Overgeneration in the Higher Infinite

SALVATORE FLORIO
University of Birmingham
s.florio@bham.ac.uk

LUCA INCURVATI
ILLC, University of Amsterdam
l.incurvati@uva.nl

This is a preprint of an article due to appear in Gil Sagi and Jack Woods (eds.), *The Semantic Conception of Logic*, Cambridge University Press.

Abstract. The Overgeneration Argument is a prominent objection against the model-theoretic account of logical consequence for second-order languages. In previous work we have offered a reconstruction of this argument which locates its source in the conflict between the neutrality of second-order logic and its alleged entanglement with mathematics. Some cases of this conflict concern small large cardinals. In this article, we show that in these cases the conflict can be resolved by moving from a set-theoretic implementation of the model-theoretic account to one which uses higher-order resources.

George Boolos once remarked that first-order Zermelo-Fraenkel set theory is 'a satisfactory if not ideal theory of infinite numbers' (Boolos 1971, p. 229). An ideal theory, he added, 'would decide the continuum hypothesis, at least'. Despite its success in capturing facts about the cardinality of infinite collections, ZF, even when augmented with the Axiom of Choice, fails to settle many natural questions about which sizes there are.[1] In particular, ZFC cannot decide whether a number of interesting cardinality notions expressible in its language are instantiated.

In a first-order setting, these notions are expressed by using the non-logical symbol for the membership relation. However, the language of second-order logic affords us the resources to characterize analogues of these notions without resorting to non-logical vocabulary. One can then formulate counterparts of the cardinality questions left open by ZFC in the pure language of second-order logic.

On a Tarskian account, also known as *semantic* account, logical validity is identified with truth in every model. Given this account, a special relation appears to exist between certain cardinality statements in set theory and the logical validity

[1]This, of course, holds *modulo* relevant consistency assumptions.

of their counterparts in second-order logic. More specifically, there are cardinality statements that are true if and only if their second-order counterpart is true in every set-theoretic model. A number of authors have claimed this *entanglement* of second-order logic with mathematics to be problematic (Etchemendy 1990; Gómez-Torrente 1996; Gómez-Torrente 1998; Hanson 1997; Hanson 1999; Priest 1995; Ray 1996; Etchemendy 2008). Others have recently defended the opposite view (Parsons 2013; Paseau 2013; Griffiths and Paseau 2016). In previous work (Florio and Incurvati forthcoming), we have exhibited a conflict between the entanglement and the view that logic should be neutral. This can be seen as vindicating the claim that the entanglement is problematic and provides the basis for a plausible reconstruction of John Etchemendy's (1990) *Overgeneration Argument*.

Etchemendy concluded that the entanglement undermines the semantic account of logical validity. Other writers have retained the semantic account, taking the entanglement to cast doubt on the view that second-order logic is pure logic (Parsons 2013; Koellner 2010). Both reactions are premature in that they assume a specific implementation of the semantic account, one in which models are set-theoretic constructions. After all, the conflict between the entanglement and the neutrality of logic might be due to this assumption.

Indeed, earlier results of ours suggest that adopting a different semantics for second-order logic might provide an alternative way of resisting our reconstruction of the Overgeneration Argument. In particular, if we embrace higher-order resources in the metatheory and take models to be higher-order entities, the paradigmatic example of the conflict either disappears or is no longer problematic. This example involves the Continuum Hypothesis (CH), which states that every set has cardinality aleph-1 if and only if it has the cardinality of the continuum.

In this article, we extend this approach to other well-known cases of apparent entanglement involving the existence of certain small large cardinals, i.e. cardinals whose existence is consistent with the Axiom of Constructibility. We consider what happens to these cases of entanglement when second-order logic is given a higher-order semantics. Our findings are consistent with the optimistic outlook suggested by our earlier results. The adoption of a higher-order semantics for second-order logic helps avoid the conflict between the entanglement and the neutrality of logic.

1 Overgeneration

Using a set-theoretic semantics, i.e. a semantics that construes models as sets, it is provable in ZFC that CH is true if and only if the following sentence of second-order logic is valid:[2]

$$\textbf{(CH2)} \qquad \forall X(\textsc{aleph-1}(X) \leftrightarrow \textsc{continuum}(X))$$

The cardinality notions involved in **CH2** can be defined by pure formulas of second-order logic (see, e.g., Shapiro 1991, pp. 101–5). As our discussion will make clear, what matters is the provability in a given theory of the biconditional stating that CH is true if and only if **CH2** is valid. We will refer to this provability as the *entanglement of second-order logic with* CH. Note that one can also formulate a sentence **NCH2** of pure second-order logic which is entangled with the negation of CH when ZFC is used as the background theory and the semantics is construed set-theoretically (Shapiro 1991, p. 105).

In our discussion, it will be important to distinguish between a sentence of a formal language and the English (or mathematical English) sentence it formalizes. To this end, we shall adopt the following typographical convention: boldface indicates that a sentence belongs to a formal language, whereas its informal counterpart is indicated using roman characters. So, for example, CH2 stands for the English sentence that **CH2** formalizes.[3] If we read the second-order quantifier as ranging over properties, CH2 states that every property is ALEPH-1 if and only if it is CONTINUUM, where being ALEPH-1 and being CONTINUUM are characterized in terms of properties and their relations. Note that CH2 is different from CH, which states that every *set* is aleph-1 if and only if it is continuum, where these notions of size are characterized using the membership relation. Another point related to the distinction between formal and informal sentences: we use 'validity' to denote the property a formal sentence has just in case it is true in every model. Validity is meant to correspond to the notion of *logical truth*, which applies to informal sentences.

As noted above, the claim that the entanglement of second-order logic with CH is problematic has recently been disputed. Against this, we have identified two arguments in support of the claim, both of which take the entanglement to be problematic insofar as it is in tension with the neutrality of second-order logic (Florio and Incurvati forthcoming). Their difference lies in the way in which they articulate the notion of neutrality. In both cases, the intended conclusion is that second-order logic fails

---

[2]As usual, we indicate second-order variables by means of upper-case letters.

[3]We are assuming that for every formal sentence there exists a unique informal counterpart, and that for every informal sentence we consider there exists a unique formalization. The uniqueness assumption is made for convenience but is not strictly needed for the purposes of the paper. What is needed is, roughly put, that the formalization operation and its inverse respect the informal relation of logical equivalence. The existence assumption is a presupposition of the debate. It may be justified by appealing to natural language sentences that seem to require variable binding of predicate positions (Higginbotham 1998, p. 3). Alternatively, it may be justified by appealing to a second-order translation of sentences involving plural quantification (Boolos 1984; Boolos 1985) or non-nominal quantification of the kind explored by Rayo and Yablo (2001).

to be *sound with respect to logical truth*, i.e. there are second-order validities whose informal counterpart is not a logical truth. In this sense, the arguments purport to show that second-order logic *overgenerates*.

The first argument is based on the idea that logic is *dialectically* neutral: it should be able to serve as a neutral arbiter in disputes. The argument uses the entanglement of second-order logic with CH, soundness of second-order logic with respect to logical truth and *completeness of second-order logic with respect to logical truth*, i.e. the thesis that there are no logical truths whose second-order formalization is not valid. Let us spell out the argument. If second-order logic is to serve as a neutral arbiter in a legitimate dispute, settling the question of whether a statement is a logical truth should not close that dispute. But now consider a dispute over CH in the context of ZFC. Suppose, on the one hand, that we agree that CH2 is a logical truth. By completeness, we have that **CH2** is a higher-order validity. Using the entanglement of second-order logic with CH, we conclude that CH is true. Hence, if we agree that CH2 is a logical truth, we also have to agree that CH is true. Suppose, on the other hand, that we agree that CH2 is *not* a logical truth. By soundness, we have that **CH2** is not a higher-order validity. Using the entanglement of second-order logic with CH again, we conclude that CH is false. Hence, if we agree that CH2 is not a logical truth, we also have to agree that CH is false. Therefore, settling the question of whether CH2 is a logical truth settles the question of whether CH is true, thereby violating dialectical neutrality.[4] The proponent of the Overgeneration Argument concludes that second-order logic is not sound with respect to logical truth.[5]

The second argument is based on the idea that logic is *informationally* neutral: it should not be a source of new information. This idea licenses the following principle of *informational neutrality*:

(IN) if a theory T does not informationally contain p and p is a consequence of T together with q, then q is not a logical truth.

For had q been a logical truth, it could not have brought about such an increase in informational content. With the principle (IN) in mind, we can now turn to the

[4]There is an alternative version of the argument which involves NCH2. This version uses the entanglement of second-order logic with the negation of CH together with soundness and completeness of second-order logic with respect to logical truth. Below we will see that there are statements which are not entangled with second-order logic whereas their negation is.

[5]To save soundness, the defender of second-order logic might be tempted to deny that second-order logic is *complete* with respect to logical truth. In fact, the argument uses only a particular instance of completeness, namely the one which states that if CH2 is a logical truth, then **CH2** is valid. Hence, denying completeness is acceptable only if one can point to features of CH2 which are not captured by the formalization. Compare with the case of propositional logic: it is usually agreed that this logic is not complete with respect to logical truth, but this is acceptable only because we can point to features of certain natural language sentences (e.g. the presence of predicates and quantifiers) which are not captured by a propositional language.

argument. The argument is more complex than the first one, and we divide it into to steps.

The first step is as follows. Suppose that CH is true. By the entanglement of second-order logic with CH, it follows that **CH2** is valid. By soundness, we conclude that CH2 is a logical truth.

The second step relies on this key assumption: ZFC together with the thesis that second-order logic is complete with respect to logical truth does not informationally contain CH. However, CH can be derived from ZFC plus completeness if we assume that CH2 is a logical truth. Using the key assumption, we derive from (IN) that it is not a logical truth that CH2 is a logical truth. On the supposition that logical truth can be iterated (i.e. if φ is a logical truth, then it is a logical truth that φ is a logical truth), we conclude that CH2 is not a logical truth.

Therefore, the conclusion of the second step of the argument contradicts the conclusion of the first step. The argument supposes that CH is true but a parallel argument can be formulated using the assumption that CH is false. This parallel argument relies on the entanglement of second-order order logic with the negation of CH. The proponent of the Overgeneration Argument concludes, once again, that second-order logic is not sound with respect to logical truth.

Both arguments make essential use of the entanglement of second-order logic with CH. The second argument, in addition, makes use of the entanglement of second-order order logic with the negation of CH. Both cases of entanglement, recall, are provable in ZFC when second-order logic is given a set-theoretic semantics. But this is just *one* implementation of the semantic conception of logic. So the question arises of what happens when second-order logic is given a different semantics.

2 Higher-Order Semantics

Higher-order semantics is an alternative way of providing a semantics for second-order logic which has gained popularity. On this semantics, one adopts higher-order resources in the metatheory and uses second-order entities, rather than sets, as the values of object-language second-order variables (Boolos 1984; Boolos 1985; Rayo and Uzquiano 1999; Rayo 2002; Rayo and Williamson 2003; Yi 2005; Yi 2006; McKay 2006; Oliver and Smiley 2013). This is compatible with various interpretations of the second-order resources. For example, one may interpret these resources in terms of properties, plurals, or Fregean concepts. Alternatively, one may take the second-order resources to be *sui generis* and not reducible to more familiar linguistic constructions (Williamson 2003, p. 459). For convenience, we will read the second-order quantifier as ranging over properties.

Higher-order semantics allows us to define validity for second-order logic without ascending beyond second-order resources (see McGee 1997). Let us focus on a language whose sole non-logical symbol is the membership predicate of set theory ($\in$). For such a language, the definition of validity amounts to truth under any reinterpretation of $\in$ in any second-order domain. More formally, let $\varphi[E/\in]$ be the result of replacing all occurrences of $\in$ in φ with E and the superscript indicates the restriction of the formula's quantifiers to U. Then we say that φ is a *higher-order validity* if for every non-empty property U and every binary relation E, $\varphi[E/\in]^U$ holds. For instance, $\exists x \exists y\, x \in y$ is a higher-order validity just in case the following statement is true: for every non-empty property U and every binary relation E, there are x and y having U such that Exy. But the latter statement can be refuted in pure second-order logic and hence $\exists x \exists y\, x \in y$ is not a higher-order validity.

If we are concerned only with arguments with finitely many premises, entailment can be defined in terms of validity. In particular, we can say that $\gamma_1, \ldots, \gamma_n$ *entail* φ just in case $(\gamma_1 \wedge \ldots \wedge \ldots \gamma_n) \rightarrow \varphi$ is a higher-order validity.

A key issue concerns the resources used in the metatheory. As we will see, the strength of the resources allowed has an impact on whether second-order logic is entangled with mathematics. In terms of logical resources, we will work with an axiomazation of second-order logic obtained by adding to first-order logic the rules for the second-order quantifiers, full second-order Comprehension and the second-order principle of Property Choice, stating that there is a choice function corresponding to every relation. As for the non-logical resources, we will consider two options. For the first option, we need the notion of a second-order closure of a theory **T**. This is denoted by $\mathbf{T}^*$ and defined as the the set of sentences derivable from **T** in the axiomatization of second-order logic just described. The first option is then to use the informal counterpart of $\mathbf{ZFC}^*$. The second option is to use ZFC2, the informal counterpart of the theory obtained by replacing **ZFC**'s Replacement Schema with the corresponding second-order axiom. Since this theory has finitely many axioms, we can take it to be axiomatized by their conjunction. Thus, unlike the first option, the second option involves a theory whose non-logical axioms are formulated using distinctively higher-order resources.

It is worth noting that second-order logic with higher-order semantics retains its expressive power. For example, work by Shapiro (unpublished manuscript) and Väänänen and Wang (2015) shows that an internal version of Zermelo's (n.d.) categoricity theorem is provable in pure second-order logic: there is a quasi-isomorphism between any two higher-order models of **ZFC2**.[6] This implies that **ZFC2** semantically decides any statement concerning the hierarchy up to V_κ, where κ is the first inaccessible. Since **CH** concerns the first few infinite levels of the hierarchy, it follows that **ZFC2** semantically decides **CH**. That is, it is provable in pure second-order logic that **ZFC2** entails either **CH** or it entails $\neg$**CH**.

[6]A function F is a *quasi-isomorphism* between two higher-order models $\langle U_1, E_1 \rangle$ and $\langle U_2, E_2 \rangle$ if F is a bijection between U_1 and a subproperty of U_2 or between U_2 and a subproperty of U_1 such that F preserves the relations E_1 and E_2.

3 The Continuum Hypothesis

We are now in a position to answer the question left open: 'What happens to the entanglement of second-order logic with CH when a higher-order semantics is adopted?'. As we have shown in Florio and Incurvati forthcoming, the answer turns on the metatheory used. In particular, assume ZFC and suppose that **ZFC** has an ω-model, i.e. a model which is correct about the natural numbers. Then we have:

First Negative Theorem. **ZFC*** *does not prove that if* **CH2**, *then* **CH**.

Second Negative Theorem. **ZFC*** *does not prove that if* **CH**, *then* **CH2**.

In addition, the following lemma is provable in pure second-order logic.

Equivalence Lemma. **CH2** *is true if and only if* **CH2** *is a higher-order validity.*

Combining these results, one can conclude that **ZFC*** proves neither direction of the following biconditional:

CH if and only if **CH2** is a higher-order validity.

Does the informal counterpart of this result hold? That is, can we prove in ZFC* that CH if and only if CH2 is a logical truth? We think not. As an analogy, consider the case of standard independence results. The independence of **CH** from **ZFC** (Gödel 1939; Cohen 1963) is usually taken to show that CH is independent of ZFC. For, presumably, any purported informal proof of CH or its negation from ZFC could be turned into a corresponding formal proof, contradicting the Gödel-Cohen result. Similarly, suppose one could give an informal proof in ZFC* that CH if and only if CH2 is a logical truth. Such an informal proof could then be turned into a formal proof in **ZFC*** that **CH** if and only if **CH2** is a higher-order validity. This point extends to the relation between formal and informal statements in the case of small large cardinals considered below.

While the entanglement of second-order logic with CH is not provable in ZFC*, matters are different if we strengthen the set-theoretic background. In particular, **ZFC2** proves:

First Positive Theorem. *If* **CH2**, *then* **CH**.

Second Positive Theorem. *If* **CH**, *then* **CH2**.

Thus, given the Equivalence Lemma, we have that **ZFC2** proves that **CH** if and only if **CH2** is a higher-order validity. It follows that the informal counterpart of this biconditional may be established in ZFC2.

Recall that the argument from dialectical neutrality and the argument from informational neutrality make essential use of the entanglement of second-order logic with

CH. Thus, if we use ZFC* as our background theory, both arguments are blocked. What about the case in which we use ZFC2? Luckily for the defender of second-order logic, the arguments can be blocked in that case too. The issue turns on a property of **ZFC2** mentioned above, namely that it semantically decides **CH**.

Consider the argument from dialectical neutrality first. The idea was that second-order logic could not serve as a neutral arbiter in a dispute over CH in the context of ZFC. That is because, in the presence of the entanglement, settling whether CH2 is a *logical* truth settles whether CH is true. One might reformulate the argument so that it applies to ZFC2, given that this theory proves the entanglement.

However, note that the original version of the argument assumed that we need to allow for a dispute over CH in the context of ZFC. This assumption is made plausible by the fact that **ZFC**, if consistent, entails neither **CH** nor ¬**CH**. The alternative version of the argument would replace this assumption with the new assumption that we need to allow for the same dispute in the context of ZFC2. But the fact that **ZFC2** semantically decides **CH** gives us reason to reject the new assumption. For the fact that either **ZFC2** entails **CH** or it entails its negation means that one of **ZFC2** → **CH** and **ZFC2** → ¬**CH** is a higher-order validity. Therefore, in the context of ZFC2, it is not possible to agree on the logic while disagreeing on CH. Once it is agreed whether **ZFC2** → **CH** or **ZFC2** → ¬**CH** is a higher-order validity, the dispute over CH can be settled by an appeal to soundness of second-order logic with respect to logical truth followed by a simple *modus ponens*. Hence, the alternative version of the argument is undermined.[7]

Similar considerations apply to the argument from informational neutrality. The argument relied on the assumption that neither CH nor its negation is implicitly contained in ZFC. A reformulation of the argument would therefore rest on the assumption that neither CH nor its negation is implicitly contained in ZFC2. But, again, the assumption is undermined by the fact that **ZFC2** semantically decides **CH**.

The defender of second-order logic has thus reasons to switch to higher-order semantics. By adopting such a semantics, she can respond to both arguments against the soundness of second-order logic with respect to logical truth, no matter whether she chooses ZFC* or ZFC2 as her background theory. The viability of this response turns on two facts. First, the entanglement of second-order logic with CH is not provable in ZFC*. Second, even on a higher-order semantics, **ZFC2** semantically decides any statement concerning sets occurring below V_κ (with κ the first inaccessible) and hence **CH**.

However, there are statements other than CH whose truth is equivalent over ZFC

[7]It may be observed that the alternative version of the argument from dialectical neutrality uses the entanglement of second-order logic with the *negation* of CH. However, analogues of the results described in this section can be proved for the case of **NCH2** and ¬**CH** (Florio and Incurvati forthcoming). Thus, the alternative version of the argument is also blocked when moving to a higher-order semantics.

to the set-theoretic validity of their second-order counterparts. The entanglement of these statements with second-order logic has also been deemed problematic. What's more, some of these statements concern V_κ or higher levels of the hierarchy. Therefore, despite Zermelo's categoricity theorem and its internal version mentioned in Section 2, these statements are not semantically decided by **ZFC2** in either set-theoretic or higher-order semantics (modulo the appropriate consistency assumptions). Can the move to higher-order semantics help with these cases of entanglement beyond categoricity? In what follows, we focus on one notable set of examples concerning the higher infinite, namely statements asserting the existence of certain small large cardinals.[8]

4 Inaccessible cardinals

As usual, we say that a cardinal is inaccessible if it is uncountable, regular and strong limit. We use IC to denote the assertion that there is such a cardinal. As is well known, ZFC proves that κ is inaccessible if and only if V_κ is a set-theoretic model of **ZFC2**. This fact suggests the standard second-order counterpart of **IC**, namely the pure statement that there is a domain and a binary relation that satisfies **ZFC2**. In symbols:

(**IC2**) $$\exists U \exists R\, \mathbf{ZFC2}[R/\in]^U$$

Now, there is no entanglement of second-order logic with IC in set-theoretic semantics (as well as in higher-order semantics). Indeed, **IC2** is false in any model with a finite domain and hence the semantics easily refutes the validity of **IC2**. Therefore, **IC** does not imply that **IC2** is valid. However, the other direction of the entanglement holds so we have what we might call *semi-entanglement*. For the fact that **IC2** is not valid trivially implies that

if **IC2** is valid, then **IC**.

Things change if we consider the *negation* of **IC**. In particular, let **NIC2** be the negation of **IC2**. Then, given a set-theoretic semantics, it is provable in **ZFC** that

$\neg$**IC** if and only if **NIC2** is valid.

The argument from informational neutrality as stated requires both directions of the entanglement for a statement *and* its negation. So it is not applicable in the

[8] To see that **ZFC2** does not semantically decide these statements given the appropriate consistency assumption, suppose that a small large cardinal κ exists. Then, one can obtain higher-order models of **ZFC2** which disagree over the existence of κ by truncating at κ. The existence of the corresponding set-theoretic models can be established by assuming that there is an inaccessible above κ.

present case. The argument from dialectical neutrality, on the other hand, only needs that a statement *or* its negation be entangled with second-order logic. Thus it can be formulated using the entanglement of the negation of **IC**.

Does this case of entanglement carry over to higher-order semantics? The answer is negative, no matter whether the background theory is **ZFC*** or **ZFC2**. Let us begin with **ZFC2**. The following two theorems hold.

Theorem 1. **ZFC2** *proves that if* **NIC2** *is a higher-order validity, then* $\neg$**IC**.

Theorem 2. **ZFC2** *does not prove that if* $\neg$**IC**, *then* **NIC2** *is a higher-order validity.*

In both cases, the key observation is that, by a simple existential introduction, **ZFC2** proves that **IC2** holds and hence that **NIC2** is *not* a higher-order validity. This statement is the negation of the antecedent of the conditional in Theorem 1. Thus the theorem follows trivially. For Theorem 2, suppose towards a contradiction that **ZFC2** proved that if $\neg$**IC**, then **NIC2** is a higher-order validity. Since **ZFC2** proves that **NIC2** is not a higher-order validity, it would follow that **ZFC2** proved **IC** and hence its own consistency. But this cannot be for standard consistency reasons.

Let us now turn to **ZFC***. To start with, note that **ZFC*** is a subtheory of **ZFC2**. Therefore, Theorem 2 yields immediately:

Theorem 3. **ZFC*** *does not prove that if* $\neg$**IC**, *then* **NIC2** *is a higher-order validity.*

Unlike **ZFC2**, **ZFC*** does not prove the other direction of the conditional in Theorem 3. This follows from the next two results, namely Lemma 4 and Theorem 5.

Lemma 4 is a version of the Equivalence Lemma above and is provable in pure second-order logic.

Lemma 4. **NIC2** *if and only if* **NIC2** *is a higher-order validity.*

The right-to-left direction is straightforward. For the other direction, assume **NIC2** and suppose towards a contradiction that **NIC2** is not a higher-order validity. This means that there is a non-empty U^* and there are U and R such that $\big(\mathbf{ZFC2}[R/\in]^U\big)^{U^*}$. By second-order Comprehension we have a property Z which something has if and only it has both U and U^*. It is easy to verify that $\mathbf{ZFC2}[R/\in]^Z$, which contradicts **NIC2**.

The proof of Theorem 5 uses the assumption that there is an ω-model of the theory consisting of **ZFC** and the assertion that there are two inaccessibles. Note that the theorem is an underivability result. So, in the course of the proof, it is permissible to give the higher-order metalanguage a set-theoretic interpretation and use countermodels constructed according to the set-theoretic semantics.

Theorem 5. **ZFC*** *does not prove that if* **NIC2**, *then* $\neg$**IC**.

Proof. Let us abbreviate the statement that there are two inaccessibles in the constructible hierarchy $\mathbf{L}$ as $\mathbf{2IC^L}$. First, we use the consistency assumption to construct an ω-model m of $\mathbf{ZFC} + \neg\mathbf{IC} + \mathbf{2IC^L}$. Start from an ω-model with at least two inaccessibles and obtain by truncation a model with exactly two inaccessibles. Using forcing, make the two inaccessible cardinals of the model only weakly inaccessible (thus destroying their strong inaccessibility). This gives the desired model, since weakly inaccessible cardinals are strongly inaccessibles in $\mathbf{L}$.

Next, we use m to establish the consistency of the theory $\mathbf{ZFC}^* + \mathbf{NIC2} + \mathbf{IC}$. Let κ be the second cardinal of m which is inaccessible in $\mathbf{L}$. From the perspective of m, L_κ is a model of $\mathbf{ZFC} + \mathbf{IC}$. Hence, again from the perspective of m, $\langle L_\kappa, \mathscr{P}(L_\kappa)\rangle$ is a model of $\mathbf{ZFC}^* + \mathbf{IC}$. Now since m satisfies $\mathbf{ZFC}$, it thinks that every set-theoretic model satisfies $\mathbf{NIC2}$ if and only if $\neg\mathbf{IC}$ is true. But m is a model of $\neg\mathbf{IC}$ and therefore it thinks that $\langle L_\kappa, \mathscr{P}(L_\kappa)\rangle$ is also a model of $\mathbf{NIC2}$. Hence m thinks that the theory $\mathbf{ZFC}^* + \mathbf{NIC2} + \mathbf{IC}$ is consistent. Since m is an ω-model, it is correct about consistency facts. Therefore the theory $\mathbf{ZFC}^* + \mathbf{NIC2} + \mathbf{IC}$ really is consistent. $\square$

The situation, therefore, is the following. If the background theory is ZFC*, neither direction of the entanglement of second-order logic with ¬IC is provable in higher-order semantics. If we strengthen the background theory to ZFC2, one direction of the entanglement becomes provable. Either way, the argument from dialectical neutrality is blocked. However, one might worry that the provability of semi-entanglement for the case of **IC** and ¬**IC** still poses a threat to the neutrality of logic. We address this issue in Section 6.

Of course, there are natural strengthenings of ZFC2 which do prove both directions of the entanglement of second-order logic with ¬IC in higher-order semantics. For instance, the result of adding to ZFC2 the assertion that there inaccessible cardinals trivially proves that if there are no inaccessibles, then **NIC2** is a higher-order validity. However, the dispute over the existence of an inaccessible is closed by such a theory quite independently of the entanglement. Therefore, its availability poses no problems for the dialectical neutrality of second-order logic. The same holds for other theories which imply the existence of inaccessibles, e.g. the theory obtained by adding a form of second-order reflection to ZFC2.

5 Mahlos and weakly compacts

The fact that **ZFC2** proves exactly one direction of the entanglement of ¬**IC** with second-order logic is a direct consequence of the definition of **NIC2**. That is, **ZFC2** proves that there is an inaccessible property, contradicting the higher-order validity of **NIC2**. Thus one may wonder what the situation is with respect to other small large cardinals, such as Mahlos and weakly compacts. For as well as not proving the existence of Mahlo and weakly compact *sets*, **ZFC2** does not prove the existence of

Mahlo and weakly compact *properties*. As it turns out, perfect analogues of the results obtained for $\neg$**IC** can be obtained for the case of Mahlos and weakly compacts. Let us consider these cases.

Recall that a cardinal κ is *Mahlo* if every normal function on κ has an inaccessible fixed point.[9] We denote the set-theoretic statement that there is a Mahlo cardinal with M. Now, it can be shown in ZFC that κ is Mahlo if and only if V_κ is a set-theoretic model of **ZFCM2**, i.e. **ZFC2** plus the following axiom:

$$\forall F(\mathbf{normal}(F) \rightarrow \exists\kappa(\mathbf{Inacc}(\kappa) \;\&\; F(\kappa,\kappa)))$$

Here $\mathbf{normal}(F)$ and $\mathbf{Inacc}(\kappa)$ formalize the statements that F is a normal functional property and κ is an inaccessible set.

As before, this suggests the standard second-order counterpart of **M**, namely the pure statement that there is a domain and a binary relation that satisfies **ZFCM2**:

(**M2**) $$\exists U \exists R\ \mathbf{ZFCM2}[R/\in]^U$$

Again, there is only semi-entanglement of second-order logic with M in set-theoretic semantics (as well as in higher-order semantics). On the other hand, if we let **NM2** be the negation of **M2**, the following is provable in **ZFC** given the set-theoretic semantics:

$\neg$**M** if and only if **NM2** is valid.

We now switch to higher-order semantics. Again, we first examine what happens when the metatheory is **ZFC2**. In that case, we have:

Theorem 6. **ZFC2** *proves that if* **NM2** *is a higher-order validity, then* $\neg$**M**.

Theorem 7. **ZFC2** *does not prove that if* $\neg$**M**, *then* **NM2** *is a higher-order validity.*

For Theorem 6, we reason in **ZFC2** and show that if there is a Mahlo cardinal, then **NM2** is not a higher-order validity. Let κ be Mahlo. It can be verified that the axioms of **ZFCM2** hold when restricted to V_κ. So the property of being a set in V_κ and the relation of membership restricted to V_κ provide witnesses to the existential quantifiers of **M2**. So **M2** is true and hence **NM2** is not a higher-order validity.

For Theorem 7, we reason about provability in **ZFC2** and establish that the relevant conditional is unprovable by providing a set-theoretic countermodel. We make

[9]A function on the ordinals is normal if it is strictly increasing and continuous. Since in the definition of a Mahlo cardinal one restricts attention to normal functions on a set of ordinals κ, the notion of being Mahlo is first-order definable. Note that the definition of a Mahlo cardinal given above is equivalent over the ZFC axioms to the following definition, which is perhaps more standard: a cardinal κ is Mahlo if it is regular and the set of inaccessibles below κ is stationary in κ. A set $a \subseteq \kappa$ is stationary in κ if $a \cap c \neq \emptyset$ for each closed unbounded subset c of κ.

the consistency assumption that there is a set-theoretic model of **ZFCM2**. This model must be a V_κ with κ a Mahlo cardinal. Now truncate this model at the first Mahlo cardinal μ. (If there are none, let μ simply be κ.) Clearly, V_μ is a model of $\neg$**M**. On the other hand, since μ is Mahlo, V_μ is a model of **ZFCM2** and hence of **M2**. Therefore V_μ is a model of the statement that **NM2** is not a higher-order validity.

The perfect parallel between the case of inaccessibles and that of Mahlos holds even if we weaken the background theory to **ZFC***. Obviously, the direction of the entanglement unprovable in **ZFC2** is also unprovable in **ZFC***:

Theorem 8. **ZFC*** *does not prove that if* $\neg$**M***, then* **NM2** *is a higher-order validity.*

Moreover, the following two results can be established by adapting the proofs of Lemma 4 and Theorem 5.

Lemma 9. **NM2** *if and only if* **NM2** *is a higher-order validity.*

Theorem 10. **ZFC*** *does not prove that if* **NM2***, then* $\neg$**M***.*

Thus, **ZFC*** also fails to prove the other direction of the entanglement.

How about weakly compact cardinals? One can formulate a second-order statement which, when added to **ZFC2**, yields a theory whose set-theoretic models are all and only the V_κ with κ a weakly compact cardinal (Hellman 1989, 86 ff.). Proceeding as before, one can then find a pure second-order statement (**WKC2**) whose set-theoretic validity implies the assertion that there is a weakly compact (**WKC**). As in the previous cases, **ZFC** proves that the negation of this statement (**NWKC2**) is true in every set-theoretic model if and only if there are no weakly compacts ($\neg$**WKC**). When moving to a higher-order semantics, one can prove analogues of the theorems obtained above. This means that we have no entanglement when the background theory is **ZFC*** and only semi-entanglement when the background theory is **ZFC2**. These results can be established using the techniques employed in the the case of Mahlo cardinals.

6 Semi-entanglement

We now want to reassess the arguments from dialectical and information neutrality in the light of the results proved in the previous section. As mentioned, the argument from informational neutrality requires the entanglement for a statement and its negation. Therefore it does not apply in the case of small large cardinals, even on a set-theoretic semantics. In contrast, the argument from dialectical neutrality requires the entanglement of a statement or its negation. So it continues to apply with respect to small large cardinals. In these cases, higher-order semantics provides a clear way out: none of the relevant statements which are entangled in set-theoretic semantics continues to be entangled in higher-order semantics. So the new instances of the argument from dialectical neutrality are blocked.

If the background theory is ZFC2 rather than ZFC*, we still have semi-entanglement. Thus one might wonder whether semi-entanglement suffices to run a *modified* version of the arguments we have articulated. In the case of informational neutrality, it is hard to see how the argument could be modified so as to rely only on semi-entanglement. The case of dialectical neutrality requires closer scrutiny.

Let S be any of the statements involving small large cardinals which are entangled with second-order logic in set-theoretic semantics (i.e. ¬IC, ¬M and ¬WKC) and let S2 be its second-order counterpart. Then the argument from dialectical neutrality concerning S goes as follows. By logic, either S2 is a logical truth or it isn't. If it is, using completeness of second-order logic with respect to logical truth and one direction of the entanglement, we derive that S is true. It follows that S is settled. If it isn't a logical truth, using soundness of second-order logic with respect to logical truth and the other direction of the entanglement, we derive that S is false. It follows, again, that S is settled. Therefore, no matter what logical status S2 has, S is settled. This means that second-order logic cannot be a neutral arbiter in a dispute over S.

Since in the context of higher-order semantics we only have semi-entanglement when the background theory is **ZFC2**, it is not the case that S is settled independently of the logical status of S2. For each of S and ¬S is consistent with the assumption that S2 is not a logical truth. So the conclusion that second-order logic cannot be a neutral arbiter in a dispute over S is not available in general. However, one may point out that the conclusion *is* available if it is assumed that S2 is a logical truth. For, in the presence of semi-entanglement, this assumption together with completeness of second-order logic with respect to logical truth entail that S is true. So it follows that S is settled.

To start with, note that this modified version of the argument leads at best to denying completeness, not soundness. Thus it cannot be considered an *overgeneration* argument. Rather, the denial of completeness results in the claim that second-order logic *undergenerates*: there are logical truths whose second-order formalization is not valid.

In any case, the question we should ask is: what reasons do we have to assume that S2 is true and hence a logical truth? To answer this question, it is useful to consider first another case, namely that of CH2 and NCH2. As Michael Potter remarks:

> [W]e do not seem to have any intuitions about whether these second-order principles [i.e. CH2 and NCH2] that could settle the continuum hypothesis are themselves true or false. So this observation [that there is entanglement] does not seem especially likely to be a route to an argument that will actually settle the continuum hypothesis one way or the other. (Potter 2004, p. 271)

The same could be said about S2. Of course, one might deny this and insist that either we have no intuitions about whether S2 is true or we have intuitions that it is false. But suppose that we did think that S2 is true. Presumably this is because we have certain views about the extent of the set-theoretic hierarchy. For example, it is because we think that the hierarchy does not have Mahlo height that we think that there is no Mahlo property and hence that NM2 is true. Clearly, this is also a reason for thinking that the hierarchy contains no Mahlo set: if there was such a set, one could obtain a Mahlo property by simply considering the property of being a member of that set. Therefore the reason to accept NM2 *ipso facto* closes the dispute over the existence of a Mahlo.

Note that this response does not carry over to the situation in which we have full entanglement. For not every reason to deny the existence of a Mahlo is *ipso facto* a reason to accept NM2. For instance, one might think that the universe is V_κ with κ the first Mahlo cardinal. In that case, one would hold that the hierarchy has Mahlo height without thereby thinking that there is a Mahlo set.

So, in all these cases of semi-entanglement, if we have reasons to take the second-order statement to be true, we also have reasons to take its set-theoretic counterpart to be true independently of semi-entanglement. Thus there appears to be no conflict between the relevant cases of semi-entanglement and dialectical neutrality.

7 Conclusion

The semantic account of logic identifies validity with truth in every model. When models are construed set-theoretically, this account gives rise to cases of entanglement which are in conflict with neutrality of second-order logic. This conflict is captured by the arguments from dialectical neutrality and informational neutrality. These arguments make use of the assumption that second-order logic is dialectically neutral and the assumption that second-order logic is informationally neutral. Thus, one could block them by rejecting those assumptions.

We have shown that the defender of second-order logic has an alternative way of blocking the arguments, namely by adopting a higher semantics. This semantics embodies a different implementation of the semantic account of logic in that it takes models to be higher-order constructions rather than sets. Of course, one might challenge the legitimacy of higher-order semantics. But this would require a separate argument targeting the use of second-order resources in semantic theorizing. Pending such an argument, the defender of second-order logic may resort to higher-order semantics in response to the Overgeneration Argument.

In this article, we have focused on some cases of entanglement in the higher infinite. We have shown that these cases of entanglement, previously thought to be problematic, disappear when a higher-order semantics is adopted. In the absence of entanglement, both of our reconstructions of the Overgeneration Argument are blocked. Moreover, a modified version of the argument from dialectical neutrality based on semi-entanglement fails to establish its intended conclusion.

This provides additional evidence that the proponent of the semantic account can hold on to the neutrality of second-order logic by employing a higher-order semantics. When one takes into account models that are not set-sized, denying the existence of small large cardinals does not translate into a commitment to the validity of certain second-order sentences.

There remains one well-known case of entanglement in set-theoretic semantics which we have not considered. This is the one involving the Generalized Continuum Hypothesis, and we plan to explore it in future work. Together with the results obtained in this article, this new investigation will provide a clearer picture of the extent to which higher-order semantics can help vindicate the neutrality of second-order logic.[10]

[10]This work has received funding from a Leverhulme Research Fellowship held by Salvatore Florio and from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 758540) within Luca Incurvati's project *From the Expression of Disagreement to New Foundations for Expressivist Semantics*. For comments and discussion, we would like to thank Michael Glanzberg, Mario Gómez-Torrente, Graham Leach-Krouse, Beau Madison Mount, Ian Rumfitt, Gil Sagi, Kevin Scharp, Gila Sher, Jack Woods and Elia Zardini. Earlier versions of this material were presented at the workshop *The Semantic Conception of Logic* in Munich and the conference *Disagreement within Philosophy* in Bonn. We are grateful to the members of these audiences for their valuable feedback.

Works Cited

Boolos, George. "Nominalist Platonism". *Philosophical Review*, vol. 94, 1985, pp. 327–44.

---. "The Iterative Conception of Set". *Journal of Philosophy*, vol. 68, 1971, pp. 215–31.

---. "To Be Is to Be a Value of a Variable (or to Be Some Values of Some Variables)". *Journal of Philosophy*, vol. 81, 1984, pp. 430–49.

Cohen, Paul J. "The Independence of the Continuum Hypothesis". *Proceedings of the National Academy of Sciences of the United States of America*, vol. 50, 1963, pp. 1143–8.

Etchemendy, John. "Reflections on Consequence". *New Essays on Tarski and Philosophy*, edited by Douglas Patterson, Oxford UP, 2008, pp. 263–99.

---. *The Concept of Logical Consequence*. Harvard UP, 1990.

Florio, Salvatore and Luca Incurvati. "Metalogic and the Overgeneration Argument". *Mind*, forthcoming, `https://doi.org/10.1093/mind/fzy059`.

Gödel, Kurt. "Consistency Proof for the Generalized Continuum Hypothesis". *Proceedings of the National Academy of Sciences of the United States of America*, vol. 25, 1939, pp. 220–4.

Gómez-Torrente, Mario. "On a Fallacy Attributed to Tarski". *History and Philosophy of Logic*, vol. 19, 1998, pp. 227–34.

---. "Tarski on Logical Consequence". *Notre Dame Journal of Formal Logic*, vol. 37, 1996, pp. 125–51.

Griffiths, Owen and Alexander Paseau. "Isomorphism Invariance and Overgeneration". *Bulletin of Symbolic Logic*, vol. 22, 2016, pp. 482–503.

Hanson, William H. "Ray on Tarski on Logical Consequence". *Journal of Philosophical Logic*, vol. 28, 1999, pp. 605–16.

---. "The Concept of Logical Consequence". *Philosophical Review*, vol. 106, 1997, pp. 365–409.

Hellman, Geoffrey. *Mathematics Without Numbers: Towards a Modal-Structural Interpretation*. Oxford UP, 1989.

Higginbotham, James. "On Higher-Order Logic and Natural Language". *Philosophical Logic*, edited by Timothy Smiley, Oxford UP, 1998, pp. 1–27.

Koellner, Peter. "Strong Logics of First and Second Order". *Bulletin of Symbolic Logic*, vol. 16, 2010, pp. 1–36.

McGee, Vann. "How We Learn Mathematical Language". *Philosophical Review*, vol. 106, 1997, pp. 35–68.

McKay, Thomas J. *Plural Predication*. Oxford UP, 2006.

Oliver, Alex and Timothy Smiley. *Plural Logic*. Oxford UP, 2013.

Parsons, Charles. "Some consequences of the entanglement of logic and mathematics". *Reference, Rationality, and Phenomenology. Themes from Føllesdal*, edited by Michael Frauchiger, De Gruyter, 2013, pp. 153–78.

Paseau, Alexander. "The Overgeneration Argument(s): A Succinct Refutation". *Analysis*, vol. 74, 2013, pp. 40–7.

Potter, Michael. *Set Theory and Its Philosophy. A Critical Introduction*. Oxford UP, 2004.

Priest, Graham. "Etchemendy and Logical Consequence". *Canadian Jounal of Philosophy*, vol. 25, 1995, pp. 283–92.

Ray, Greg. "Logical Consequence: A Defense of Tarski". *Journal of Philosophical Logic*, vol. 25, 1996, pp. 617–77.

Rayo, Agustín. "Word and Objects". *Noûs*, vol. 36, 2002, pp. 436–464.

Rayo, Agustín and Gabriel Uzquiano. "Toward a Theory of Second-Order Consequence". *Notre Dame Journal of Formal Logic*, vol. 40, 1999, pp. 315–25.

Rayo, Agustín and Timothy Williamson. "A Completeness Theorem for Unrestricted First-Order Languages". *Liars and Heaps*, edited by JC Beall, Oxford UP, 2003, pp. 331–56.

Rayo, Agustín and Stephen Yablo. "Nominalism Through De-Nominalization". *Noûs*, vol. 35, 2001, pp. 74–92.

Shapiro, Stewart. *Foundations without Foundationalism. A Case for Second-order Logic*. Oxford UP, 1991.

Väänänen, Jouko and Tong Wang. "Internal Categoricity in Arithmetic and Set Theory". *Notre Dame Journal of Formal Logic*, vol. 56, 2015, pp. 121–34.

Williamson, Timothy. "Everything". *Philosophical Perspectives*, vol. 17, 2003, pp. 415–65.

Yi, Byeong-Uk. "The Logic and Meaning of Plurals. Part I". *Journal of Philosophical Logic*, vol. 34, 2005, pp. 459–506.

---. "The Logic and Meaning of Plurals. Part II". *Journal of Philosophical Logic*, vol. 35, 2006, pp. 239–88.

Zermelo, Ernst. "Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre". *Fundamenta Mathematicae* vol. 16, 1930, pp. 29-47. Translated as "On Boundary Numbers and Domains of Sets: New Investigations in the Foundations of Set Theory", *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, vol. 2, edited by William Ewald, Oxford UP, 1996, pp. 1219-33.