Σ^0_1 SOUNDNESS ISN'T ENOUGH

SHARON BERRY

Abstract. It's sometimes suggested that we can (in a sense) settle the truth-value of some statements in the language of number theory by stipulation, adopting either ϕ or $\neg\phi$ as an additional axiom. For example, in [4] and a series of recent APA Presentations, Clarke-Doane suggests that any Σ^0_1 sound expansion of our current arithmetical practice would express a truth. In this paper I'll argue that (given a certain popular assumption about the model-theoretic representability of languages like ours) we can't know ourselves to have any such freedom.

1. Introduction

It's sometimes suggested that we're (in a sense) free to stipulate either answer to certain undecidable questions in the language of number theory[1]. For example, in [4] and a series of recent APA presentations, Clarke-Doane suggests that mathematicians are free to adopt any Σ^0_1 sound variant on (or expansion of) our current mathematical axioms.

Call the claim that there's *some* number-theoretic sentence ϕ, such that we could safely add ϕ or $\neg\phi$ to our current number theoretic practice, the Number-Theoretic Indeterminacy (NTI) thesis. Claiming that many different such variants on our actual number theoretic practice are acceptable can be appealing because it promises to help solve access worries. For, crudely put, the wider the range of acceptable choices of axioms of number theory we might have adopted, the less puzzling it is that our axioms are acceptable.

In this paper, I'll present what I take to be a serious consideration against the Number-Theoretic Indeterminacy thesis (and every larger philosophy of mathematics implying it). I'll argue that, accepting NTI theory commits one to either

.
[1]Here, by 'undecidable questions' I mean questions about the truth of a sentence in the language of number theory which we can neither prove nor refute using our current axioms and inference rules.

making an intuitively unjustified assumption about the physical world or rejecting a certain popular assumption about the model theoretic representability of natural languages. Thus, philosophers (who accept the relevant assumption) shouldn't take themselves to know NTI and can't use it to help answer access worries. .

In §2 I will clarify the number theoretic indeterminacy thesis above, of which the theory that all Σ^0_1 sound expansions of our current number theory are on par is a particularly attractive instance. In §3 I will explain the popular assumption about model theoretic representability referenced above. In §4 I'll make my main argument and in §5 I'll make some reflections on it.

2. Number Theoretic Indeterminacy

2.1. **The NTI Thesis.** So let me begin by introducing the NTI thesis I mean to argue against, and explaining why it is appealing.

> **Number-Theoretic Indeterminacy Thesis**: For some number-theoretic sentence ϕ, we could equally well add either ϕ or $\neg\phi$ to our current number-theoretic axioms yielding a total mathematical practice which is 'on par' with our current one.

Note that the Number-Theoretic Indeterminacy Thesis asserts that we can add either ϕ or $\neg\phi$ to our current *overall* mathematical practice, which I take to include a priori acceptance of both first order logical pure mathematical principles (like the Peano Axioms) and some applied mathematical principles like the following:

- All instances of the first order induction schema in our current language including ones that involve non-mathematical vocabulary.[2]
- Certain quasi-analytic principles about counting.

This attention to applied mathematics is important because everyone agrees that all consistent extensions of first order Peano Arithmetic correspond to legitimate topics of mathematical investigation (many such axiom systems are studied in nonstandard number theory). The controversial claim in question is that one can

[2] One such instance of the induction schema is 'If Jake likes 0, and for each number n, if Jake likes n then Jake likes n+1 then Jake likes all the numbers.'

harmoniously add whichever one of ϕ and $\neg\phi$ you prefer to the whole of our current mathematical practice. We can only hope to use Number Theoretic Indeterminacy to reduce access worries if we can show that variants on our actual mathematical reasoning methods - which include a priori acceptance of the impure mathematical principles above- would have been equally truth good/truth conducive.[3]

I take the claim that adding either ϕ or $\neg\phi$ as an axiom would yield a mathematical practice on par with our current one to imply at least that if we someone made one of these choices then: the axiom they've added would express a truth, while all our current mathematical axioms and the pure number theoretic principles mentioned above would continue to express truths.

2.2. **Two Types of NTI Theories.** Now why accept the above number theoretic indeterminacy thesis? As noted above, accepting NTI seems to help with access worries by allowing that a wider range of variant mathematical theories would express truths. Perhaps it can also be motivated by Putnamian model theoretic arguments suggesting that we can't pin down a unique intended natural number structure[8]. However it will be helpful to distinguish two kinds of number theoretic indeterminist views.

One type of number theoretic indeterminist view says mathematical truth shouldn't outrun proof, and so takes *every* number theoretic sentence ϕ which can't be proved or refuted[4] to be indeterminate in a way that implies the mathematical practices which stipulate either its truth or falsehood are on par with each other (and our own mathematical practice). It holds that all *syntactically consistent* variants on

[3]To put this point another way, perhaps one might say that both ϕ and $\neg\phi$ were, somehow, pure mathematically on par (e.g., on par qua pure mathematical first order theories), despite the fact that (at most) one of them had sufficient harmony which our counting practices for the counting rules above to continue expressing truths. But in this case merely noting this kind of limited parity wouldn't do anything to reduce access worries. They would simply shift access worries about how we manage to express pure mathematical truths to access worries about how we get pure mathematical concepts which have the right kind of harmony with our counting procedures (and other apparently necessary a priori applications of mathematics.

[4]That is, every claim which can't be proved or refuted from our current axioms

our current number theoretic practice would express truths[5]. I will, following [6], call this view Extreme Anti-Objectivism.

Famously, Extreme Anti-Objectivism faces a worry involving the apparent connection between number theory and facts about consistency[6]. For, the above proposal appeals to the idea that there are definite facts about consistency. But consistent variants on our current arithmetical practice can (seemingly) take different views about provabilty and consistency through implying different arithmetical Con() sentences.[7] And it seems odd to be a realist about consistency while saying that variants on our mathematical practice which, seemingly, get consistency facts wrong (e.g., claim that a consistent theory is actually inconsistent) would be on par with our current practice.

Another type of view implying NTI is what I'll call 'moderate anti-objectivism'. Philosophers of this stripe avoid the problem above by saying certain incompatible ways of extending our current number-theoretic practice are legitimate, but only ones *which don't get consistency wrong* (in the sense above). The simplest example of such a view is Justin Clarke-Doane's suggestion (in [4] and a series of APA presentations) that all (and only) Σ^0_1 sound extensions of number theory (those which don't prove any false sentences in the language of arithmetic with only a single unbounded existential quantifier) are acceptable. This proposal amounts to

[5]Henceforth by 'consistent' I will mean syntactically consistent, rather than any richer notion of coherence like that invoked in [2], or having a model which treats second order quantifiers standardly.

[6]For example see Hartry Field's discussion of 'extreme mathematical anti-objectivism' in [6]

[7]By Gödel, there are consistent extensions of Peano Arithmetic (or any extension thereof) which disagree on the truth of $\mathrm{CON}(T+\phi)$ where T is our current number-theoretic practice and ϕ is a sentence in the language of number theory. Specifically, we're inclined to accept all instances of the schema 'Con(S) iff S syntactically consistent' for a suitably concrete descriptions of some axiom system S, so it seems good interpretations of number theory should make this claim true.

saying that *all and only* expansions of our number-theoretic practice which don't get consistency facts wrong are acceptable and on par[8].[9]

Mere Σ^0_1 soundness and its ilk can seem to provide a principled and attractive middle ground between realism and antirealism concerning the numbers. On one hand, we avoid the problem for extreme antiobjectivists above by saying a that all claims of the form "no natural number codes a proof of '0=1' in formal system S" have definite right answers; (just as one would hope) they're determiantely true if the the formal system in question is consistent and determinately false otherwise. On the other hand, one reduces the strength of access worries (compared to more truthvalue realist views) by saying getting mathematics right doesn't require recognizing facts about the coherence of theories logically powerful enough to pin down right answers to all questions in number theory (like Second Order Peano Arithmetic) or truth in some favored platonic model of the natural numbers[10].[11]

However, I'll argue that, a certain popular assumption about model theoretic representability is difficult to reconcile with the Number Theoretic Indeterminacy thesis (and hence with even modest antiobjectivist views like the mere Σ^0_1 soundness thesis).

[8]Specifically it holds that all (and only) Σ^0_1 sound extensions of number theory (those which don't let one prove any false sentences in the language of arithmetic with only a single unbounded existential quantifier) are acceptable.For Σ^0_1 sound extensions of number theory can't get consistency facts wrong, as the claim that some (computable) theory is inconsistent (and all theories we could accept are computable) is a Σ^0_1 claim and no consistent theory (extending Peano arithmetic) can falsely prove an inconsistent theory is consistent. And an inconsistent theory T, by definition, has some finite proof of contradiction which ensures that any theory extending Peano arithmetic proves $\neg\,\mathrm{con}(T)$

[9]I will drop this caveat going forward with the understanding that all theories mentioned in this paper can be presumed to extend Peano arithmetic.

[10]For example, if you take our determinate reference to the natural numbers to be secured by employing full second-order quantification, the same resources will also let us pin down the width of the hierarchy of sets and hence the truth-value of CH.

[11]A further appealing feature of this weakening is that any theory which avoids proving false Con sentences will automatically avoid the kind of counterargument by imaginable physical experiment involving infinitary computers discussed in [2]. For these hypothetical experiments with (simple, less controversially imaginable) hypercomputers showed the falsehood of some Σ^0_1 sentence independent of our current arithmetic.

3. Model-Theoretic Representability Expectations

I take it that many philosophers accept something like the following idea about the philosophical illumination provided by set theoretic models and Tarski's definition of truth.

We can illuminatingly think of truth values for our utterances in English and other natural languages as being built up from ground level facts about the references of names and extensions of predicates etc. in accordance with Tarski's recursive definition of truth. Thus (if we temporarily bracket certain issues about the size of the universe etc. discussed in appendix B) we'd expect there to be a set theoretic model which assigns references and extensions to terms in our language correctly in such a way that applying the Tarskian recursive definition yields correct truth values for all our sentences.

Furthermore, something similar goes for speakers who mildly change the meaning of certain terms by adopting an additional mathematical axiom. Specifically, adding such axioms can't and won't change the overall Tarskian structure of their/our language. It also plausibly won't change the meaning of certain concrete/physical vocabulary like 'cat' and 'bites'. Rather it can (at most) lead to a person starting to about a different coherent mathematical structure[12].

Thus, (continuing to ignore concerns about size etc.) there should also be a set theoretic model for their language such that

- This model interprets physical vocabulary like 'cat' and 'bites' standardly.
- Applying Tarskian definitions to this model yields correct truth-values for the all speaker's utterances.

Note that an interpretation which attractively models a person's language in this sense needn't make all their beliefs come out true. For example, a number theoretic

[12]I mean to allow that adding some such new axioms regarding a certain mathematical structure could also make certain portions of one's language be undefined, e.g., if you think that we can express full second order quantification then maybe expanding your axioms of number theory in a syntactically consistent but second order incoherent way would have this effect.

determinist like myself (i.e., someone who believes all arithmetic sentences have determinate truth-values), might well think that (in most cases) if someone attempts to adopt some independent sentence ϕ which currently expresses a falsehood as an additional axiom of number theory, they'll wind up still talking about the (essentially) same natural numbers structure that we do[13], but holding a certain false belief about them. And if this picture of how language works is correct, then the model theoretic interpretations for the unlucky stipulator's language invoked above must reflect that fact, by making the sentence ϕ come out false.[14]

Some have tried to motivate this model theoretic interpretability assumption by a kind of inference to the best explanation. It has been suggested that Tarksi's theory provides the best (or only known decent) candidate for an explanation for the finite learnability of human languages[5] or the complex pattern of truth conditions English sentences take on a general theory of truth[15]

---

[13]One might take this to be true because this speaker's firmer acceptance of something like second order Peano Arithmetic anchors them to this and/or something about reference magnetism.

[14]Note that nothing about the above model theoretic representability claim forbids acceptance of new mathematical axioms from expressing truths by getting one to speak a more ontologically profligate language than our current one (as Quantifier Variantists have long suggested adopting new mathematical axioms can do[3]). For - barring the limitations of size noted above- we can always interpret talk about objects we don't recognize as referring to sets, i.e., if I believe only in fundamental particles and you also believe there are dogs I can consider a model which interprets your talk of dogs to referring to a set.

However, an important consequence of this technique for avoiding begging the question against quantifier variance theorists is that the acceptable set theoretic interpretations needn't be intuitively acceptable interpretations. We don't want to say that someone one speaking a more ontologically profligate language really refers to the sets when he talks about these objects. Rather we expect that someone who stipulatively started talking in terms of such extra objects would not change the large scale Tarskian structure of how the meanings/truth-values for complex sentences in their language are built up from more basic facts, and that we should be able to provide a set theoretic model which reflects these facts (about a set theory free fragment of their language) by associating their terms with some sets and relations between sets that have the structure which these new objects are supposed to form in the actual world.

[15]In [1] Benacerraf writes, "any theory of mathematical truth [should] be in conformity with a general theory of truth-a theory of truth theories ...which certifies that the property of sentences that the account calls 'truth' is indeed truth. This... can be done only on the basis of some general theory for at least the language as a whole (I assume that we skirt paradoxes in some suitable fashion)... the semantical apparatus of mathematics [should] be seen as part and parcel of that of the natural language in which it is done, and thus whatever semantical account we are inclined to give of names or, more generally, of singular terms, predicates, and quantifiers in the mother tongue [should] include those parts of the mother tongue which we classify as mathematics.

I take it that we have only one such account: Tarski's, and that its essential feature is to define truth in terms of reference (or satisfaction) on the basis of a particular kind of syntactico-semantical analysis of the language, and thus that any putative analysis of mathematical truth must be an analysis of a concept which is a truth concept at least in Tarski's sense. "

Now let's turn the problem for NTI theorists.

4. Main Argument

If number-theoretic indeterminacy is true, two speakers Alice and Bob can make syntactically incompatible but acceptable choices for which axioms of number theory to add. In particular, the NTI theorist thinks the following is possible.

> **Acceptable Divergent Axiom Choice** We start by accepting Peano Arithmetic together with certain conceptually core principles about applied mathematics[16]. Alice and Bob do the same. But, for some number theoretic sentence ϕ independent of our current number theoretic practice, Alice adds[17] ϕ to her theory of arithmetic and Bob adds $\neg\phi$ to his theory. Afterwards all the sentences Alice now takes to be axioms or conceptually central truths about the numbers (as per §2) express truths when spoken by her. The same goes for Bob.

Note that, by the completeness theorem there are (i.e., our current set theory includes) models of the 'natural numbers' which make Alice's axioms come out true and other models which make Bob's axioms come out true. However, at least one of those models must interpret Alice's term 'number' non-standardly, adding 'infinite' natural numbers which are larger than all finite natural numbers (while being cleverly structured to keep all instances of the first-order induction schema in the language of number true). Thus, at least one of Alice and Bob can only be model theoretically interpreted (from the point of view of our set theory) by taking them to be talking about something which isn't isomorphic to the natural numbers.

Without loss of generality, assume that Alice is the one whose pure mathematical axioms require us to interpret her non-standardly. This fact alone doesn't

[16]Note that the standard induction schema in Peano Arithmetic automatically lets us perform induction on any predicate definable in our language, even predicates which make use of non-number-theoretic vocabulary.
[17]This talk of adding is merely a verbal flourish. It is enough to imagine Alice simply accepting our practice as well as the truth of ϕ.

prevent Alice from being model theoretically representable. For recall that the model theoretic representability assumption doesn't require that we interpret her mathematical expressions like 'number' as referring to the same objects as ours does. merely requires interpreting physical vocabulary like 'Emperor' and 'ruled after' standardly.

However, I'll argue that if the epistemically possible physical scenario below holds then Alice cannot be attractively model theoretically represented (in the sense defined in 3). That is, working within our current set theory we can prove the following conditional, 'If there are an ω sequence of emperors then there are not attractive model theoretic representations of both Alice and Bob'. Note that I'm showing there can't even be *separate* models that attractively interpret both Alice and Bob's speech; I'm not assuming that Alice and Bob can refer to each others' number concepts.[18]

Now I will introduce the epistemic possibility which makes it difficult to attractively model theoretically represent Alice.

> **ω-sequence of Emperors**: The Roman Empire will be reestablished, and time will turn out to be infinite and the laws of entropy an illusion etc. so that Virgil's dream of imperium sine fine literally comes true. Specifically there is (what our current set theory with ur-elements classifies as) an ω sequence of emperors.

To see why Alice can't be attractively interpreted in this case, first note that if we interpret 'is the nth emperor' standardly[19] then we're forced to interpret Alice's '... is the ...th emperor' as relating the emperors to exactly the standard initial segment of this nonstandard model. For it seems this relation should pair Augustus with the first of Alice's 'numbers', Tiberius with the least 'number' after that, Caligula

[18] Thus I won't make the assumption that Alice can somehow refer to Bob's 'numbers' and use them to define an initial ω sequence and thus contradict the induction axiom as per [7]. For the NTI theorist isn't committed to the possibility of a unified language which (so to speak) quantifies over both Alice's and Bob's numbers (and makes each of their axioms come out true about their own copy of the natural numbers).

[19] That is if we identify the initial segment of Alice's 'numbers' with our numbers then the extension of the relation 'is the n-th emperor' would have the same extension for Alice as it does for us.

with the least after that etc. until we have mapped the emperors to the standard initial segment of the Alice's 'numbers', at which point we have no emperors left to pair with the remaining 'nonstandard numbers' in our nonstandard model. In this way we map onto all and only the standard initial segment of Alice's numbers (Note that an attractive interpretation can't change the extension of emperor). But then we must interpret the property $Q(n)$ "there is an n-th emperor" as applying to 1 and the successor of every 'number' it applies to, but not to all numbers. So we must interpret the following claim as coming out false.[20]

> **Informal Emperor Induction**: If there's a first Roman Emperor and, for every number n, if there's an n-th emperor then there's an n+1st emperor then for all numbers n there's an nth emperor.

However, on our description of of the NTI's Acceptable Divergent Axiom Choice scenario this instance of the induction schema expresses a truth in Alice's language. Thus, there can be no model theoretic interpretation which gets Alice's truth conditions right

OK but what if we don't interpret the two place relation'is the n-th emperor' standardly? Insofar as this relation is a somewhat mathematical one, it seems reasonable that adopting a new axiom of number theory could change the meaning of this mathematical sounding relation. However, allowing this kind of wiggle room turns out not to help. For it turns out that merely ensuring the truth of certain principles which seem to be (or follow from) conceptually core truths about counting and the numbers forces us to interpret 'is the nth emperor' standardly in the sense above, i.e., as relating the emperors to exactly the standard initial segment of (our model for) Alice's 'numbers'. I lay out this argument in detail in appendix A.

So the argument above goes through just the same. Thus if there are an ω sequence of emperors, any attempt to attractively model theoretically represent Alice must fail.

---

[20]The last two points follow from the clause concerning Tarskian recursive definitions of truth in our definition of model theoretic representability

So it follows from the model-theoretic interpretability assumption (and the fact that we can't know there won't be an ω sequence of Roman Emperors) that Number Theoretic Indeterminism is an unattractive view. More specifically, anyone who takes themselves to know the truth of the NTI thesis is committed to the possibility of superficially similar but deeply model-theoretically unrepresentable languages.

5. Reflection

One can think of what is going on here as follows. Initially, if seems like we can reconcile NTI with model theoretic intepretability by simply modeling speakers who stipulate different consistent extensions of the Peano Axioms as speaking about nonstandard models of the natural numbers. The fact that some of these models must not be isomorphic to our current model of the numbers seems like a small cost (if any). For it's plausible that adding new axioms of number theory could change the kind of structure your number talk would have to be about.

However it's implausible that adding new axioms of number theory could change how terms like 'emperor' and 'ruled before' would apply. And this creates a problem for interpreting Alice using a nonstandard model as follows.

Although these nonstandard models are cooked up to satisfy all our our first order logical pure mathematical axioms, including induction for purely mathematical predicates, they don't satisfy real second order induction. And appealing to physical stuff can let us reveal that by picking out portions of mathematical structures we couldn't pick out using pure mathematical language alone [21]. There are some ways which we consider it epistemically possible for 'emperor' and 'successor' to apply which would let her define subsets she couldn't with pure mathematical vocabulary alone. More specifically, if our numbers are an initial segment of hers and the emperors have the structure of our numbers, her predicate 'There is an emperor

[21]For another example of how the physical world can let us secure reference we might not otherwise be able to, consider reference to physical constants, i.e., constants in the laws of physical nature, could let us refer to real numbers which we can't provide definite descriptions for in our countable pure mathematical language.

who is the n-th emperor' applies to a successor closed initial segment of Alice's natural numbers –hence induction fails for this predicate.

6. Conclusion

In this paper, I've argued that accepting any form of number theoretic indeterminacy has grave costs — or at least an important consequence. This antirealism about the numbers turns out to be deeply in tension with ideas about about the model representability of natural languages that many philosophers find attractive.

For, working within our current language and set theory, we can prove the following conditional, 'If there is an ω sequence of emperors then there are not attractive model theoretic representations of both Alice and Bob (people who have successfully stipulated opposite truth-values for some number theoretic sentence ϕ as per NTI)'. Thus any philosopher[22] who accepts both NTI and MTI is committed to the apparently unjustified claim that 'There aren't an ω sequence of emperors'.

Let me end with some quick caveats. The argument in 4 isn't intended to suggest (or assume) that anyone could recognize that there's an ω sequence of emperors[23] if there happens to be one. Nor does it suggest any way of learning *which* one of two people who've adopted syntactically inconsistent number theoretic axioms got things right. It also doesn't (in its own) suffice to answer Putnamian model theoretic worries. For we don't know whether there's an ω sequence of emperors (or the like). So we don't know that any physical objects form the kind of short linear orders needed to block nonstandard interpretations of our first order number theoretic axioms.[24]

[22]Strictly speaking this only applies to those who accept standard ZFC axioms of set theory, as true of our current number concept

[23]It also doesn't imply that they could determinately refer to the is structure which the emperors under successor happen to have, in any other way than as 'whatever structure the total pluarality Roman emperors happen to form when considered under the relation 'ruled before'.

[24]However, at the end of REDACTED I suggest that combining Davidsonian theses about charitable interpretation with (an earlier version of) the argument above might do the trick.

Appendix A. Details of Non-Modelability

To articulate my argument that one can't attractively model-theoretically represent Alice (given the assumptions about Alice and the Roman Emperors made in S4) more rigorously, I'll need to formalize a relevant fragment of Alice's language and counting talk.

Consider our use of the natural numbers to talk about temporally ordered objects and events such as 'the 4th U.S. President', 'the 37th successful rickrolling', or 'the 3rd Roman emperor'. Now I take all the principles below to be (or be derivable from) conceptually central truths - which any acceptable interpretation of Alice must make true - regarding counting (temporal) sequences of events using the natural numbers.

COUNTING RULES[25]

- An object x is the 0th emperor, i.e., $countemp(0, x)$ iff x is a emperor and all other emperors happen after x[26].
- If x is the nth emperor, then y is the $n + 1$th emperor iff y occurs after x and no other emperor occurs between x and y [27].
- Only emperors can be the nth emperor[28].
- No two distinct numbers correspond to the same emperor [29].

One can formulate these principles in first order logic as indicated in the footnote above (where $emperor(x)$ denotes 'x is a Roman emperor', $countemp(n, x)$ denotes 'x is the n-th emperor' and $before(x, y)$ denotes that the emperor x ruled before

[25] See REDACTED (under review) for a related argument that shows (using an analog of the rules below) that if Putnam's model-theoretic challenge can be resisted for certain physical vocabulary, then it can be resisted for the natural numbers as well.

[26] That is, $(\forall x)[countemp(0, x) \leftrightarrow emperor(x) \wedge (\forall y)(countemp(y) \to before(x, y) \vee x = y))]$. Here I am departing from common practice in counting up from 0 rather than 1, for the sake of simplicity. Also by writing this formula with 0, I abbreviate the corresponding claim about the unique object satisfying some definite description of 0 in terms of $\mathbb{N}.S$, e.g., 'the unique object that has a successor but isn't a successor.' .

[27] That is,

$$(\forall n)(\forall x)(\forall y)(\mathbb{N}(n) \wedge countemp(n, x) \to [(countemp(S(n), y) \leftrightarrow \\ emperor(y) \wedge before(x, y) \wedge (\forall z)\neg(emperor(z) \wedge before(x, z) \wedge before(z, y))]$$

[28] That is $(\forall x)(\exists n)(countemp(n, x) \to emperor(x))$.

[29] That is $(\forall n)(\forall m)[emperor(n, x) \wedge emperor(m, x) \to m = n]$

emperor y). And we can use the same vocabulary to formalize the predicate $Q(n)$ "there is an n-th emperor" above $Q(n) \overset{\text{def}}{=} (\exists x)(countemp(n, x))$. So the instance of the induction schema above becomes the following.

So our Alice will also accept the following formal sentence as a conceptually central mathematical truth.

EMPEROR INDUCTION $Q(0) \wedge (\forall n)\,[Q(n) \rightarrow Q(S(n))] \rightarrow (\forall n)[\mathbb{N}(n) \rightarrow Q(n)]$

Now to see why we can't simultaneously interpret Alice's number talk as referring to a nonstandard model and her terms emperor() and ruled before() as referring standardly (in the scenario under consideration) while making both her sentences COUNTING RULES and EMPEROR INDUCTION come out true, consider our predicament when choosing an extension for her relation '*countemp*()' (i.e., '..is the ...th Roman Emperor').

We are forced to interpret '*countemp*()' as relating the emperors to exactly the standard initial segment of Alice's 'numbers' for the following reason. We can only satisfy the first conjunct of (while interpreting 'emperor()' and 'ruledbefore()' standardly), if we interpret '*countemp*()' as relating 0 to the temporally first Roman Emperor (Augustus). And, given that we've done this, we can only satisfy the second conjunct of COUNTING RULES if we interpret '*countemp*()' as relating 1 to the second Roman emperor (Tiberius), and so on for all the objects in the standard initial segment of the nonstandard model. But this 'uses up' all the Roman emperors (and recall that we're taking Alice's term 'emperor()' to apply to all and only the Roman emperors). Since, by the last conjunct of COUNTING RULES, no emperor can be counted again (i.e., associated with another putative natural 'number') it follows that all and only the 'true' (i.e., standard) natural numbers are paired with emperors.

So we must make '$Q(0)$' and 'Whenever Q applies to some number n it also applies to $Q(n+1)$' true. But as we takes 'the numbers' to refer to a larger structure than this standard initial segment, our interpretation also makes '$\forall n Q(n)$' false. So our interpretation renders EMPEROR INDUCT false.

Appendix B. Weakening MTI

Now the NTI theorist might respond to the argument above by noting that there are other reasons why we might expect the model theoretic interpretability assumption stated above is somewhat too strong. There are certain intuitively harmless reasons we'd expect model theoretic interpretability to fail. In this section I'll discuss these caveats and note that weakening this assumption to recognize caveats makes no difference to my argument above, and hence can't help the NTI theorist.

First, (as mentioned in the Benacerraf quote above) there are concerns about representing truth conditions in a language containing a general truth predicates. One might also worry about fitting claims involving primitive physical or metaphysical possibility operators into the picture above. I will handle this problem by restricting our attractive modelability requirement to the collection of sentences in Alice's/Bob's languages which (is quantifier restricted to and) talks only about numbers and non-modal properties of physical objects.

Second, there can be limitation of size problems when set theoretically modeling things. If your language quantifies over sets then you can't have a model-theoretic interpretation which gets the extension of sets right (even up to isomorphism) because that would have to be proper class sized. Perhaps we should allow the modelability requirement to fail when limitations of size are all that prevent modeling. However, weakening the model theroetic interpertability assumption in this way does nothing to help the NTI theorist. For our problem was that all set theoretic interpretations made Alice's 'number' structure too large, not that our models of her language didn't have a large enough domain of objects to work with.[30]

Third one might have a blameless failure of model theoretic representability because a model theoretic representation (of the kind we've been considering) won't

[30]Here's another way to put that point. Temporarily adopting an expanded language which adds some layers of proper classes atop our hierarchy of sets helps with the problem above, but wouldn't block the argument above.

ever assign a sentence an indeterminate truth value. And surely it's extremely natural that Number Theoretic Indeterminacy theorists to say that many mathematical sentences are neither determinedly true or false.

So it's important not to beg the question against this kind of position. However we can avoid doing this by simply weakening our model theoretic interpretability assumptions slightly. For note that many popular ways of thinking about vagueness and other indeterminacy still take our language to have a broadly Tarskian structure, and hence still generate serious model theoretic representability assumptions. Consider supervaluationist theories, on which we get indeterminate sentences because there are a range of different acceptable ways of sharpening/specifying the extension of a predicate 'bald' or 'bank', and the determinately true sentences are those which come out true when we apply standard Taraskian conditions to all acceptable interpretations. If we think about indeterminacy in something like this way (as is popular to do) then we're still committed to some/all (mundane physical reference honoring) model theoretic interpretations which make all sentences determinately true in the speakers language come out true. And this turns out to be all my argument needs.

So relaxing this requirement to allow for vagueness in this way makes no difference to my argument.

References

[1] Paul Benacerraf. Mathematical Truth. *Journal of Philosophy*, 70:661–80, 1973.

[2] Sharon Berry. Malament-Hogarth Machines and Tait's Axiomatic Conception of Mathematics. *Erkenntnis*, 79(4):893–907, 2014.

[3] Rudolf Carnap. Empiricism, Semantics, and Ontology. *Revue Internationale de Philosophie*, 4(11):20–40, 1950.

[4] Justin Clarke-Doane. *Morality and Mathematics*. Oxford University Press.

[5] Donald Davidson. Truth and Meaning. *Synthese*, 17(1):304–323, 1967. bibtex:davidson1967.

[6] Hartry Field. *Realism, Mathematics & Modality*, volume 16. Basil Blackwell, 1988.

[7] Vann McGee. How We Learn Mathematical Language. *Philosophical Review*, 106(1):35–68, 1997.

[8] Hillary Putnam. Models and Reality. In *Realism and Reason*. Cambridge University Press, 1983.

Email address: `seberry@invariant.org`