The Review of Symbolic Logic
Volume 19, Number 1, March 2026

SATISFACTION IS NOT ABSOLUTE

JOEL DAVID HAMKINS
Department of Philosophy, University of Notre Dame
and
RUIZHI YANG
School of Philosophy, Fudan University

Abstract. We prove that the satisfaction relation $\mathcal{N} \models \varphi[\vec{a}]$ of first-order logic is not absolute between models of set theory having the structure $\mathcal{N}$ and the formulas φ all in common. Two models of set theory can have the same natural numbers, for example, and the same standard model of arithmetic $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle$, yet disagree on their theories of arithmetic truth; two models of set theory can have the same natural numbers and the same arithmetic truths, yet disagree on their truths-about-truth, at any desired level of the iterated truth-predicate hierarchy; two models of set theory can have the same natural numbers and the same reals, yet disagree on projective truth; two models of set theory can have the same $\langle H_{\omega_2}, \in \rangle$ or the same rank-initial segment $\langle V_\delta, \in \rangle$, yet disagree on which assertions are true in these structures.

On the basis of these mathematical results, we argue that a philosophical commitment to the determinateness of the theory of truth for a structure cannot be seen as a consequence solely of the determinateness of the structure in which that truth resides. The determinate nature of arithmetic truth, for example, is not a consequence of the determinate nature of the arithmetic structure $\mathbb{N} = \{ 0, 1, 2, \dots \}$ itself, but rather, we argue, is an additional higher-order commitment requiring its own analysis and justification.

§1. Introduction. Many mathematicians and philosophers regard the natural numbers $0, 1, 2, \dots$, along with their usual arithmetic structure, as having a privileged mathematical existence, a Platonic realm of numbers in which assertions have definite, absolute truth values, independently of our ability to prove or discover them. Although there are some arithmetic assertions that we can neither prove nor refute—such as the consistency of the background theory in which we undertake our proofs—the view is that nevertheless there is a fact of the matter about whether any such arithmetic statement is true or false in the intended interpretation. The definite nature of arithmetic truth is often seen as a consequence of the definiteness of the structure of arithmetic $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle$ itself, for if the natural numbers exist in a clear and distinct totality in a way that is unambiguous and absolute, then (on this view) the first-order theory of truth residing in that structure—arithmetic truth—is similarly clear and distinct.

Received: August 19, 2025.
2020 *Mathematics Subject Classification*: Primary 03C62, 00A30.
Key words and phrases: truth, satisfaction class, set theory.

© The Author(s), 2025. Published by Cambridge University Press on behalf of The Association for Symbolic Logic. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

doi:10.1017/S1755020325100907

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

Feferman provides an instance of this perspective when he writes:

> In my view, the conception [of the bare structure of the natural numbers] is *completely clear*, and thence *all arithmetical statements are definite*. [5, pp. 6–7] (emphasis original)

It is Feferman's 'thence' to which we call attention, and many mathematicians and philosophers seem to share this perspective. Martin writes:

> What I am suggesting is that the real reason for confidence in first-order completeness is our confidence in the full determinateness of the concept of the natural numbers. [26, p. 13]

The truth of an arithmetic statement, to be sure, does seem to depend entirely on the structure $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle$, with all quantifiers restricted to $\mathbb{N}$ and using only those arithmetic operations and relations, and so if that structure has a "definite" nature, then it would seem that the truth of the statement should be similarly definite.

Nevertheless, in this article we should like to tease apart these two ontological commitments, arguing that the definiteness of truth for a given mathematical structure, such as the natural numbers, the reals or higher-order structures such as H_{ω_2} or V_δ, does not follow from the definite nature of the underlying structure in which that truth resides. Rather, we argue that the commitment to a theory of truth for a structure is a higher-order ontological commitment, going strictly beyond the commitment to a definite nature for the underlying structure itself. This is, of course, a commitment that we expect many of those authors to want to make—our main point is that it does not come for free.

We shall make our argument by first proving, as a strictly mathematical matter, that different models of set theory can have a structure identically in common, even the natural numbers, yet disagree on the theory of truth for that structure.

- Models of set theory can have the same structure of arithmetic $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle$, yet disagree on arithmetic truth.
- Models of set theory can have the same reals, yet disagree on projective truth.
- Models of set theory can have a transitive rank initial segment V_δ in common, yet disagree about whether it is a model of ZFC.

In these cases and many others, the theory of a structure is not absolute between models of set theory having that structure identically in common. This is a stronger kind of non-absoluteness phenomenon than the usual observation, via the incompleteness theorem, that models of set theory can disagree on arithmetic truth, for here we have models of set theory, which disagree about arithmetic truth, yet agree completely on the structure in which that truth resides. Our mathematical claims will be made in §2 through §5. Afterwards, on the basis of these mathematical observations, we shall draw our philosophical conclusions in §6.

Satisfaction is absolute! This slogan, heard by the first author in his graduate-student days—a fellow logic student would assert it with exaggerated double-entendre—was meant to evoke the idea that the satisfaction relation $\mathcal{N} \models \varphi[\vec{a}]$ is absolute between all the various models of set theory able to express it. The set-theoretic universe V, for example, has the same arithmetic truths as the constructible universe L; and it doesn't matter, when asking whether $\mathcal{N} \models \varphi[\vec{a}]$, whether one determines the answer in the universe V or in a forcing extension $V[G]$ of it. So the slogan is true to a very

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

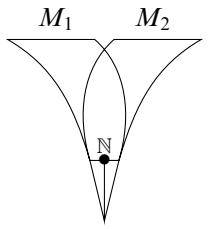
high degree, such as in the case of Shoenfield absoluteness and Lévy absoluteness, and in particular, it is true whenever the formula φ has standard-finite length in the meta-theory or between any two models of set theory for which at least one has access to the satisfaction relation of the other, since any model of set theory that can see two satisfaction relations for a structure will see by induction on formulas that they must agree. Nevertheless, the main theorems of this article show that the slogan is not true when one uses a broader concept of absoluteness, namely, satisfaction is *not* absolute, for there can be models of set theory with a structure $\mathcal{N}$ and sentence σ in common, but which disagree about whether $\mathcal{N} \models \sigma$.

Before proceeding further, we should like to remark on the folklore nature of some of the mathematical arguments and results contained in §2 through §5 of this article. Theorems 1 and 5 and their consequences, for example, are proved here using only well-known classical methods. These arguments could be considered as a part of the mathematical folklore of the subject of models of arithmetic, a subject filled with many fascinating results about automorphisms of nonstandard models of arithmetic and of set theory and the images of non-definable sets in computably saturated models. For example, Schlipf [27] proved many basic results about computably saturated models, including the case of $\langle \mathcal{M}, X\rangle$ where X is not definable in M, a case which figures in our theorems 1 and 5. For example, Schlipf proves that if $\mathcal{M} \models \text{ZF}$ is resplendent, then there is a cofinal set of indiscernibles I in the ordinals of $\mathcal{M}$ such that for each $\alpha \in I$ we have $V_\alpha^{\mathcal{M}} \prec \mathcal{M}$ and $V_\alpha^{\mathcal{M}} \cong \mathcal{M}$; if $\mathcal{M}$ is also countable and $V_\alpha^{\mathcal{M}} \prec \mathcal{M}$, then there are 2^ω many distinct isomorphisms $V_\alpha^{\mathcal{M}} \cong \mathcal{M}$; and $\mathcal{M}$ is isomorphic to some topless initial segment $\mathcal{N} \prec \mathcal{M}$, that is, for which $\mathcal{M}$ has no supremum to the ordinals of $\mathcal{N}$. Kossak and Kotlarski [16] identified circumstances under which a nondefinable subset X of a countable model $\mathcal{M} \models \text{PA}$ must have the maximum number of automorphic images in $\mathcal{M}$, including the case where X is an inductive satisfaction class. Schmerl subsequently proved that every undefinable class X in such a model $\mathcal{M}$ has continuum many automorphic images. In other work, Kossak and Kotlarski [17] proved that if $\mathcal{M}$ is a model of PA with a full inductive satisfaction class, then it has full inductive satisfaction classes S_1 and S_2 which disagree on a set of sentences that is coinitial with the standard cut. So the topic is well-developed and much is known. Concerning the specific results we prove in this article, experts in the area seem instinctively to want to prove them by means of resplendency and the other sophisticated contemporary ideas that frame the current understanding of the subject—showing the depth and power of those methods—and indeed one may prove the theorems via resplendency. Nevertheless, our arguments here show that elementary methods suffice.

§2. Indefinite arithmetic truth. Let us begin with what may seem naively to be a surprising case, where we have two models of set theory with the same structure of arithmetic $\langle \mathbb{N}, +, \cdot, 0, 1, <\rangle$, but different theories of arithmetic truth.

THEOREM 1. *Every consistent extension of ZFC has two models M_1 and M_2, which agree on the natural numbers and on the structure $\langle \mathbb{N}, +, \cdot, 0, 1, <\rangle^{M_1} = \langle \mathbb{N}, +, \cdot, 0, 1, <\rangle^{M_2}$, but which disagree their theories of arithmetic truth, in the sense that there is in M_1 and M_2 an arithmetic sentence σ, such that M_1 thinks σ is true, but M_2 thinks it is false.*

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press



Thus, two models of set theory can agree on which natural numbers exist and agree on all the details of the standard model of arithmetic, yet disagree on which sentences are true in that model. The proof is elementary, but before giving the proof, we should like to place the theorem into the context of some classical results on arithmetic truth, particularly Krajewski's work [19, 20] on incompatible satisfaction classes, explained in theorem 2.

Inside every model of set theory $M \models \text{ZFC}$, we may extract a canonical model of arithmetic, the structure $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle^M$, which we henceforth denote simply by $\mathbb{N}^M$, arising from what M views as the standard model of arithmetic. Namely, $\mathbb{N}^M$ is the structure whose objects are the objects that M thinks to be natural numbers and whose operations and relations agree with what M thinks are the standard arithmetic operations and relations on the natural numbers. Let us define that a *ZFC-standard* model of arithmetic, or just a *standard* model of arithmetic (as opposed to *the* standard model of arithmetic), is a model of arithmetic that arises in this way as $\mathbb{N}^M$ for some model $M \models \text{ZFC}$. In other words, *a* standard model of arithmetic is one that is thought to be *the* standard model of arithmetic from the perspective of some model of ZFC. More generally, for any set theory T we say that a model of arithmetic is a *T-standard* model of arithmetic, if it arises as $\mathbb{N}^M$ for some $M \models T$.

Every model of set theory $M \models \text{ZFC}$ has what it thinks is the true theory of arithmetic TA^M, the collection of σ thought by M to be (the Gödel code of) a sentence in the language of arithmetic, true in $\mathbb{N}^M$. (In order to simplify notation, we shall henceforth identify formulas with their Gödel codes.) The theory TA^M is definable in M by means of the recursive Tarskian definition of truth-in-a-structure, although it is not definable in $\mathbb{N}^M$, by Tarski's theorem on the non-definability of truth. Note that when $\mathbb{N}^M$ is nonstandard, this theory will include many nonstandard sentences σ, which do not correspond to any actual assertion in the language of arithmetic from the perspective of the metatheory, but nevertheless, these sentences gain a meaningful truth value inside M, where they appear to be standard, via the Tarski recursion as carried out inside M. In this way, the ZFC-standard models of arithmetic can be equipped with a notion of truth, which obeys the recursive requirements of the Tarskian definition.

These relativized truth predicates are instances of the more general concept of a satisfaction class for a model of arithmetic (see [15, 22] for general background). Given a model of arithmetic $\mathcal{N} = \langle N, +, \cdot, 0, 1, < \rangle$, we say that a subclass $\text{Tr} \subseteq N$ is a *truth predicate* or a *full satisfaction class* for $\mathcal{N}$—we shall use the terms interchangeably—if every element of Tr is a sentence in the language of arithmetic as viewed by $\mathcal{N}$, and such that Tr obeys the recursive Tarskian definition of truth:

1. (atomic) For each atomic sentence σ in N, we have $\sigma \in \text{Tr}$ just in case N thinks σ is true. (Note that the value of any closed term may be uniquely evaluated inside

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

$\mathcal{N}$ by an internal recursion, and so every model of arithmetic has a definable relation for determining the truth of atomic assertions.)

2. (conjunction) $\sigma \wedge \tau \in \text{Tr}$ if and only if $\sigma \in \text{Tr}$ and $\tau \in \text{Tr}$.
3. (negation) $\neg\sigma \in \text{Tr}$ if and only if $\sigma \notin \text{Tr}$.
4. (quantifiers) $\exists x\, \varphi(x) \in \text{Tr}$ if and only if there is some $n \in N$ such that $\varphi(\bar{n}) \in \text{Tr}$, where $\bar{n}$ is the corresponding term $\underbrace{1 + \cdots + 1}_{n}$ as constructed in $\mathcal{N}$.

Note that Tr is applied only to assertions in the language of arithmetic, not to assertions in the expanded language using the truth predicate itself (but we look at iterated truth predicates in §3). A truth predicate Tr for a model $\mathcal{N} = \langle N, +, \cdot, 0, 1, < \rangle$ is *inductive*, if the expanded structure $\langle N, +, \cdot, 0, 1, <, \text{Tr} \rangle$ satisfies $\text{PA}(\text{Tr})$, the theory of PA in the language augmented with a predicate symbol for Tr, so that mentions of Tr may appear in instances of the induction axiom. The theory of true arithmetic TA^M arising in any model of set theory M is easily seen to be an inductive truth predicate for the corresponding ZFC-standard model of arithmetic $\mathbb{N}^M$ arising in that model of set theory, simply because ZFC proves that $\mathbb{N}$ satisfies the second-order Peano axioms, and so ZFC proves that every subset of $\mathbb{N}$ is inductive.

A principal case of Tarski's theorem on the non-definability of truth is the fact that no model of arithmetic $\mathcal{N}$ can have a truth predicate that is definable in the language of arithmetic. This is simply because for every arithmetic formula $\tau(x)$ there is by the Gödel fixed-point lemma a sentence σ such that $\text{PA} \vdash \sigma \leftrightarrow \neg\tau(\sigma)$, and so we would have either that σ is true in $\mathcal{N}$ while $\tau(\sigma)$ fails, or that σ is false in $\mathcal{N}$ while $\tau(\sigma)$ holds; either of these possibilities would mean that the collection of sentences satisfying τ could not satisfy the recursive Tarskian truth requirements, applied up to the logical complexity of σ, which is a standard-finite sentence.

Krajewski observed that a model of arithmetic can have different incompatible truth predicates, a fact we find illuminating for the context of this paper, and so we presently give an account of it. The argument is pleasantly classical, relying principally only on Beth's implicit definability theorem and Tarski's theorem on the non-definability of truth.

THEOREM 2 [19, 20]. *There are models of arithmetic with different incompatible inductive truth predicates. Indeed, every model of arithmetic $N_0 \models \text{PA}$ that admits an (inductive) truth predicate has an elementary extension N that admits several incompatible (inductive) truth predicates.*

Proof. Let $\mathcal{N}_0 = \langle N_0, +, \cdot, 0, 1, < \rangle$ be any model of arithmetic that admits a truth predicate (for example, the standard model $\mathbb{N}^M$ arising in any model of set theory M). Let T be the theory consisting of the elementary diagram $\Delta(\mathcal{N}_0)$ of this model, in the language of arithmetic with constants for every element of N_0, together with the assertion "Tr is a truth predicate," which is expressible as a single assertion about Tr, namely, the assertion that it satisfies the recursive Tarskian truth requirements. The theory T is consistent, because by assumption, $\mathcal{N}_0$ itself admits a truth predicate. Furthermore, any model of the theory T provides an elementary extension $\mathcal{N}$ of $\mathcal{N}_0$, when reduced to the language of arithmetic, together with a truth predicate for $\mathcal{N}$. Suppose toward contradiction that every elementary extension of $\mathcal{N}_0$ that admits a truth predicate has a unique such class. It follows that any two models of T with the same reduction to the language of the diagram of $\mathcal{N}_0$ must have the same interpretation for the predicate Tr. Thus, the predicate Tr is implicitly definable in T, in the sense of

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

the Beth implicit definability theorem (see [1, theorem 2.2.22]), and so by that theorem, the predicate Tr must be explicitly definable in any model of T by a formula in the base language, the language of arithmetic with constants for elements of $\mathcal{N}_0$. But this violates Tarski's theorem on the non-definability of truth, which implies that no model of arithmetic can have a definable truth predicate. Thus, there must be models of T with identical reductions $\mathcal{N}$ to the language of the diagram of $\mathcal{N}_0$, but different truth predicates on $\mathcal{N}$. In other words, $\mathcal{N}$ is an elementary extension of $\mathcal{N}_0$ having at least two different incompatible truth predicates. (And it is not difficult by similar reasoning to see that there must be an $\mathcal{N}$ having infinitely many distinct truth predicates.)

In the case where $\mathcal{N}_0$ admits an inductive full satisfaction class, we simply add $\mathrm{PA}(\mathrm{Tr})$ to the theory T, with the result by the same reasoning that the elementary extension $\mathcal{N}$ will also have multiple inductive full satisfaction classes, as desired. □

In fact, Krajewski [20] proves that we may find elementary extensions $\mathcal{N}$ having at least any desired cardinal κ many such full satisfaction classes.

We observe next the circumstances under which these various satisfaction classes can become the true theory of arithmetic inside a model of set theory. First, let's note the circumstances under which a model of arithmetic is a ZFC-standard model of arithmetic. Let $\mathrm{Th}(\mathbb{N})^{\mathrm{ZFC}}$ be the set of sentences σ in the language of arithmetic, such that $\mathrm{ZFC} \vdash (\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle \models \sigma)$. These are the arithmetic consequences of ZFC, the sentences that hold in every ZFC-standard model of arithmetic. If ZFC is consistent, then so is this theory, since it holds in the standard model of arithmetic of any model of ZFC. More generally, for any set theory T proving the existence of the standard model $\mathbb{N}$ and its theory TA, we have the theory $\mathrm{Th}(\mathbb{N})^T$, consisting of the sentences σ in the language of arithmetic that T proves to hold in the standard model. Thanks to Roman Kossak (recently) and Ali Enayat (from some time ago) for discussions concerning the following proposition. The result appears in [2] with further related analysis of the ZFC-standard models of arithmetic.

Proposition 3 (Enayat). *The following are equivalent for any countable nonstandard model of arithmetic* $\mathcal{N}$.

(1) $\mathcal{N}$ *is a* ZFC*-standard model of arithmetic. That is,* $\mathcal{N} = \mathbb{N}^M$ *for some* $M \models \mathrm{ZFC}$.
(2) $\mathcal{N}$ *is a computably saturated model of* $\mathrm{Th}(\mathbb{N})^{\mathrm{ZFC}}$.

Proof. $(1 \to 2)$ Suppose that $\mathcal{N}$ is a countable nonstandard ZFC-standard model of arithmetic, arising as $\mathcal{N} = \mathbb{N}^M$ for some model of set theory $M \models \mathrm{ZFC}$, which we may assume is countable. Clearly, $\mathcal{N}$ satisfies the theory $\mathrm{Th}(\mathbb{N})^{\mathrm{ZFC}}$. Note also that $\mathcal{N}$ admits an inductive satisfaction class, namely, the collection of arithmetic truths TA^M as they are defined inside M. An easy overspill argument shows that any model of PA with an inductive satisfaction class is computably saturated (e.g., [28, lemma 1.1]). Namely, suppose that $p(x, \vec{n})$ is a computable type in the language of arithmetic with parameters $\vec{n} \in \mathcal{N}$ and that p is finitely realized in $\mathcal{N}$. Since the type p is computable, the model $\mathcal{N}$ computes its own version of the type $p^{\mathcal{N}}$, and this will agree with p on all standard formulas. Furthermore, the structure $\left\langle \mathcal{N}, \mathrm{TA}^M \right\rangle$ can see that all the standard-finite initial segments of $p^{\mathcal{N}}$ are satisfiable according to the truth predicate TA^M, which by induction in the meta-theory agrees with actual truth on the standard-finite length formulas. Thus, by overspill (since we have induction for this predicate), it follows that some nonstandard length initial segment of $p^{\mathcal{N}}$ is satisfied in $\mathcal{N}$ according

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

to the truth predicate, and since that truth predicate agrees with actual satisfaction on standard-finite formulas, it follows in particular that p itself is satisfied in $\mathcal{N}$. So $\mathcal{N}$ is a countable computably saturated model of $\mathrm{Th}(\mathbb{N})^{\mathrm{ZFC}}$, as desired.

$(2 \to 1)$ Conversely, suppose that $\mathcal{N}$ is a countable computably saturated model of $\mathrm{Th}(\mathbb{N})^{\mathrm{ZFC}}$. Consider the theory T consisting of the ZFC axioms together with the assertions $\sigma^{\mathbb{N}}$ for every sentence $\sigma \in \mathrm{Th}(\mathcal{N})$, which is consistent precisely because $\mathcal{N} \models \mathrm{Th}(\mathbb{N})^{\mathrm{ZFC}}$. Indeed, $\mathcal{N}$ agrees on the consistency of any particular finite subtheory of T, because for any finite fragment φ of ZFC and any arithmetic sentence σ, we have $\mathrm{ZFC} \vdash \sigma^{\mathbb{N}} \to \mathrm{Con}(\varphi + \sigma^{\mathbb{N}})$ by the reflection theorem, and so this statement is true in $\mathcal{N}$, along with its antecedent σ, and so $\varphi + \sigma^{\mathbb{N}}$ is consistent in $\mathcal{N}$.

By computable saturation, the theory T is coded in $\mathcal{N}$, since we may write down a computable type for this, and so there is a nonstandard finite theory $t \in \mathcal{N}$ whose standard part is exactly T.

Since as we have mentioned $\mathcal{N}$ thinks the finite subtheories of T are consistent, by overspill we may assume by cutting down to a nonstandard initial segment if necessary that t is consistent in $\mathcal{N}$. Inside $\mathcal{N}$, we can now build the canonical complete consistent Henkin theory H extending t, and let $M \models H$ be the corresponding Henkin model. In particular, $M \models \mathrm{ZFC}$, since this is a part of t, and so $\mathbb{N}^M$ is a ZFC-standard model of arithmetic and hence computably saturated. Note also that $\mathbb{N}^M$ has the same theory as $\mathcal{N}$, because this is part of t and hence H. The structure $\mathcal{N}$ can construct an isomorphism from itself with an initial segment of $\mathbb{N}^M$, because for every $a \in \mathcal{N}$ it has a Henkin constant $\check{a}$ witnessing $\check{a} = \overbrace{1 + \cdots + 1}^{a}$ and it must be part of the theory that any $x < \check{a}$ is some $\check{b}$ for some $b < a$ in $\mathcal{N}$, since otherwise $\mathcal{N}$ would think H is inconsistent. It follows that $\mathcal{N}$ and $\mathbb{N}^M$ have the same standard system. But any two countable computably saturated models of arithmetic with the same standard system are isomorphic, by the usual back-and-forth argument, and so $\mathcal{N} \cong \mathbb{N}^M$, showing that $\mathcal{N}$ is ZFC-standard, as desired. □

Thus, the countable nonstandard ZFC-standard models of arithmetic are precisely the countable computably saturated models of $\mathrm{Th}(\mathbb{N})^{\mathrm{ZFC}}$. One may similarly show that every uncountable model of $\mathrm{Th}(\mathbb{N})^{\mathrm{ZFC}}$ has an elementary extension to a ZFC-standard model of arithmetic, since the theory $\mathrm{ZFC} + \{\, \sigma^{\mathbb{N}} \mid \sigma \in \Delta(\mathcal{N}) \,\}$ is finitely consistent, where $\Delta(\mathcal{N})$ refers to the elementary diagram of $\mathcal{N}$ in the language with constants for every element of $\mathcal{N}$, and any model M of this theory will have $\mathbb{N}^M$ as an elementary extension of $\mathcal{N}$, as desired. The equivalence stated in Proposition 3 does not generalize to uncountable models, for there are uncountable computably saturated models of $\mathrm{Th}(\mathbb{N})^{\mathrm{ZFC}}$ that are ω_1-like and rather classless, which means in particular that they admit no inductive truth predicates and therefore are not ZFC-standard.

We shall now extend Proposition 3 to the case where the model carries a truth predicate. Let $\mathrm{Th}(\mathbb{N}, \mathrm{TA})^{\mathrm{ZFC}}$ be the theory consisting of all sentences σ for which $\mathrm{ZFC} \vdash (\langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{TA} \rangle \models \sigma)$, where σ is in the language of arithmetic augmented with a truth predicate and TA refers to the ZFC-definable set of true arithmetic assertions. If ZFC is consistent, then so is $\mathrm{Th}(\mathbb{N}, \mathrm{TA})^{\mathrm{ZFC}}$, since it holds in the standard model of arithmetic, with the standard interpretation of TA, arising inside any model of ZFC. Note that $\mathrm{Th}(\mathbb{N}, \mathrm{TA})^{\mathrm{ZFC}}$ includes the assertion that TA is a truth predicate, as well as the induction scheme in the language with this predicate.

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

PROPOSITION 4. *The following are equivalent for any countable nonstandard model of arithmetic $\mathcal{N}$ with a truth predicate* Tr.

(1) $\langle \mathcal{N}, \mathrm{Tr}\rangle$ *is a* ZFC*-standard model of arithmetic and arithmetic truth. That is,* $\mathcal{N} = \mathbb{N}^M = \langle \mathbb{N}, +, \cdot, 0, 1, <\rangle^M$ *for some* $M \models \mathrm{ZFC}$ *in which* $\mathrm{Tr} = \mathrm{TA}^M$ *is the theory of true arithmetic.*
(2) $\langle \mathcal{N}, \mathrm{Tr}\rangle$ *is a computably saturated model of* $\mathrm{Th}(\mathbb{N}, \mathrm{TA})^{\mathrm{ZFC}}$.

Proof. The proof of Proposition 3 adapts to accommodate the expanded structure. If a countable nonstandard model $\langle \mathcal{N}, \mathrm{Tr}\rangle$ arises as $\langle \mathbb{N}, \mathrm{TA}\rangle^M$ for some $M \models \mathrm{ZFC}$, then it admits an inductive truth predicate in the expanded language, and this implies that it is computably saturated just as above. Conversely, any countably computably saturated model of $\mathrm{Th}(\mathbb{N}, \mathrm{TA})^{\mathrm{ZFC}}$ can build as before a model of the corresponding Henkin theory extending $\mathrm{ZFC} + \{\, \sigma^{\langle \mathbb{N}, \mathrm{TA}\rangle} \mid \langle \mathcal{N}, \mathrm{Tr}\rangle \models \sigma \,\}$. The corresponding Henkin model M will have $\langle \mathcal{N}, \mathrm{Tr}\rangle$ as an initial segment of $\langle \mathbb{N}^M, \mathrm{TA}\rangle$, and so these two models have the same standard system, and since they also are elementarily equivalent and computably saturated, they are isomorphic by the back-and-forth construction. So $\langle \mathcal{N}, \mathrm{Tr}\rangle$ is ZFC-standard. □

Let us now finally prove Theorem 1. We shall give two proofs, one as a corollary to Proposition 4, and another simpler direct proof.

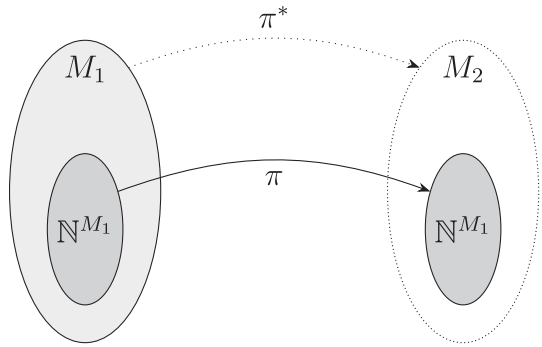
Proof of Theorem 1 (as corollary to Proposition 4). By Beth's theorem as in the Proof of Theorem 2, we may find two different truth predicates on the same model of arithmetic, with both $\langle \mathcal{N}, \mathrm{Tr}_1\rangle$ and $\langle \mathcal{N}, \mathrm{Tr}_2\rangle$ being computably saturated models of $\mathrm{Th}(\mathbb{N}, \mathrm{TA})^{\mathrm{ZFC}}$. It follows by Proposition 4 that they arise as the standard model inside two different models of set theory, with $\mathcal{N} = \mathbb{N}^{M_1} = \mathbb{N}^{M_2}$ and $\mathrm{Tr}_1 = \mathrm{TA}^{M_1}$ and $\mathrm{Tr}_2 = \mathrm{TA}^{M_2}$, establishing theorem 1. □

We also give a simpler direct proof, as follows. (Thanks to W. Hugh Woodin for pointing out a further simplification in this argument.)

Proof of Theorem 1 (direct argument). Suppose that M_1 is any countable ω-nonstandard model of set theory. It follows that M_1's version of the standard model of arithmetic $\langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{TA}\rangle^{M_1}$, augmented with what M_1 thinks is the true theory of arithmetic, is countable and computably saturated. Since TA^{M_1} is not definable in the reduced structure $\mathbb{N}^{M_1} = \langle \mathbb{N}, +, \cdot, 0, 1, <\rangle^{M_1}$, it follows by saturation that there is an arithmetic sentence $\sigma \in \mathrm{TA}^{M_1}$ realizing the same 1-type in $\mathbb{N}^{M_1}$ as another sentence τ in $\mathbb{N}^{M_1}$ with $\tau \notin \mathrm{TA}^{M_1}$; that is, $\mathbb{N}^{M_1} \models \varphi(\sigma) \leftrightarrow \varphi(\tau)$ for every formula φ in the language of arithmetic. To see this, consider the type $p(s, t)$ containing all formulas $\varphi(s) \leftrightarrow \varphi(t)$ for φ in the language of arithmetic, plus the assertions $s \in \mathrm{TA}$ and $t \notin \mathrm{TA}$. This is a computable type, and it is finitely realized in $\langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{TA}\rangle^{M_1}$, precisely because otherwise we would be able to define TA^{M_1} in $\mathbb{N}^{M_1}$. Thus, there are arithmetic sentences σ and τ in M_1, whose Gödel codes realize the same 1-type in $\mathbb{N}^{M_1}$, but such that M_1 thinks that σ is true and τ is false in $\mathbb{N}^{M_1}$. Since these objects have the same 1-type in $\mathbb{N}^{M_1}$, it follows by the back-and-forth construction that there is an automorphism $\pi : \mathbb{N}^{M_1} \to \mathbb{N}^{M_1}$ with $\pi(\tau) = \sigma$.

Since M_1 sits over $\mathbb{N}^{M_1}$ in the way that it does and $\mathbb{N}^{M_1}$ is isomorphic to its image under π, we may build a copy M_2 of M_1 in such a way that extends the isomorphism π.

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press



That we may do so is an instance of the trivial general observation about structures that whenever you have made an isomorphic copy of part of a structure, then you may extend this to an isomorphic copy of the whole structure. In our case, we had an isomorphism π of $\mathbb{N}^{M_1}$ with its copy (which happened to be $\mathbb{N}^{M_1}$ itself), and so we may extend this to an isomorphism π^* of M_1 with a suitable structure M_2 extending $\mathbb{N}^{M_1}$. So we have a model M_2 which is a copy of M_1, witnessed by the isomorphism $\pi^* : M_1 \to M_2$ extending π, in such a way that every element $m \in \mathbb{N}^{M_1}$ sits inside M_1 the same way that $\pi(m)$ sits inside M_2.

Since π was an automorphism of $\mathbb{N}^{M_1}$, the situation is therefore that M_1 and M_2 have exactly the same natural numbers $\mathbb{N}^{M_1} = \mathbb{N}^{M_2}$. Yet, M_1 thinks that the arithmetic sentence σ is true in the natural numbers, while M_2 thinks σ is false, because M_1 thinks τ is false, and so M_2 thinks $\pi(\tau)$ is false, but $\pi(\tau) = \sigma$. □

§3. Satisfacton is not absolute. In this section we aim to show that the non-absoluteness phenomenon is pervasive. For any sufficiently rich structure $\mathcal{N}$ in any countable model of set theory M, there are elementary extensions M_1 and M_2, which have the structure $\mathcal{N}^{M_1} = \mathcal{N}^{M_2}$ in common, yet disagree about the satisfaction relation for this structure, in that M_1 thinks $\mathcal{N} \models \sigma[\vec{a}]$ for some formula σ and parameters $\vec{a}$, while M_2 thinks that $\mathcal{N} \models \neg\sigma[\vec{a}]$ (see Corollary 6). Even more generally, Theorem 5 shows that for any non-definable class S in any structure $\mathcal{N}$ in any countable model of set theory M, there are elementary extensions M_1 and M_2 of M, which have the structure $\mathcal{N}^{M_1} = \mathcal{N}^{M_2}$ in common, yet disagree on the interpretation of the class $S^{M_1} \neq S^{M_2}$. We provide several curious instances as corollary applications.

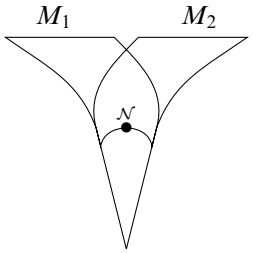
THEOREM 5. *Suppose that M is a countable model of set theory, $\mathcal{N}$ is structure in M in a finite language and $S \subseteq \mathcal{N}$ is an additional predicate, $S \in M$, that is not definable in $\mathcal{N}$ with parameters. Then there are elementary extensions $M \prec M_1$ and $M \prec M_2$, which agree on the structure $\mathcal{N}^{M_1} = \mathcal{N}^{M_2}$, having all the same elements of it, the same language and the same interpretations for the functions, relations and constants in this language, yet disagree on the extension of the additional predicate $S^{M_1} \neq S^{M_2}$, even though by elementarity these predicates have all the properties in M_1 and M_2, respectively, that S has in M.*

Proof. Fix such a structure $\mathcal{N}$ and non-definable class S inside a countable model of set theory M. Let M_1 be any countable computably saturated elementary extension of M, in the language with constants for every element of M. In particular, the structure $\langle \mathcal{N}, S, m\rangle^{M_1}_{m\in\mathcal{N}^M}$ is countable and computably saturated (and this is all we actually

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

require). Since S is not definable from parameters, there are objects s and t in $\mathcal{N}^{M_1}$ with the same 1-type in $\langle \mathcal{N}, m\rangle^{M_1}_{m\in\mathcal{N}^M}$, yet $s \in S^{M_1}$ and $t \notin S^{M_1}$. Since this latter structure is countable and computably saturated, it follows that there is an automorphism $\pi : \mathcal{N} \to \mathcal{N}$ with $\pi(t) = s$ and $\pi(m) = m$ for every $m \in \mathcal{N}^M$. That is, π is an automorphism of $\mathcal{N}^{M_1}$ mapping t to s, and respecting the copy of $\mathcal{N}^M$ inside $\mathcal{N}^{M_1}$. Let M_2 be a copy of M_1 containing M, witnessed by an isomorphism $\pi^* : M_1 \to M_2$ extending π and fixing the elements of M. It follows that $\mathcal{N}^{M_1} = \mathcal{N}^{M_2}$ and M_1 thinks $s \in S$, but since M_1 thinks $t \notin S$, it follows that M_2 thinks $\pi(t) \notin S$ and so M_2 thinks $s \notin S$, as desired. □

COROLLARY 6. *Suppose that M is a countable model of set theory, and that $\mathcal{N}$ is a sufficiently robust structure in M, in a finite language. Then there are elementary extensions $M \prec M_1$ and $M \prec M_2$, which agree on the natural numbers $\mathbb{N}^{M_1} = \mathbb{N}^{M_2}$ and on the structure $\mathcal{N}^{M_1} = \mathcal{N}^{M_2}$, having all the same elements of it, the same language and the same interpretations for the functions, relations and constants in this language, yet they disagree on what they each think is the standard satisfaction relation $\mathcal{N} \models \sigma[\vec{a}]$ for this structure.*



$M \prec M_1, M_2 \models \text{ZFC}$

$\mathbb{N}^{M_1} = \mathbb{N}^{M_2}, \qquad \mathcal{N}^{M_1} = \mathcal{N}^{M_2}$

there are σ *and* $\vec{a}$ *for which*

M_1 *believes* $\mathcal{N} \models \sigma[\vec{a}]$

M_2 *believes* $\mathcal{N} \models \neg\sigma[\vec{a}]$.

Proof. By 'sufficiently robust' we mean that $\mathcal{N} = \langle N, \ldots\rangle$ interprets the standard model of arithmetic $\mathbb{N}$, so that it can handle the Gödel coding of formulas, and also that it has a definable pairing function, so that it contains (Gödel codes for) all finite tuples of its elements. We are assuming that $\mathcal{N}$ is sufficiently robust from the perspective of the model M. It follows by Tarski's theorem on the non-definability of truth that the satisfaction relation of M for $\mathcal{N}$—that is, the relation $\text{Sat}(\varphi, \vec{a})$, which holds just in case $\mathcal{N} \models \varphi[\vec{a}]$ from the perspective of M—is not definable in $\mathcal{N}$. And so the current theorem is a consequence of Theorem 5, using that relation. □

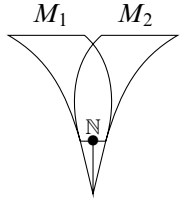
Many of the various examples of non-absoluteness that we have mentioned in this article can now be seen as instances of Corollary 6, as in the following further corollary. In particular, statement (2) of Corollary 7 is a strengthening of Theorem 1, since we now get the models M_1 and M_2 as elementary extensions of any given countable model M.

COROLLARY 7. *Every countable model of set theory M has elementary extensions $M \prec M_1$ and $M \prec M_2$, respectively, in each case, which…*

(1) *agree on their natural numbers with successor and order $\langle \mathbb{N}, S, <\rangle^{M_1} = \langle \mathbb{N}, S, <\rangle^{M_2}$, but which disagree on the even numbers, the prime numbers and*

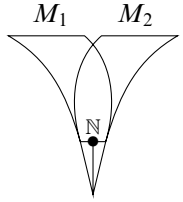
https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

the powers of two, so that M_1 thinks some n *is a large odd prime number, but M_2 thinks it is a large power of* 2.



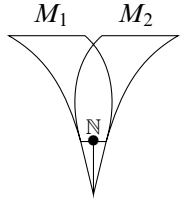
$M_1, M_2 \models \mathrm{ZFC}$
$\langle \mathbb{N}, S, < \rangle^{M_1} = \langle \mathbb{N}, S, < \rangle^{M_2}$
M_1 *believes* $\mathbb{N} \models n$ *is an odd prime*
M_2 *believes* $\mathbb{N} \models n = 2^k$ *for some* k.

(2) *agree on their natural numbers with successor, addition and order* $\langle \mathbb{N}, S, +, < \rangle^{M_1} = \langle \mathbb{N}, S, +, < \rangle^{M_2}$, *but which disagree on natural-number multiplication, so that M_1 thinks $a \cdot b = c$ for some particular natural numbers, but M_2 disagrees.*



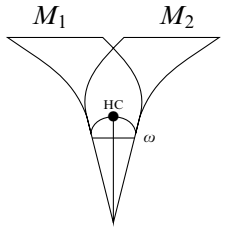
$M_1, M_2 \models \mathrm{ZFC}$
$\langle \mathbb{N}, S, +, < \rangle^{M_1} = \langle \mathbb{N}, S, +, < \rangle^{M_2}$
M_1 *believes* $\mathbb{N} \models a \cdot b = c$
M_2 *believes* $\mathbb{N} \models a \cdot b \neq c$.

(3) *agree on their standard model of arithmetic* $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle^{M_1} = \langle \mathbb{N}, +, \cdot, 0, 1, < \rangle^{M_2}$, *but which disagree on their theories of arithmetic truth.*



$M_1, M_2 \models \mathrm{ZFC}$
$\mathbb{N}^{M_1} = \mathbb{N}^{M_2}$
M_1 *believes* $\mathbb{N} \models \sigma$
M_2 *believes* $\mathbb{N} \models \neg\sigma$.

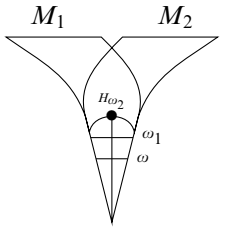
(4) *agree on their natural numbers* $\mathbb{N}^{M_1} = \mathbb{N}^{M_2}$, *their reals* $\mathbb{R}^{M_1} = \mathbb{R}^{M_2}$ *and their hereditarily countable sets* $\langle \mathrm{HC}, \in \rangle^{M_1} = \langle \mathrm{HC}, \in \rangle^{M_2}$, *but which disagree on their theories of projective truth.*



$M_1, M_2 \models \mathrm{ZFC}$
$\mathbb{N}^{M_1} = \mathbb{N}^{M_2}$ $\quad$ $\mathbb{R}^{M_1} = \mathbb{R}^{M_2}$
$\langle \mathrm{HC}, \in \rangle^{M_1} = \langle \mathrm{HC}, \in \rangle^{M_2}$
M_1 *believes* $\mathrm{HC} \models \sigma$
M_2 *believes* $\mathrm{HC} \models \neg\sigma$.

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

(5) *agree on the structure* $\langle H_{\omega_2}, \in\rangle^{M_1} = \langle H_{\omega_2}, \in\rangle^{M_2}$, *but which disagree on truth in this structure.*



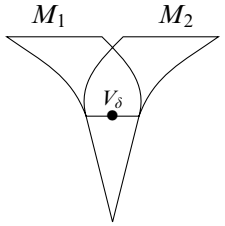
$M_1, M_2 \models \text{ZFC}$

$\langle H_{\omega_2}, \in\rangle^{M_1} = \langle H_{\omega_2}, \in\rangle^{M_2}$

M_1 *believes* $H_{\omega_2} \models \sigma$

M_2 *believes* $H_{\omega_2} \models \neg\sigma$.

(6) *have a transitive rank-initial segment* $\langle V_\delta, \in\rangle^{M_1} = \langle V_\delta, \in\rangle^{M_2}$ *in common, but which disagree on truth in this structure.*



$M_1, M_2 \models \text{ZFC}$

$V_\delta^{M_1} = V_\delta^{M_2} \models \text{ZFC}$

M_1 *believes* $V_\delta \models \sigma$

M_2 *believes* $V_\delta \models \neg\sigma$.

Proof. What we mean by "respectively" is that each case may be exhibited separately, using different pairs of extensions M_1 and M_2. Each statement in the theorem is an immediate consequence of Corollary 6 or of Theorem 5. Thanks to Roman Kossak for pointing out statement (2), which follows from Theorem 5 because multiplication is not definable in Presburger arithmetic $\langle \mathbb{N}, +, <\rangle$, as the latter is a decidable theory. Statement (1) is proved similarly, since the set of primes and the powers of two are not definable in $\langle \mathbb{N}, S, <\rangle$, since any nonstandard model of this structure consists of infinitely many $\mathbb{Z}$ chains above the standard part, and these admit numerous automorphisms by translation. By picking a suitable translation, we may move a power of two to an odd prime, and apply the argument in the Proof of Theorem 1. For statement (3), we apply Corollary 6 to the structure $\mathcal{N} = \langle \mathbb{N}, +, \cdot, 0, 1, <\rangle^M$. For statement (4), one should clarify exactly what is meant by 'projective truth,' since it can be viewed variously as the full second-order theory of the standard model of arithmetic, using $P(\mathbb{N})^M$, or as the theory of the ordered real field $\langle \mathbb{R}, \mathbb{Z}, +, \cdot, 0, 1, <\rangle$ with a predicate for the integers, or as the theory of the set-theoretic structure $\langle V_{\omega+1}, \in\rangle$, or of the structure of hereditarily countable sets $\text{HC} = \langle \text{HC}, \in\rangle$. Nevertheless, these models are all bi-interpretable with each other, and one can view projective truth as the satisfaction relation for any of them; in each case, we can make the conclusion of statement (4) by using that structure in Corollary 6. Notice that one can similarly view arithmetic truth as residing in the structure of hereditarily finite sets $\langle \text{HF}, \in\rangle$, which is mutually interpretable with the standard model of arithmetic via the Ackermann encoding, and in this case, statements (3), (4) and (5) can be seen as a progression, concerning the structures H_ω, H_{ω_1} and H_{ω_2}, a progression which continues, of course, to higher orders in set theory. Statement (6) is similarly an immediate consequence of Corollary 6, using the structure $\mathcal{N} = \langle V_\delta, \in\rangle^M$. □

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

Let us explore a bit further how indefiniteness arises in the iterated truth-about-truth hierarchy. Beginning with the standard model of arithmetic $\mathbb{N}_0 = \langle \mathbb{N}, +, \cdot, 0, 1, < \rangle$, we may define the standard truth predicate Tr_0 for assertions in the language of arithmetic, and consider the structure $\mathbb{N}_1 = \langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{Tr}_0 \rangle$, in the expanded language with a predicate for the truth of arithmetic assertions. In this expanded language, we may make assertions both about arithmetic and about arithmetic truth. Climbing atop this structure, let $\mathrm{Tr}_1 = \mathrm{Th}(\mathbb{N}_1)$ be its theory and form the next level of the iterated truths-about-truth hierarchy $\mathbb{N}_2 = \langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{Tr}_0, \mathrm{Tr}_1 \rangle$ by appending this new truths-about-truth predicate. We may easily continue in this way, building the finite levels of the hierarchy, each new truth predicate telling us about the truth of arithmetic assertions involving only previous levels of the iterated truth hierarchy. There is a rich literature on various aspects of this iterated truth hierarchy, from Tarski [29] to Kripke [21] and many others, including the related development of the revision theory of truth [6]. Feferman [3] treats the iteration of truth as an example where once one accepts certain statements about $\langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{Tr}_0, \dots, \mathrm{Tr}_n \rangle$, then one ought accept certain other statements about $\langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{Tr}_0, \dots, \mathrm{Tr}_{n+1} \rangle$.

COROLLARY 8. *For every countable model of set theory M and any natural number* n, *there are elementary extensions M_1 and M_2 of M, which have the same natural numbers $\mathbb{N}^{M_1} = \mathbb{N}^{M_2}$, the same iterated arithmetic truth predicates $(\mathrm{Tr}_k)^{M_1} = (\mathrm{Tr}_k)^{M_2}$ for $k < n$ and hence the same iterated truth structure up to* n,

$$\langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{Tr}_0, \dots, \mathrm{Tr}_{n-1} \rangle^{M_1} = \langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{Tr}_0, \dots, \mathrm{Tr}_{n-1} \rangle^{M_2},$$

but which disagree on the theory of this structure, and hence disagree on the next order of truth, $(\mathrm{Tr}_n)^{M_1} \neq (\mathrm{Tr}_n)^{M_2}$.

Proof. In other words, we will have $(\mathbb{N}_n)^{M_1} = (\mathbb{N}_n)^{M_2}$, yet $(\mathrm{Tr}_n)^{M_1} \neq (\mathrm{Tr}_n)^{M_2}$. This is an immediate consequence of Corollary 6, using the model $\mathcal{N} = (\mathbb{N}_n)^M$. □

In particular, even the cases $n = 0, 1$ or 2 are interesting. The case $n = 0$ amounts to Theorem 1, and the case $n = 1$ shows that one can have models of set theory M_1 and M_2 which have the same standard model of arithmetic $\mathbb{N}^{M_1} = \mathbb{N}^{M_2}$ and the same arithmetic truth $\mathrm{Tr}_0^{M_1} = \mathrm{Tr}_0^{M_2}$, yet disagree in their theory of the theory of arithmetic truth $\mathrm{Tr}_1^{M_1} \neq \mathrm{Tr}_1^{M_2}$. Thus, even if one assumes a definite nature for the structure of arithmetic, and also for arithmetic truth, then still there is indefiniteness as to the nature of truths about arithmetic truth, and so on throughout the iterated truth-about-truth hierarchy; indefiniteness arises at any particular level.

The process of iterating the truth hierarchy of course continuous transfinitely, for as long as we have some natural way of representing the ordinals inside $\mathbb{N}$, in order to undertake the Gödel coding of formulas in the expanded language and retain the truth predicates as subclasses of $\mathbb{N}$. If α is any computable ordinal, for example, then we have a representation of α inside $\mathbb{N}$ using a computable relation of order type α, and we may develop a natural Gödel coding for ordinals up to α and formulas in the language $\{ +, \cdot, 0, 1, <, \mathrm{Tr}_\xi \}_{\xi<\alpha}$. If $\eta < \alpha$ and $\mathbb{N}_\eta = \langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{Tr}_\xi \rangle_{\xi<\eta}$ is defined up to η, then we form $\mathbb{N}_{\eta+1}$ by adding the η^{th} order truth predicate Tr_η for assertions in the language of $\mathbb{N}_\eta$, which can make reference to the simpler truth predicates Tr_ξ for $\xi < \eta$ using the Gödel coding established by the computability of α. The higher levels of this truths-about-truth hierarchy provide truth predicates for assertions about lower-level truths-about-truth for arithmetic. We note that indefiniteness cannot arise

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

at limit ordinal stages, since when λ is a limit ordinal, then a sentence σ is true at stage λ just in case it is true at any stage after which all the truth predicates appearing in σ have arisen. In other words, $\mathrm{Tr}_\lambda = \bigcup_{\xi<\lambda} \mathrm{Tr}_\xi$ for any limit ordinal λ. Nevertheless, we find it likely that there is version of Corollary 8 revealing indefiniteness in the transfinite realm of the iterated truth hierarchy.

We conclude this section with some further applications of Theorem 5, using non-definable predicates other than a satisfaction predicate.

Corollary 9.

(1) *Every countable model of set theory M has elementary extensions $M \prec M_1$ and $M \prec M_2$, which agree on their standard model of arithmetic $\mathbb{N}^{M_1} = \mathbb{N}^{M_2}$ and have a computable linear order $\lhd$ on $\mathbb{N}$ in common, yet M_1 thinks $\langle \mathbb{N}, \lhd\rangle$ is a well-order and M_2 does not.*

(2) *Similarly, every such M has such M_1 and M_2, which agree on their standard model of arithmetic $\mathbb{N}^{M_1} = \mathbb{N}^{M_2}$ and thus agree on the computational behavior of all programs, yet they disagree on Kleene's $\mathcal{O}$, with $\mathcal{O}^{M_1} \neq \mathcal{O}^{M_2}$.*

(3) *Every countable model of set theory M which thinks $0^\sharp$ exists has elementary extensions M_1 and M_2, which agree on the ordinals up to any desired uncountable cardinal $\kappa \in M$, on the constructible universe $L_\kappa^{M_1} = L_\kappa^{M_2}$ up to κ and on the facts that κ is an uncountable cardinal and $0^\sharp$ exists, yet disagree on which ordinals below κ are the Silver indiscernibles. Similarly, we may ensure that they disagree on $0^\sharp$, so that $(0^\sharp)^{M_1} \neq (0^\sharp)^{M_2}$.*

Proof. These are each consequences of Theorem 5. For statement (1), consider the structure $\mathcal{N} = \mathbb{N}^M$, with the predicate $S = \mathrm{WO}^M$, the set of indices of computably enumerable well-orderings on $\mathbb{N}$ in M. This is a Π^1_1-complete set of natural numbers, and hence not first-order definable in the structure $\mathcal{N}$ from the perspective of M. Thus, by Theorem 5, we get models of set theory M_1 and M_2, elementarily extending M, which agree on $\mathbb{N}^{M_1} = \mathbb{N}^{M_2}$, yet disagree on $\mathrm{WO}^{M_1} \neq \mathrm{WO}^{M_2}$. So there is some c.e. relation $\lhd$ in common, yet M_1 (we may assume) thinks it is a well-order and M_2 does not. Since every c.e. order is isomorphic to a computable order, we may furthermore assume that $\lhd$ is computable, and the models will compute it with the same program, and both see that it is a linear order, as desired. Statement (2) follows as a consequence, since using the program computing $\lhd$, we may construct deviations between $\mathcal{O}^{M_1}$ and $\mathcal{O}^{M_2}$; or alternatively, statement (2) follows immediately from Theorem 5, since $\mathcal{O}$ is not definable in $\mathbb{N}$. Statement (3) also follows from Theorem 5, since the class of Silver indiscernibles below κ is not definable in $\langle L_\kappa, \in\rangle$ in M, as from it we could define a truth predicate for that structure. So there must be extensions which agree on $\langle L_\kappa, \in\rangle^{M_1} = \langle L_\kappa, \in\rangle^{M_2}$ and on the fact that κ is an uncountable cardinal and that $0^\sharp$ exists, yet disagree on the Silver indiscernibles below κ. Similarly, s $0^\sharp$ also is not definable in $\langle L_\kappa, \in\rangle$, we may make them disagree on $0^\sharp$. □

We expect that the reader will be able to construct many further instances of the phenomenon. Here is another example we found striking.

Theorem 10. *Every countable model of set theory M has elementary extensions M_1 and M_2, which agree on the structure of their standard natural numbers $\langle \mathbb{N}, +, \cdot, 0, 1, <\rangle^{M_1} = \langle \mathbb{N}, +, \cdot, 0, 1, <\rangle^{M_2}$, and which have a set $A \subseteq \mathbb{N}$ in common, extensionally identical in M_1 and M_2, yet M_1 thinks A is first-order definable in $\mathbb{N}$ and M_2 thinks it is not.*

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

The proof relies on the following lemma, which was conjectured by the first author and asked on MathOverflow [11] specifically in connection with this application. The question was answered there by Andrew Marks, whose proof we adapt here.

LEMMA 10.1 (Andrew Marks [11]). *There is a subset $B \subseteq \mathbb{N} \times \mathbb{N}$, such that the set $\{ n \in \mathbb{N} \mid B_n$ is arithmetic$\}$ is not definable in the structure $\langle \mathbb{N}, +, \cdot, 0, 1, <, B \rangle$, where $B_n = \{ k \mid (n,k) \in B \}$ denotes the n^{th} section of B.*

Proof. We identify $\mathbb{N} \times \mathbb{N}$ with $\mathbb{N}$ via Gödel pairing. We use X' and $X^{(n)}$ to denote the Turing jump of X and the n^{th} Turing jump of X, respectively. Recall that a set $X \subseteq \mathbb{N}$ is *n-generic* if for every Σ^0_n subset $S \subseteq 2^{<\omega}$, there is an initial segment of X that either is in S or has no extension in S. A set is *arithmetically generic* if it is n-generic for every n. It is a standard fact that if X is 1-generic, then $X' \equiv_T 0' \oplus X$, and if X is n-generic and Y is 1-generic relative to $X \oplus 0^{(n-1)}$, then $X \oplus Y$ is n-generic. Let $A = 0^{(\omega)} = \oplus_n 0^{(n)}$, which is Turing equivalent to the set of true sentences of first order arithmetic. We shall construct a set B with the following features:

(1) $\{ n \in \mathbb{N} \mid B_n$ is arithmetic $\} = A$. More specifically, if $n \in A$, then B_n is $(n+1)$-generic and computable from $0^{(n+1)}$, and if $n \notin A$, then B_n is arithmetically generic.
(2) For each natural number k, the set $C_k = \oplus_{i \in \omega} B_{m_i}$ is $(k+1)$-generic, where m_i is the ith element of the set $\{ m \in \mathbb{N} \mid m \notin A$ or $k \leq m \}$.

Any set B with these features, we claim, fulfills the lemma. To see this, we argue first by argue by induction that $B^{(n)} \equiv_T 0^{(n)} \oplus C_n$ for any natural number n. This is immediate for $n = 0$, since $C_0 = \oplus_m B_m$. If $B^{(n)} \equiv_T 0^{(n)} \oplus C_n$, then $B^{(n+1)} \equiv_T (0^{(n)} \oplus C_n)' \equiv_T 0^{(n+1)} \oplus C_n$, since C_n is $(n+1)$-generic and hence 1-generic relative to $0^{(n)}$, but $0^{(n+1)} \oplus C_n \equiv_T 0^{(n+1)} \oplus C_{n+1}$ because either $n \notin A$ and so $C_n = C_{n+1}$ or $n \in A$ so $C_{n+1} \equiv_T B_n \oplus C_n$, since $B_n \leq_T 0^{(n+1)}$. So we have established $B^{(n)} \equiv_T 0^{(n)} \oplus C_n$. Since C_n is $(n+1)$-generic, it follows that $0^{(n)} \oplus C_n$ does not compute $0^{(n+1)}$, and so also $B^{(n)}$ does not compute $0^{(n+1)}$. In particular, $B^{(n)}$ does not compute A for any n, and so A is not arithmetically definable from B, as desired.

It remains to construct the set B with features (1) and (2). We do so in stages, where after stage n we will have completely specified $B_0, B_1, \ldots, B_n$ and finitely much additional information about B on larger coordinates. To begin, let B_0 be any set satisfying the requirement of condition (1). We will ensure inductively that after each stage n, the set $C_{k,n} = B_{m_0} \oplus \ldots \oplus B_{m_j}$ is $(k+1)$-generic, where $k \leq n$ and $m_0, \ldots, m_j$ are the elements of $\{ m \in \mathbb{N} \mid m \notin A$ or $k \leq m \} \cap \{0, \ldots, n\}$.

At stage $n > 0$, for each of the finitely many pairs (i,k) with $i, k < n$, we let $S_{i,k}$ be the ith Σ^0_{k+1} subset of $2^{<\omega}$, and if possible, we make a finite extension to our current approximation to B so that the resulting approximation to C_k extends an element of $S_{i,k}$, thereby ensuring this instance of (2). If there is no such extension, then since inductively $C_{k,n-1}$ is $(k+1)$-generic, there is already a finite part of our current approximation to B that cannot be extended to extend an element of $S_{i,k}$, and this also ensures this instance of (2).

We complete stage n by specifying B_n. If $n \notin A$, then we simply extend our current approximation to B by ensuring that B_n is arithmetically generic relative to $B_0 \oplus \ldots \oplus B_{n-1}$. This ensures this instance of (1) while maintaining our induction assumption that $C_{k,n}$ is $(k+1)$-generic for each $k < n$, since $C_{k,n-1}$ is $(k+1)$-generic and B_n is $(k+1)$-generic relative to it; and similarly, $C_{n,n}$ is now $(n+1)$-generic. If $n \in A$, then we let

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

B_n be any $0^{(n+1)}$-computable $(n+1)$-generic set extending the finitely many bits of B_n specified in the current approximation. For each $k < n$, let $j_0, \ldots, j_t$ be the elements of A in the interval $[k, n)$. Since our new B_n is 1-generic relative to $0^{(n)}$, which can compute $B_{j_0} \oplus \ldots \oplus B_{j_t}$, it follows that that $B_{j_0} \oplus \ldots \oplus B_{j_t} \oplus B_n$ is $(k+1)$-generic, and so $C_{k,n}$ is $(k+1)$-generic, as the remaining elements in the finite join defining $C_{k,n}$ are mutually arithmetically generic with this; and since $C_{n,n}$ is $n+1$-generic, we maintain our induction assumption.

This completes the construction. We have fulfilled (1) explicitly by the choice of B_n, and we fulfilled (2) by systematically deciding all the required sets $S_{i,k}$. □

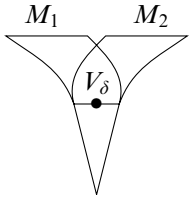
Proof of Theorem 10. Fix any countable model of set theory M. Apply Lemma 10.1 inside M, to find a predicate $B \subseteq \mathbb{N} \times \mathbb{N}$ in M, such that the set $S = \{ n \in \mathbb{N} \mid B_n \text{ is arithmetic} \}$ is not definable in the structure $\mathcal{N} = \langle \mathbb{N}, +, \cdot, 0, 1, <, B \rangle$ from the perspective of M. It follows by Theorem 5 applied to this structure that there are elementary extensions M_1 and M_2 of M, which agree on $\mathcal{N}^{M_1} = \mathcal{N}^{M_2}$ and in particular on $\mathbb{N}^{M_1} = \mathbb{N}^{M_2}$ and $B^{M_1} = B^{M_2}$, but not on $S^{M_1} \neq S^{M_2}$. In particular, there is some section $A = B_n$ that is arithmetic in M_1 (we may assume), but not in M_2. But since the models agree on the predicate B, they agree on all the sections of B and in particular have the set A extensionally in common. □

§4. Indefiniteness for specific types of sentences. Earlier in this article, we proved that the satisfaction relation $\mathcal{N} \models \sigma$ for a first-order structure $\mathcal{N}$ is not generally absolute between the various models of set theory containing that model and able to express this satisfaction relation. But the proofs of non-absoluteness did not generally reveal any specific nature for the sentences on which truth can differ in different models of set theory. We should now like to address this issue by presenting an alternative elementary proof of non-absoluteness, using reflection and compactness, which shows that the theory of a structure can vary on sentences whose specific nature we can identify.

A cardinal δ is defined to be Σ_n-*correct*, if $V_\delta \prec_{\Sigma_n} V$. The reflection theorem shows that there is a proper class club $C^{(n)}$ of such cardinals. The cardinal δ is *fully correct*, if it is Σ_n-correct for every n. This latter notion is not expressible as a single assertion in the first-order language of ZFC, but one may express it as a scheme of assertions about δ, in a language with a constant for δ. Namely, let "$V_\delta \prec V$" denote the theory asserting of every formula φ in the language of set theory, that $\forall x \in V_\delta [\varphi(x) \leftrightarrow \varphi(x)^{V_\delta}]$. Since every finite subtheory of this scheme is proved consistent in ZFC by the reflection theorem, it follows that the theory is finitely consistent and so, by the compactness theorem, ZFC + $V_\delta \prec V$ is equiconsistent with ZFC. The reader may note that since $V_\delta \prec V$ asserts the elementarity separately for each formula, we may not deduce in this theory that $V_\delta \models$ ZFC, but rather only for each axiom of ZFC separately, that V_δ satisfies that axiom.

THEOREM 11. *Every countable model of set theory* $M \models$ ZFC *has elementary extensions* M_1 *and* M_2, *with a transitive rank-initial segment* $\langle V_\delta, \in \rangle^{M_1} = \langle V_\delta, \in \rangle^{M_2}$ *in common, such that* M_1 *thinks that the least natural number* n *for which* V_δ *violates* Σ_n*-collection is even, but* M_2 *thinks it is odd.*

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press



$M_1, M_2 \models \text{ZFC}$
$V_\delta^{M_1} = V_\delta^{M_2}$
n *is least with* $\neg\Sigma_n$*-collection in* V_δ
M_1 *believes* n *is even*
M_2 *believes* n *is odd*.

Proof. Suppose that M is a countable model of ZFC, and consider the theory:

$$T_1 = \Delta(M) + V_\delta \prec V \ + \ \{\, m \in V_\delta \mid m \in M \,\} \\ + \text{ the least } n \text{ such that } V_\delta \not\models \Sigma_n\text{-collection is even},$$

where $\Delta(M)$ is the elementary diagram of M, in the language of set theory having constants for every element of M. Note that the first three components of T_1 as it is described above are each infinite schemes, whereas the final assertion "the least n...," which we take also to assert that there is such an n, is expressible as a single sentence about δ in the language of set theory, using the ZFC-definable satisfaction relation for $\langle V_\delta, \in\rangle$. We claim that this theory is consistent. Consider any finite subtheory $t \subseteq T_1$. We shall find a δ in M such that M with this δ will satisfy every assertion in t. Let k be a sufficiently large odd number, so that every formula appearing in any part of t has complexity at most Σ_k, and let δ be the next Σ_k-correct cardinal in M above the largest rank of an element of M whose constant appears in t. We claim now that $\langle M, \in^M, \delta, m\rangle_{m\in M} \models t$. First, it clearly satisfies all of $\Delta(M)$; and since δ is Σ_k-correct in M, we have $V_\delta^M \prec_{\Sigma_k} M$, and since also δ was large enough to be above any of the constants of M appearing in t, we attain any instances from the first three components of the theory that are in the finite sub-theory t; since δ is Σ_k-correct in M, it follows that V_δ^M satisfies every instance of Σ_k-collection; but since δ is not a limit of Σ_k-correct cardinals (since it is the "next" one after a certain ordinal), it follows that V_δ^M does not satisfy Σ_{k+1}-collection, and so the least n such that $V_\delta^M \not\models \Sigma_n$-collection is precisely $n = k + 1$, which is even. So T_1 is finitely consistent and thus consistent. Similarly, the theory

$$T_2 = \Delta(M) + V_\delta \prec V \ + \ \{\, m \in V_\delta \mid m \in M \,\} \\ + \text{ the least } n \text{ such that } V_\delta \not\models \Sigma_n\text{-collection is odd}$$

is also consistent.

Let $\langle M_1, M_2\rangle$ be a computably saturated model pair, such that $M_1 \models T_1$ and $M_2 \models T_2$. It follows that $\left\langle V_\delta^{M_1}, V_\delta^{M_2}\right\rangle$ is a computably saturated model pair of elementary extensions of $\langle M, \in^M\rangle$, which are therefore elementarily equivalent in the language of set theory with constants for elements of M, and hence isomorphic by an isomorphism respecting those constants. So we may assume without loss of generality that $\langle M, \in^M\rangle \prec \langle V_\delta, \in\rangle^{M_1} = \langle V_\delta, \in\rangle^{M_2}$. Meanwhile, M_1 thinks that this V_δ violates Σ_n-collection first at an even n and M_2 thinks it does so first for an odd n, since these assertions are part of the theories T_1 and T_2, respectively. □

An alternative version of Theorem 11, with essentially the same proof, produces elementary extensions M_1 and M_2 of M with a rank initial segment $V_\delta^{M_1} = V_\delta^{M_2}$ in

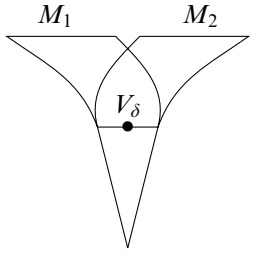
https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

common, but M_1 thinks the least n for which V_δ is not Σ_n-correct is even, but M_2 thinks it is odd.

By looking not just at the parity of the least n where Σ_n-collection (or Σ_n-correctness) fails, but rather, say, at the k^{th} binary digit, we can easily make infinitely many different elementary extensions $M_1, M_2, \ldots$ of M, with the natural numbers $\mathbb{N}^{M_k}$ and $V_\delta^{M_k}$ all in common, but such that M_k thinks that the least n for which this V_δ violates Σ_n-collection is a number with exactly k many prime factors. In particular, even though they have the same structure $\langle V_\delta, \in\rangle^{M_k}$, they each think specific incompatible things about the theory of this structure.

§5. "Being a model of ZFC" is not absolute. In this section, we prove that the question of whether a given transitive rank initial segment V_δ of the universe is a model of ZFC is not absolute between models of set theory with that rank initial segment in common. Recall that a cardinal δ is *worldly*, if $V_\delta \models \text{ZFC}$.

THEOREM 12. *If M is a countable model of set theory in which the worldly cardinals form a stationary proper class (it would suffice, for example, that $M \models \text{Ord}$ is inaccessible), then there are elementary extensions M_1 and M_2, which have a transitive rank initial segment $\langle V_\delta, \in\rangle^{M_1} = \langle V_\delta, \in\rangle^{M_2}$ in common, such that M_1 thinks $V_\delta \models \text{ZFC}$ but M_2 thinks $V_\delta \not\models \text{ZFC}$. Moreover, such extensions can be found for which δ is fully correct in both M_1 and M_2, and furthermore in which $M \prec V_\delta^{M_1} = V_\delta^{M_2}$.*



$M_1, M_2 \models \text{ZFC}$

$\langle V_\delta, \in\rangle^{M_1} = \langle V_\delta, \in\rangle^{M_2}$

M_1 *believes* $V_\delta \models \text{ZFC}$

M_2 *believes* $V_\delta \not\models \text{ZFC}$.

Proof. Fix any countable model $M \models \text{ZFC}$, such that the worldly cardinals form a stationary proper class in M. That is, every definable proper class club $C \subseteq \text{Ord}$ in M contains some δ that is worldly in M. Let T_1 be the theory consisting of the elementary diagram $\Delta(M)$ plus the scheme of assertions $V_\delta \prec V$, in the language with a new constant symbol for δ, plus the assertions $a \in V_\delta$ for each constant symbol a for an element $a \in M$, plus the assertion "δ is worldly," which is to say, the assertion that $V_\delta \models \text{ZFC}$. Suppose that t is a finite subtheory of T_1, which therefore involves only finitely many instances from the $V_\delta \prec V$ scheme. Let n be large enough so that all the formulas φ in these instances have complexity at most Σ_n. Since the Σ_n-correct cardinals form a closed unbounded class and the worldly cardinals are stationary, there is a Σ_n-correct worldly cardinal δ in M. It follows that M satisfies all the formulas in t using this δ, and so the theory T_1 is finitely consistent and hence consistent.

Let T_2 be the theory consisting of the elementary diagram $\Delta(M)$, the scheme $V_\delta \prec V$ plus the assertion "δ is not worldly." This theory also is finitely consistent, since if $t \subseteq T_2$ is finite, then let n be beyond the complexity of any formula appearing as an instance of $V_\delta \prec V$ in t, and let δ be a Σ_n correct cardinal in M that is not worldly (for example,

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

we could let δ be the next Σ_n-correct cardinal after some ordinal; this can never be worldly since V_δ will not satisfy Σ_n-reflection). It follows that M with this δ satisfies every assertion in t, showing that T_2 is finitely consistent and hence consistent.

$$\begin{array}{ccc} & \prec & \langle V_\delta, \in\rangle^{M_1} \prec \langle M_1, \in^{M_1}\rangle \\ \langle M, \in^M\rangle & & = \\ & \succ & \langle V_\delta, \in\rangle^{M_2} \prec \langle M_2, \in^{M_2}\rangle. \end{array}$$

Let $\langle M_1, M_2\rangle$ be a computably saturated model pair, where $M_1 \models T_1$ and $M_2 \models T_2$. It follows that $\left\langle V_\delta^{M_1}, V_\delta^{M_2}\right\rangle$ is also a computably saturated model pair of models of set theory, and these both satisfy the elementary diagram of M. Consequently, they are isomorphic by an isomorphism that respects the interpretation of M in them, and so by replacing with an isomorphic copy, we may assume that $\langle V_\delta, \in\rangle^{M_1} = \langle V_\delta, \in\rangle^{M_2}$. The theories T_1 and T_2 ensure that $V_\delta \prec V$ in both M_1 and M_2, that $M \prec V_\delta$, and furthermore, that $M_1 \models \delta$ *is worldly* and $M_2 \models \delta$ *is not worldly*, or in other words, $M_1 \models (V_\delta \models \text{ZFC})$, but $M_2 \models (V_\delta \not\models \text{ZFC})$, as desired. □

The hypothesis that the worldly cardinals form a stationary proper class is a consequence of the (strictly stronger) Lévy scheme, also known as Ord *is Mahlo*, asserting in effect that the inaccessible cardinals form a stationary proper class. This is in turn strictly weaker in consistency strength than the existence of a single Mahlo cardinal, since if κ is Mahlo, then $V_\kappa \models$ Ord *is Mahlo*. So these are all rather weak large cardinal hypotheses, in terms of the large cardinal hierarchy. Meanwhile, the conclusion already explicitly has large cardinal strength, since $M_1 \models \delta$ is worldly. Furthermore, the "Moreover,..." part of the conclusion makes the hypothesis optimal, since if $M_1 \models \delta$ is worldly and fully correct, then the worldly cardinals of M_1 form a stationary proper class in $V_\delta^{M_1}$, as any definable class club there extends to a class club in M_1 containing δ.

As an example to illustrate the range of non-absoluteness, consider the case where M is a model of ZFC in which the worldly limits of Woodin cardinals are stationary. It follows from Theorem 12 that there are elementary extensions M_1 and M_2 of M with a common rank initial segment $\langle V_\delta, \in\rangle^{M_1} = \langle V_\delta, \in\rangle^{M_2}$, such that

$$M_1 \models (V_\delta \models \text{ZFC} + \text{there are a proper class of Woodin cardinals}),$$

but

$$M_2 \models (V_\delta \not\models \text{ZFC} + \text{there are a proper class of Woodin cardinals}).$$

The proof shows that from the perspective of M_2, the common V_δ does not satisfy Σ_n-replacement for sufficiently large n, even though it satisfies full replacement from the perspective of M_1, which has all the same formulas.

We conclude this section with an observation that many have found curious. We learned this result from Brice Halimi [8], who emphasized that the proof is classical and the result may be folklore. If $\mathcal{M} = \langle M, \in^M\rangle$ is any model of set theory, and $\mathcal{M} \models \langle m, E\rangle$ is a first-order structure, then we may regard $\langle m, E\rangle$ as an actual first-order structure in V, with domain $\{\, a \in M \mid \mathcal{M} \models a \in m \,\}$ and relation $a\ E\ b \leftrightarrow \mathcal{M} \models a\ E\ b$. Thus, we extract the existential content of the structure from $\mathcal{M}$ to form an actual structure in V.

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

THEOREM 13 [8]. *Every model of ZFC has an element that is a model of ZFC. Specifically, if* $\langle M, \in^M \rangle \models \text{ZFC}$, *then there is an object* $\langle m, E \rangle$ *in* $\mathcal{M}$, *which when extracted as an actual structure in* V, *satisfies ZFC.*

Proof. Consider first the case that $\mathcal{M}$ is ω-nonstandard. By the reflection theorem, any particular finite fragment of ZFC is true in some $\langle V_\delta^M, \in \rangle^{\mathcal{M}}$, and so by overspill there must be some δ in $\mathcal{M}$ for which $\mathcal{M}$ believes $\langle V_\delta, \in \rangle^{\mathcal{M}}$ satisfies a nonstandard fragment of ZFC, and so in particular it will satisfy the actual ZFC, as desired. Meanwhile, if $\mathcal{M}$ is ω-standard, then it must satisfy Con(ZFC) and so it can build the Henkin model of ZFC. So in any case, there is a model of ZFC inside $\mathcal{M}$. □

What surprises some is that the argument succeeds even when $\mathcal{M} \models \neg \text{Con(ZFC)}$, for although $\mathcal{M}$ thinks there is no model of ZFC, nevertheless it has many actual models of ZFC, which it rejects because its ω-nonstandard nature causes it to have a false understanding of what ZFC is. In the ω-nonstandard case of Theorem 13, the model $\langle V_\delta, \in \rangle^{\mathcal{M}}$ that is produced is actually a model of ZFC, although $\mathcal{M}$ thinks it fails to satisfy some nonstandard part of ZFC. Perhaps one way of highlighting the issue is to point out that the theorem asserts that every model $\mathcal{M}$ of ZFC has an element $\langle m, E \rangle$, such that for every axiom σ of ZFC, the model $\mathcal{M}$ believes $\langle m, E \rangle \models \sigma$; this is different from $\mathcal{M}$ believing that $\langle m, E \rangle$ satisfies every axiom of ZFC, since when $\mathcal{M}$ is ω-nonstandard it has additional axioms not present in the metatheory.

§6. Conclusions. Let us now return to the motivating philosophical issue with which we began this article. In contention is whether the definiteness of arithmetic truth can be seen as a consequence of the definiteness of the underlying mathematical objects and structure, or to put it simply, whether definiteness-about-truth follows from definiteness-about-objects. Feferman implicitly supports this in his conclusion:

> In my view, the conception [of the bare order structure of the natural numbers $\mathbb{N}$, with its least element and the attendant operations of successor and predecessor] is *completely clear*, and thence *all arithmetical statements are definite*. [5, pp. 6–7] (emphasis original)

Donald Martin is more explicit in his article for the same conference series:

> The concept of the natural numbers is first-order complete: it determines truth values for all sentences of first-order arithmetic. That is, it implies each first-order sentence or its negation. (p. 3)
>
> What I am suggesting is that the real reason for confidence in first-order completeness is our confidence in the full determinateness of the concept of the natural numbers. [26, p. 13]

Let us give a name to the general principle underlying these remarks, in order to help focus the discussion:

> **Objects→Truth determinateness principle**. Necessarily, if the objects and underlying operations and relations of a mathematical structure are definite and determined, then the corresponding theory of truth for that structure also is definite and determined.

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

According to this principle, there can be no indeterminateness in our theory of arithmetic, if the mathematical structure in which that truth resides is definite and determined. No blurriness or fuzziness will sneak into our judgment of truth in a structure, if there is none in the underlying mathematical objects and structure. In short, the view is that whenever you know exactly what a structure is, then truth in that structure is fully determined.

Our main philosophical thesis is that we don't necessarily get determinate truth in this way, for free, from the determinateness of the underlying objects, but rather, determinate truth constitutes an additional higher-order commitment, which must be argued for and justified separately. In particular, we do not legitimately deduce the determinateness of arithmetic truth based solely on our determinate understanding of the numbers 0, 1, 2, 3, and so on and their arithmetic structure.

6.1. What is definiteness? What does it mean exactly to say that mathematical objects and mathematical structure are definite or that truth in a structure is definite or determined? We are inclined to place the burden of providing a full account of definiteness on those who are making the positive claim, that is, on our opponents who are supporting (if implicitly) the Objects→Truth determinateness principle.

Meanwhile, however, in order to support our criticism of that too-easy jump from objects to truth, let us mention a few general ideas that we would find reasonable for any such notion of definiteness on offer to exhibit. To begin, in our view the notions of definiteness and determinateness of objects, structure, and truth are inherently concerned with a comparison of the objects and structure in different models or universe concepts and with the evaluation of truth judgments about that structure in those various worlds. That is, definiteness at bottom is an absoluteness claim made about a structure with respect to a class of worlds or universe concepts, asserting that the relevant objects, structure, and truth judgments are invariant across those worlds.

In this light, we would find it to be a sufficient condition for the definiteness of objects in a mathematical structure amongst a collection of worlds for those objects and that structure to exist identically in those worlds—all the worlds in question agree on which objects there are in the given structure and furthermore they are in perfect agreement about the fundamental atomic features of that structure, such as the interpretations of relations and operations in a first-order language. In such a case of agreement, we hold, the objects and structure could be said to be definite across those various worlds.

Similarly, we would find it to be a necessary condition for the definiteness of a given truth assertion about such a definite structure across a collection of worlds that they all come to the same judgement about the truth value of that assertion. That is, if the truth assertion is definite with respect to a collection of worlds or universe concepts, then at the very least, the assertion must have the same truth value in all those worlds.

6.2. Criticism of Objects→Truth. In light of these considerations, let us now argue for our central thesis, which is that the definiteness of the theory of truth for a structure does not necessarily follow as a consequence of the definiteness of the structure in which that truth resides. In particular, we deny the Objects→Truth determinateness principle.

Even in the case of arithmetic truth and the standard model of arithmetic $\mathbb{N}$, we claim, it is a philosophical error to conclude that arithmetic truth is definite just on the basis that the natural numbers themselves and their structure $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle$ are definite. Definiteness of truth does not ride for free upon definiteness of objects, but rather one must say something more about it.

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

Our argument is a simple appeal to the mathematical results of the previous sections of this article, particularly Theorems 1 and 5 and their corollaries, which show how it can be that we have definiteness about the objects and structural relations of a certain structure across a collection of worlds, while there is no definiteness for the truths residing in that structure across those worlds.

Namely, in the context of Theorem 1, we have two set-theoretic worlds, M_1 and M_2, which agree completely on the objects and atomic structure of what they both agree is the standard model of arithmetic $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle^{M_1} = \langle \mathbb{N}, +, \cdot, 0, 1, < \rangle^{M_2}$, and yet they disagree on their judgments of arithmetic truth in this structure. They have a sentence σ in common, which they both agree makes a certain arithmetic assertion—they both agree on the Gödel code of the sentence and on the syntactic structure of this sentence—and yet M_1 thinks that σ is true in $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle$ while M_2 thinks that σ is not true there.

Because the two models M_1 and M_2 agree exactly on the objects and primitive structural relations of this structure, they fulfill the sufficient condition for the definiteness of objects for this structure. And indeed we could consider these models in the context of the larger collection of worlds that agree with them on the structure $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle$. This structure exists identically in all these worlds, and so we have definiteness of arithmetic objects across these worlds. But since these models do not all agree on their accounts of arithmetic truth, they violate the necessary requirement for the definiteness of truth. For this reason, the situation constitutes a counterexample to the Objects→Truth definiteness principle.

We should like furthermore to emphasize that, for all we know, the situation of this example could be the actual state of affairs for our own mathematical ontology in regard to our arithmetic and set-theoretic concepts. For all we know, after all, we might be amongst the people living happily in such a world as M_1, taking the set-theoretic ontology of that world as our one true set-theoretic background—there seems no way in principle that we could ever know this was not the case. But in this event, we would look upon $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle^{M_1}$ as the standard model of arithmetic, blissfully unaware that there is another possible world M_2 with exactly this same structure of arithmetic and yet having different arithmetic truths. For this reason, we would be in error to deduce the definiteness of our theory of arithmetic truth solely on the basis of the seeming definiteness of the mathematical structure in which that truth resides, as there would be other set-theoretic realms with exactly the same arithmetic structure and yet different arithmetic truths.

6.3. Definiteness of truth as a higher-order commitment. In the situation we have exhibited, how exactly does the indefiniteness about arithmetic truth arise? How is it possible that the two models agree on the natural-number structure and yet disagree on arithmetic truth? Let us explain a little more about it. In order for this indefiniteness to occur, we claim, the two models M_1 and M_2 must have different second-order arithmetic. In particular, they will have different incompatible arithmetic truth predicates. Neither model will be able to see the arithmetic truth predicates of the other because both of these predicates obey the Tarskian truth conditions and one may easily prove by induction that any two predicates satisfying the Tarskian conditions must agree universally. So neither model can have access to the arithmetic truth predicate of the other, and consequently, the models cannot agree on the power

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

set of the natural numbers; so they cannot agree on the real numbers. In this way, indefiniteness of truth requires indefiniteness in the ontology of second-order objects.

This is why we say that the commitment to a definite theory of truth for a structure is a higher-order ontological commitment, one going strictly beyond the definiteness of the underlying structure itself. The assertion that there is definite arithmetic truth is a claim about the definiteness of certain second-order objects, asserting that there is no indefiniteness for the second-order object that constitutes the truth predicate fulfilling the Tarski truth conditions. For the definiteness of arithmetic truth, it is not enough to express a definite nature for the natural numbers 0, 1, 2, and so on, but rather one seems to require a degree of definiteness for the real numbers as well.

6.4. The meaning of an assertion. Roman Kossak has objected to our argument on the grounds that the arithmetic assertion σ, which has different truth values in the two models of set theory M_1 and M_2, nevertheless has different *meanings* in these two structures. Although the two models agree on the natural number structure $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle^{M_1} = \langle \mathbb{N}, +, \cdot, 0, 1, < \rangle^{M_2}$, nevertheless, according to this objection, the meaning of the sentence σ has changed. The Proof of Theorem 1, after all, shows that M_1 and M_2 are isomorphic by an isomorphism π mapping a different sentence τ to σ, and so according to this objection, the sentence σ carries in M_2 the same meaning that sentence τ carried in M_1, rather than the meaning carried by σ itself in M_1.

Wherein does the meaning of a sentence lie? To our way of thinking, the meaning of a sentence σ is best construed as precisely what that sentence expresses. And what the sentence expresses is determined completely by its compositional syntactic structure—the parse tree—for it is this compositional hereditary syntactic structure of the sentence that provides the instructions for how to calculate its truth, and this at bottom is what constitutes its meaning. The meaning of a sentence is precisely what the sentence expresses by its form.

Tarski seems to agree that the meaning of a sentence is unambiguously determined by its syntactic form.

> restrict[ing] myself henceforth entirely to formalized languages. These can be roughly characterized as artificially constructed languages in which the sense of every expression is unambiguously determined by its form. Alfred Tarski [29, sec. 2]

According to this view, the meaning of an assertion is determined by what the assertion expresses, and thus is grounded ultimately in the exact nature of the parse tree of the assertion.

Since the models M_1 and M_2 agree on the arithmetic structure of the natural numbers, the sentence σ has an identical parse tree structure in these models. And so the meaning of σ in M_1 is precisely the same as the meaning of σ in M_2.

In particular, it is not the meaning of σ that changed in the move from M_1 to M_2, we argue, but rather merely the truth judgment of σ that has changed. There is of course an analogy between σ in M_2 and τ in M_1, an analogy that is made precise and robust by the fact that they are automorphic images of each other. In the end, we therefore find Kossak's point to be not an objection to the argument, but rather an explanation of how it could come to be that arithmetic truth can vary without the objects and structure of arithmetic changing.

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

Similarly, to argue in the context of Theorem 12 that δ refers in M_2 to something other than δ itself would seem to be a very hard task for Kossak. He replies to this by asking whether there is any meaning of δ at all other than that given by the semantics of the ambient universe? We believe that there is, since δ is a particular ordinal that both M_1 and M_2 have in common, and they agree on the cumulative hierarchy V_δ up to stage δ, and one may inquire about the properties of this common object in M_1 and M_2. Our point is that because M_1 and M_2 agree on V_δ, they will be forced to have a certain level of agreement of their theories of truth for this structure, but this agreement is not necessarily universal and ultimately M_1 and M_2 can disagree on whether certain statements are true in that common structure.

In connection with this objection, Kossak asked whether there is another method of producing examples where a given model of arithmetic has alternative truth predicates, and specifically, whether two models of set theory can have the same natural numbers, but non-isomorphic theories of arithmetic truth. That is, the question is whether one can have models for which

$$\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle^{M_1} = \langle \mathbb{N}, +, \cdot, 0, 1, < \rangle^{M_2}, \text{ but}$$

$$\langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{TA} \rangle^{M_1} \ncong \langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{TA} \rangle^{M_2}.$$

Jim Schmerl observed that indeed this situation occurs, arguing essentially as follows: the theory $T = \mathrm{TA} + \mathrm{Th}(\mathbb{N}, \mathrm{TA})^{\mathrm{ZFC}}$, in the language with a satisfaction predicate Tr, is computable from $0^{(\omega)}$, but the theory of $\langle \mathbb{N}, +, \cdot, 0, 1, <, \mathrm{TA} \rangle$ has degree $0^{(\omega+\omega)}$, and so T is not complete in the expanded language. It follows that there are computably saturated models $\langle \mathcal{N}, \mathrm{Tr}_1 \rangle$ and $\langle \mathcal{N}, \mathrm{Tr}_2 \rangle$ of T, which have the same reduct to the language of arithmetic, but which are not elementarily equivalent in the full language and hence also not isomorphic. But since these models are computably saturated models of $\mathrm{Th}(\mathbb{N}, \mathrm{TA})^{\mathrm{ZFC}}$, it follows by Proposition 4 that they are both ZFC-standard, arising as the standard model of arithmetic and arithmetic truth $\langle \mathbb{N}, \mathrm{TA} \rangle^{M_1}$ and $\langle \mathbb{N}, \mathrm{TA} \rangle^{M_2}$ inside models of ZFC, as desired. This kind of example therefore seems to address Kossak's objection on the meaning of σ, while still establishing our main point, for in this case there is no automorphism providing one with an alternative meaning for the sentence σ on which the models disagree.

6.5. A spectrum of views on pluralism. The question of definiteness of objects and truth is deeply connected with the issue of pluralism in mathematics and set theory, for pluralism itself is a kind of indefiniteness. Peter Koellner [18] has emphasized that there is a hierarchy of positions to take on pluralism in mathematics, depending on where one expects indeterminism to first arise in mathematics, if it does so at all.

> But, in fact, many people are non-pluralists with regard to certain branches of mathematics and pluralists with regard to others. For example, a very popular view in the foundations of mathematics embraces non-pluralism for first-order number theory while defending pluralism for the higher reaches of set theory. Most people would, for example, maintain that the Riemann Hypothesis (which is equivalent to a Π^0_1-statement of arithmetic) has a determinate truth-value and yet there are many who maintain that CH (a statement of set theory, indeed, of third-order arithmetic) does not have a determinate truth-value. Feferman is an example of someone who

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

> holds this position. So instead of two positions—pluralism versus non-pluralism—we really have a hierarchy of positions. [18]

Summarizing the spectrum, on one end we have what might be described as the hard-core set-theoretic Platonists, such as Isaacson [14] (and to a more qualified extent, Woodin and Maddy [23, 24]), who affirm the existence of the universe of all sets, in which set-theoretic claims such as CH have a definite truth value that we might come to know. Steel softens this position by allowing a restricted pluralism of mutually interpretable theories, such as occurs in the interaction of large cardinals, inner model theory and the theory of the axiom of determinacy. Martin [25] considers the case for determinism in set theory, but declining a full endorsement, argues instead only that there is at most one concept of set meeting his criteria, leaving explicitly open the possibility that in fact there is none. In the center of the spectrum we are considering, Feferman [4] is committed to the Platonic existence of the natural numbers, while remaining circumspect about the definiteness of higher-order set-theoretic objects and truth. He asserts that "the origin of Dedekind–Peano axioms is a clear intuitive concept," while the intuition for set theory "is a far cry from what leads one to accept the Dedekind–Peano axioms," and based on the dichotomy between the ontological status of the structure of natural numbers and of the structure of higher types, he concludes that "the Continuum Hypothesis is an inherently vague problem that no new axiom will settle in a convincingly definite way." Nik Weaver [30] similarly calls for classical logic in the realm of arithmetic, where he takes mathematical assertions to have a clear and distinct definite nature; but in the set-theoretic realm of sets of numbers or sets of sets of numbers, which he takes to have a less definite nature, we should be using intuitionistic logic.

Meanwhile, Hamkins [7, 9, 10, 12, 13] advances a more radical multiverse perspective that embraces pluralism even with respect to arithmetic truth. He criticizes the categoricity arguments for arithmetic definiteness—which are also used to justify higher-order definiteness, as in Isaacson [14]—as unsatisfactorily circular, in that they attempt to establish the definiteness of the natural number concept by appeal to second-order features involving a consideration of arbitrary subsets, which we would seem to have even less reason to think of as definite. More extreme positions, such as ultrafinitism, lie further along the spectrum.

6.6. The nonstandardness objection. It would be natural to object to our argument using M_1 and M_2 on the grounds that these models are ω-nonstandard, and furthermore, every instance of this kind of non-determinism will be with models whose natural numbers are non-standard. Such an argument, however, seems to us to beg the question, since the issue is whether definiteness of truth follows from definiteness of objects and structure, not whether it follows from definiteness of objects and structure and the knowledge that those objects are "standard" with respect to some meta-mathematical conception of standardness. One would want to know, for example, about whether that conception of standardness is itself definite, for it is easy enough to establish that the concept of whether a given model of arithmetic is standard or not is not absolute between all models of set theory.

The natural number structure we consider, after all, is precisely the standard model of arithmetic from the point of view both of M_1 and M_2, since it is $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle^{M_1}$, which is identical to $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle^{M_2}$. So the structure we consider is already "standard" from the point of view of the set-theoretic models in which we consider it.

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

It is only seen as nonstandard from a larger set-theoretic conception able to encompass both M_1 and M_2.

For our opponents to augment their argument for definite truth by making additional claims about the special nature of the natural numbers they consider, going beyond mere definiteness to a larger meta-theoretic notion of standardness, merely serves to verify our main point, which is that indeed something more needs to be said about it.

6.7. Conclusion. Ultimately, we argue that the definiteness of arithmetic truth does not flow for free from the definiteness of the numbers and the arithmetic structure $\langle \mathbb{N}, +, \cdot, 0, 1, < \rangle$, but rather amounts to a higher-order level of definiteness, for which one must argue separately. Even if two mathematicians might agree on the ontology of the numbers $0, 1, 2, \ldots$ and on their basic arithmetic structure, they may disagree on the judgments of arithmetic truth, if they do not agree on the second-order theory, as the models M_1 and M_2 in our main theorem show. A philosopher cannot legitimately claim to have a clear and definite conception of the underlying structure of arithmetic as a reason to take arithmetic truth also as definite.

Perhaps Feferman and the others will reply to the argument we have given in this article by saying that their conception of the natural numbers $0, 1, 2, \ldots$ and the natural number structure is so clear and distinct that they can use this overwhelming definiteness to see that arithmetic truth must also be definite. That is, perhaps they reply to our objection by saying that they did not base the definiteness of arithmetic truth solely on the fact that the natural number objects and structure are definite, but rather on something more, on the special nature of the definiteness of natural number objects and structure that they assert is manifest. For example, perhaps they would elaborate by explaining that arithmetic truth must be definite, because any indefiniteness would reveal a proper cut in the natural numbers: consider the natural numbers n for which Σ_n truth is definite and argue that this contains 0 and is closed under successors.

That would be fine. Our response to such a reply, after first getting in the obvious objection about the indefinite nature of definiteness and whether it can be used in a mathematical induction in this way, would be to say that yes, indeed, this kind of further discussion and justification is exactly what we call for. Our main point, after all, is that definiteness-about-truth does not follow for free from definiteness-about-objects and that one must say something more about it.

Acknowledgments. This work was begun in June 2013 and largely completed at that time. For inexplicable reasons, despite the journal acceptance modulo minor revisions only, the authors became distracted with other projects and this paper languished until 2025. We are now pleased finally to bring the paper to completion. The authors are both thankful to Roman Kossak for fruitful discussions, to Andrew Marks for proving Lemma 10.1, and to W. Hugh Woodin for suggesting a simplification in one of our arguments. Commentary concerning this article can be made at http://jdh.hamkins.org/satisfaction-is-not-absolute.

Funding statement. The first author is grateful for the support provided to him as a visitor to Fudan University in Shanghai in June 2013 and again in July 2025. The second author's research has been supported by the Major Project of the Key Research Institute of Humanities and Social Sciences under the Ministry of Education (Project No. 22JJD110002). He is thankful to the Institute for Mathematical Sciences at the

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

National University of Singapore for continuing support for him to attend the logic summer school every summer for four years, where the authors first met.

BIBLIOGRAPHY

[1] Chang, C. C., & Keisler, H. J. (1990). *Model Theory* (third edition), volume 73 of Studies in Logic and the Foundations of Mathematics, Amsterdam: North-Holland.

[2] Enayat, A. (2009). On ZF-standard models of PA, preliminary report, Mittag-Leffler Institute.

[3] Feferman, S. (1991). Reflecting on incompleteness. *Journal of Symbolic Logic*, **56**(1), 1–49.

[4] ———. (1999). Does mathematics need new axioms? *American Mathematical Monthly*, **106**(2), 99–111.

[5] ———. (2013). Comments for EFI workshop meet. In Koellner, P., editor. *Exploring the Frontiers of Incompleteness*. Cambridge, Massachusetts: Harvard, pp. 1–9.

[6] Gupta, A., & Belnap, N. (1993). *The Revision Theory of Truth*. Cambridge, Massachusetts: A Bradford Book, MIT Press.

[7] Gitman, V., & Hamkins, J. D. (2010). A natural model of the multiverse axioms. *Notre Dame Journal of Formal Logic*, **51**(4), 475–484.

[8] Halimi, B. (2009). Models and universes, manuscript, 19.

[9] Hamkins, J. D. (2011). The set-theoretic multiverse: A natural context for set theory. *Annals of the Japan Association for Philosophy of Science*, **19**, 37–55.

[10] ———. (2012). The set-theoretic multiverse. *Review of Symbolic Logic*, **5**, 416–449.

[11] ———. (2013). Is there a subset of the natural number plane, which doesn't know which of its slices are arithmetic? MathOverflow question (version: 2013-09-05). http://mathoverflow.net/q/141387

[12] ———. (2014). A multiverse perspective on the axiom of constructibility. In *Infinity and Truth*, volume 25 of Lecture Notes Series, Institute for Mathematical Sciences, National University of Singapore. Hackensack: World Scientific Publishing, pp. 25–45.

[13] ———. (2015). Is the dream solution of the continuum hypothesis attainable? *Notre Dame Journal of Formal Logic*, **56**(1), 135–145.

[14] Isaacson, D. (2011). The reality of mathematics and the case of set theory. In Novak, Z., and Simonyi, A., editors. *Truth, Reference, and Realism*. Budapest: Central European University Press, pp. 1–76.

[15] Kaye, R. (1991). *Models of Peano Arithmetic*, volume 15 of Oxford Logic Guides. Oxford: The Clarendon Press, Oxford University Press, and Oxford Science Publications.

[16] Kossak, R., & Kotlarski, H. (1988). Results on automorphisms of recursively saturated models of PA. *Fundamenta Mathematicae*, **129**(1), 9–15.

[17] ———. (1992). Game approximations of satisfaction classes and the problem of rather classless models. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, **38**, 21–26.

[18] Koellner, P. (2013). Hamkins on the multiverse. In Koellner, P., editor. *Exploring the Frontiers of Incompleteness*. Cambridge, Massachusetts: Harvard, pp. 1–27.

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press

[19] Krajewski, S. (1974). Mutually inconsistent satisfaction classes. *Bulletin de l'Académie Polonaise des Sciences, Série des Sciences Mathématiques, Astronomiques et Physiques*, **22**, 983–987.

[20] ———. (1976). Non-standard satisfaction classes. In Marek, W., Srebrny, M., and Zarach, A., editors. *Set Theory and Hierarchy Theory A Memorial Tribute to Andrzej Mostowski*, volume 537 of Lecture Notes in Mathematics. Berlin: Springer, pp. 121–144.

[21] Kripke, S. (1975). Outline of a theory of truth. *The Journal of Philosophy*, **72**(19), 690–716.

[22] Kossak, R., & Schmerl, J. H. (2006). *The Structure of Models of Peano Arithmetic*, volume 50 of Oxford Logic Guides. Oxford: The Clarendon Press and Oxford University Press, Oxford Science Publications.

[23] Maddy, P. (1988). Believing the axioms, I. *The Journal of Symbolic Logic*, **53**(2), 481–511.

[24] ———. (1988). Believing the axioms, II. *The Journal of Symbolic Logic*, **53**(3): 736–764.

[25] Martin, D. A. (2001). Multiple universes of sets and indeterminate truth values. *Topoi*, **20**(1), 5–16.

[26] ———. (2012). Completeness or incompleteness of basic mathematical concepts. In Koellner, P., editor. *Exploring the Frontiers of Incompleteness*. Cambridge, Massachusetts: Harvard, pp. 1–26, draft 3/15/12. Available at https://www.math.ucla.edu/~dam/booketc/efi.pdf.

[27] Schlipf, J. S. (1978). Toward model theory through recursive saturation. *The Journal of Symbolic Logic*, **43**(2), 183–206.

[28] Smoryński, C. (1981). Recursively saturated nonstandard models of arithmetic. *The Journal of Symbolic Logic*, **46**(2), 259–286.

[29] Tarski, A. (1983) [1935]. The concept of truth in formalized languages. In Corcoran, J., editor. *Logic, Semantics, Metamatheamtics* (second edition). Indianapolis: Hackett, pp. 152–278.

[30] Weaver, N. (2005). Predicativity beyond Γ_0. *Mathematics* Preprint, arXiv:math/0509244.

DEPARTMENT OF PHILOSOPHY
UNIVERSITY OF NOTRE DAME
100 MALLOY HALL
NOTRE DAME
IN 46556
USA
E-mail: jdhamkins@nd.edu
URL: https://jdh.hamkins.org

SCHOOL OF PHILOSOPHY
FUDAN UNIVERSITY
220 HANDAN ROAD
SHANGHAI, 200433
CHINA
E-mail: yangruizhi@fudan.edu.cn

https://doi.org/10.1017/S1755020325100907 Published online by Cambridge University Press