in some sense modalities can save naive set theory along the lines of  $(Comp\Box)$ , they probably don't produce interesting set theories on their own: as far as the modal systems we have investigated are concerned,  $(Comp\Box)$  seems to require supplementation by further set-theoretic axioms, most likely themselves modal, to yield any interesting set theory. We have not considered such additional axioms; the prospects for constructing interesting set theories based on  $Q\Lambda + (\Box Comp\Box)$  and additional modal set-theoretic axioms, for some normal modal logic  $\Lambda$ , is a wide open field for future research.

There are a number of interesting further questions left open by the results established here. One concerns a variant of  $(\Box Comp\Box)$  in which the universal quantifier and the outer modal operator are interchanged:

$$\exists y \Box \forall x (x \in y \leftrightarrow \Box \varphi). \tag{BoxCompBox'}$$

In some settings, such as **QKT**, it is not hard to see how to adapt the proof of the inconsistency of  $(\Box Comp\Box)$  to show the inconsistency of  $(\Box Comp\Box')$ , but in other settings, such as **KD4**, the consistency of this variant principle seems harder to settle. Of course, there are many more ways of modalizing set comprehension one might investigate, and one might take the negative results established here to motivate exploring such alternatives. Indeed,

one might also take the present findings to motivate returning to Krajíček's comprehension principle (MCA), mentioned above, which leads to a strong theory in **KT**. It would be very interesting to settle the consistency question of the resulting modal set theory.

Further open questions arise from normal modal logics incomparable in strength to **S5** apart from **KDDc**. A particularly interesting example is **S4M**, the normal extension of **S4** by the axiom

$$(M) \quad \Box \Diamond p \rightarrow \Diamond \Box p.$$

In frames validating **S4M**, each world can access a world which can only access itself; it follows that there can be no model of  $(\text{Comp}\Box)$  based on a frame validating **S4M**. Yet this does not mean that  $QS4M + (\text{Comp}\Box)$  is inconsistent, since  $QS4M$  is incomplete with respect to the class of such models; see Hughes & Cresswell (1996, pp. 265–270). If  $QS4M + (\text{Comp}\Box)$  turns out to be consistent, it would also be interesting to consider certain normal extensions of **S4M**, in particular **K2**, which adds the converse of  $M$ .

A final set of open questions concerns the strength of  $QKDDc + (\text{Comp}\Box)$ . Even though it was shown above not to prove certain basic set-theoretic principles, much about its strength is left open. One way of measuring its strength is in terms of interpretability: which standard theories, such as Peano Arithmetic, can be interpreted in it? As usual, different notions of interpretability are available, and in the present setting, we can additionally choose between interpreting the relevant theory in the full modal system or its nonmodal fragment. The nonmodal fragment may also be characterized directly, analogously to the characterization of the nonmodal fragment of  $QS5 + (\text{Comp}\Box)$  as  $(F) + (CF)$  in Theorem 3.1. In fact, none of the results established above immediately rules out the hypothesis that the nonmodal fragment of  $QKDDc + (\text{Comp}\Box)$  is  $(F) + (CF)$  as well.

**§7. Acknowledgments.** We thank audiences at a workshop on Modality and Set Theory in Oslo, the Stanford Faculty Club, the Vienna Center for Logic and Algorithms, the Kurt Gödel Society in Vienna, the Berkeley Logic Colloquium, and the New York Philosophy of Logic Group for questions and comments. We are also grateful to six reviewers of this journal for their numerous detailed suggestions, which have contributed greatly to the quality of the article. Each set of comments helped us improve the motivation for and presentation of our results. Some reviewers also provided us with ideas which led to new or stronger results. As suggested by the editor, Peter Koellner, we acknowledge these contributions in footnotes throughout the article. We would also like to thank Reviewer #5 in particular for three sets of extremely thorough comments on presentational matters, which have led to substantial improvements in the clarity of the article. Tiankai Liu was supported by a NSF Postdoctoral Research Fellowship under grant no. DMS-1402411.

## BIBLIOGRAPHY

- Aczel, P. & Feferman, S. (1980). Consistency of the unrestricted abstraction principle using an intensional equivalence operator. In Seldin, J. P. and Hindley, J. R., editors. *To H. B. Curry: Essays on Combinatory Logic, Lambda Calculus and Formalism*. New York: Academic Press, pp. 67–98.
- Bacon, A. (2015). Can the classical logician avoid the revenge paradoxes? *Philosophical Review*, **124**(3), 299–352.



- Balaguer, M. (2008). Fictionalism in the philosophy of mathematics. In Zalta, E. N., editor. *Stanford Encyclopedia of Philosophy*. Available at <https://plato.stanford.edu/archives/sum2015/entries/fictionalism-mathematics/>.
- Blass, A. (1990). Infinitary combinatorics and modal logic. *The Journal of Symbolic Logic*, **55**(2), 761–778.
- Boolos, G. (1995). *The Logic of Provability*. Cambridge: Cambridge University Press.
- Cantini, A. (2009). Paradoxes, self-reference and truth in the 20th century. In Woods, J. and Gabbay, D. M. editors. *Handbook of the History of Logic, Volume 5: Logic from Russell to Church*. Amsterdam: Elsevier, pp. 875–1013.
- Feferman, S. (1984). Toward useful type-free theories, I. *The Journal of Symbolic Logic*, **49**, 75–111.
- Field, H. (1980). *Science Without Numbers*. Princeton, NJ: Princeton University Press.
- Field, H. (1989). *Realism, Mathematics & Modality*. Oxford: Basil Blackwell.
- Field, H., Lederman, H., & Øgaard, T. F. (2017). Prospects for a naive theory of classes. *Notre Dame Journal of Formal Logic*, **58**(4), 461–506.
- Fine, K. (1978). Model theory for modal logic – Part II. The elimination of *de re* modality. *Journal of Philosophical Logic*, **7**, 277–306.
- Fine, K. (1981). First-order modal theories I—sets. *Noûs*, **15**, 177–205.
- Fitch, F. B. (1942). A basic logic. *The Journal of Symbolic Logic*, **7**(3), 105–114.
- Fitch, F. B. (1948). An extension of basic logic. *The Journal of Symbolic Logic*, **13**(2), 95–106.
- Fitch, F. B. (1963). The system  $C\Delta$  of combinatory logic. *The Journal of Symbolic Logic*, **28**(1), 87–97.
- Fitch, F. B. (1966). A consistent modal set theory (Abstract). *The Journal of Symbolic Logic*, **31**, 701.
- Fitch, F. B. (1967a). A complete and consistent modal set theory. *The Journal of Symbolic Logic*, **32**, 93–103.
- Fitch, F. B. (1967b). A theory of logical essences. *The Monist*, **51**, 104–109.
- Fitch, F. B. (1970). Correction to a paper on modal set theory. *The Journal of Symbolic Logic*, **35**, 242.
- Fitting, M. (2003). Intensional logic—beyond first order. In Hendricks, V. F. and Malinowski, J., editors. *Trends in Logic*. Dordrecht: Springer, pp. 87–108.
- Forster, T. E. (1995). *Set Theory with a Universal Set* (second edition). Oxford: Clarendon Press.
- Fritz, P., & Goodman, J. (2016). Higher-order contingentism, part 1: Closure and generation. *Journal of Philosophical Logic*, **45**(6), 645–695.
- Gallin, D. (1975). *Intensional and Higher-Order Modal Logic*. Amsterdam: North-Holland.
- Gilmore, P. C. (1967). The consistency of a positive set theory. Technical Report RC-1754, IBM Research Report.
- Goodman, N. (1990). Topological models of epistemic set theory. *Annals of Pure and Applied Logic*, **46**, 147–167.
- Hamkins, J. D. (2003). A simple maximality principle. *The Journal of Symbolic Logic*, **68**(2), 527–550.
- Hamkins, J. D. & Löwe, B. (2008). The modal logic of forcing. *Transactions of the American Mathematical Society*, **360**(4), 1793–1817.
- Hamkins, J. D. & Woodin, W. H. (2005). The necessary maximality principle for ccc forcing is equiconsistent with a weakly compact cardinal. *Mathematical Logic Quarterly*, **51**(5), 493–498.

- Hellman, G. (1989). *Mathematics Without Numbers: Towards a Modal-Structural Interpretation*. Oxford: Oxford University Press.
- Hughes, G. E. & Cresswell, M. J. (1996). *A New Introduction to Modal Logic*. London: Routledge.
- Kaye, R. (1993). Review of Forster, T. E., set theory with a universal set: Exploring an untyped universe. *Notre Dame Journal of Formal Logic*, **34**, 302–309.
- Krajíček, J. (1987). A possible modal formulation of comprehension scheme. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, **33**, 461–480.
- Krajíček, J. (1988). Some results and problems in the modal set theory MST. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, **34**, 123–134.
- Kripke, S. A. (1959). A completeness theorem in modal logic. *The Journal of Symbolic Logic*, **24**, 1–14.
- Kripke, S. A. (1975). Outline of a theory of truth. *Journal of Philosophy*, **72**(19), 690–716.
- Linnebo, Ø. (2010). Pluralities and sets. *The Journal of Philosophy*, **107**(3), 144–164.
- Linnebo, Ø. (2013). The potential hierarchy of sets. *The Review of Symbolic Logic*, **6**(2), 205–228.
- Parsons, C. (1983). Sets and modality. *Mathematics in Philosophy*. Ithaca: Cornell University Press, pp. 298–341.
- Rosen, G. (2001). Nominalism, naturalism, epistemic relativism. *Noûs*, **35**(s15), 69–91.
- Rundle, B. (1969). Review of Frederic B. Fitch, “A theory of logical essences” and “A complete and consistent modal set theory”. *The Journal of Symbolic Logic*, **34**, 125.
- Scott, D. (2010). *A Boolean-Valued Modal Set Theory*. Unpublished Oxford Lectures. Available at: [www.homepages.inf.ed.ac.uk/als/ScottInScotland/ScottLect3.pdf](http://www.homepages.inf.ed.ac.uk/als/ScottInScotland/ScottLect3.pdf).
- Scroggs, S. J. (1951). Extensions of the Lewis system S5. *The Journal of Symbolic Logic*, **16**(2), 112–120.
- Shapiro, S. (editor) (1985). *Intensional Mathematics*. Amsterdam: North-Holland.
- Smullyan, R. & Fitting, M. (1996). *Set Theory and the Continuum Problem*. Oxford: Oxford University Press.
- Solovay, R. M. (1976). Provability interpretations of modal logic. *Israel Journal of Mathematics*, **25**(3–4), 287–304.
- Studd, J. (2013). The iterative conception of set. A (bi-)modal axiomatisation. *Journal of Philosophical Logic*, **42**(5), 697–725.
- Williamson, T. (2013). *Modal Logic as Metaphysics*. Oxford: Oxford University Press.
- Yablo, S. (2001). Go figure: A path through fictionalism. *Midwest Studies in Philosophy*, **25**(1), 72–102.
- Zalta, E. N. (1988). *Intensional Logic and the Metaphysics of Intentionality*. Cambridge, MA: MIT Press.

IFIKK

UNIVERSITY OF OSLO  
PO BOX 1020 BLINDERN  
0315 OSLO, NORWAY

*E-mail:* peter.fritz@ifikk.uio.no

DEPARTMENT OF PHILOSOPHY  
PRINCETON UNIVERSITY  
PRINCETON, NJ 08544, USA

*E-mail:* harvey.lederman@princeton.edu

DEPARTMENT OF MATHEMATICS  
UNIVERSITY OF UTAH  
SALT LAKE CITY, UT 84112, USA  
*E-mail:* tliu@math.utah.edu

VISITING SCHOLAR  
UNIVERSITY OF CALIFORNIA, BERKELEY  
BERKELEY, CA 94720, USA  
*E-mail:* dana.scott@cs.cmu.edu