212–228
https://doi.org/10.1093/philmat/nkaf003
Philosophia Mathematica
Advance Access Publication on May 23, 2025

Fregean Metasemantics

Ori Simchen*

Department of Philosophy, University of British Columbia, Vancouver, B.C. V6T1Z1, Canada

ABSTRACT

How the semantic significance of numerical discourse gets determined is a metasemantic issue par excellence. At the sub-sentential level, the issue is riddled with difficulties on account of the contested metaphysical status of the subject matter of numerical discourse, *i.e.*, numbers and numerical properties and relations. Setting those difficulties aside, I focus instead on the sentential level, specifically, on obvious affinities between whole numerical and non-numerical sentences and how their significance is determined. From such a perspective, Frege's 1884 construction of number, while famously mathematically untenable, fares better metasemantically than extant alternatives in the philosophy of mathematics.

1. INTRODUCTION

The semantic significance of numerical discourse presents us with well known difficulties.[1] Consider the following pair:

(1) One directly follows zero.
(2) Amy loves Mary.

Taken together, (1) and (2) incline us to think that numerals should stand for numbers in much the way that names stand for individuals named. Both appear to have the familiar aRb form and both appear to be true just in case the expression on the left stands for something that bears the relation conveyed by the relational predicate to whatever the expression on the right stands for.

∗Orcid.org/0000-0002-5417-3404. E-mail: ori.simchen@ubc.ca.

[1] By 'numerical discourse' here and elsewhere I mean discourse pertaining to numbers in its pure (*e.g.*, 'Twice two is four') rather than applied (*e.g.*, 'Jupiter has ninety-five moons') variety.

Such surface similarities incline us to suppose that what we say about how the numerals stand for the numbers in (1) should follow closely what we say about how the names stand for the named individuals in (2). Unfortunately, our expectations are frustrated here, at least according to a dominant metasemantic view according to which a name stands for its bearer in virtue of some causal-historical chain of communication extending back from its contemporary use to whatever was so named. The *locus classicus* is of course [Kripke, 1980, pp. 91–97], but an earlier and vivid statement can also be found in [Geach, 1969, pp. 288–289]:

> I do indeed think that for the use of a word as a proper name there must in the first instance be someone acquainted with the object named. But language is an institution, a tradition; and the use of a given name for a given object, like other features of language, can be handed on from one generation to another; the acquaintance required for the use of a proper name may be mediate, not immediate. Plato knew Socrates, and Aristotle knew Plato, and Theophrastus knew Aristotle, and so on in apostolic succession down to our own times; that is why we can legitimately use 'Socrates' as a name the way we do. It is not our knowledge of this chain that validates our use, but the existence of such a chain; just as according to Catholic doctrine a man is a true bishop if there is in fact a chain of consecrations going back to the Apostles, not if we know that there is.

But whatever numbers are, it is difficult to see how numerals could stand in such relations to them.[2] And yet parallels between (1) and (2) are striking enough to suggest a way forward in coming to appreciate what a successful account of the significance of the likes of (1) should look like.

In what follows I set aside the familiar metasemantic problems afflicting the sub-sentential level and focus instead on the sentential level. In particular, I bracket urgent questions pertaining to how numerals get to stand for numbers, functors get to stand for functions, and so on, and explore instead the idea that the way whole sentences such as (1) come to mean what they do follows closely the way the likes of (2) come to mean what they do. I believe such parallels have not been sufficiently appreciated and that focusing on them can bring us closer to a better understanding of the overall semantic significance of numerical discourse.[3]

The present focus on the metasemantics of full sentential significance allows for a fuller appreciation than is otherwise available of the achievement of Frege's 1884 construction of number. From such a metasemantic perspective, it will be shown that Frege's logicism complies with a natural and intuitive requirement

---

[2] For an effective articulation of this worry, see [Hodes, 1984]. Recent attempts to overcome it, inspired by [Kripke, 1992], are found in [Gómez-Torrente, 2020, Ch. 4] and [Simchen, 2017, App. III].

[3] An influential discussion of parallels between mathematical and non-mathematical sentences is found in [Benacerraf, 1973], but with a semantic sub-sentential focus, as opposed to a metasemantic sentential focus, which is the focus here.

that numerical truths should be established by proofs emanating from what numbers and numerical properties and relations are. Later developments in the philosophy of mathematics, largely in reaction to the mathematical untenability of Frege's logicism, do not comply with this natural and intuitive requirement. This includes the neo-logicist program. The ensuing discussion will illustrate the broadly methodological point that metasemantic plausibility and mathematical tenability need not align. Whether or not a mathematically tenable and yet metasemantically plausible account of the semantic significance of numerical discourse is forthcoming remains an open question.

2. FULL-SENTENTIAL SIGNIFICANCE

Let us consider first how (2) gets to mean what it does. Opting for the standard truth-conditional analysis, we let the association of (2) with a certain truth condition represent (2)'s meaning what it does. How is it that (2) has its particular truth condition, namely, Amy's loving Mary? A familiar and somewhat platitudinous answer is that (2) is associated with Amy's loving Mary in virtue of what 'Amy', 'loves', and 'Mary' contribute to the meaning of (2) *inter alia*, namely Amy, the loving relation, and Mary.[4] If (2) is true, its truth is determined by what 'Amy', 'loves', and 'Mary' stand for; if (2) is false, its falsity is likewise determined by what 'Amy', 'loves', and 'Mary' stand for. We would say that (2) is true or false because of, or in virtue of, what its semantically significant sub-sentential parts stand for (and how those parts are put together, and how matters stand with Amy, Mary, and the loving relation). This is our common metasemantic lore when it comes to the determination of full-sentential significance for the likes of (2).[5]

When it comes to how (1) gets to mean what it does, the details are far less platitudinous. Letting full-sentential significance be represented again by the association with a truth condition, we would like to say that (1) is associated with the condition of one's being an immediate successor of zero in virtue of what 'one', 'directly follows', and 'zero' contribute to the meaning of (1) *inter alia*, namely the number one, the relation of immediate succession, and the number zero, respectively. Truth or falsity for (1) should be determined by what 'one', 'directly follows', and 'zero' stand for. Indeed, here, too, it would appear that (1) is true or false because of, or in virtue of, what its semantically significant sub-sentential parts stand for (and how those parts are put together, and how matters stand with one, zero, and immediate succession). But here, unlike the case of (2), (1) admits of proof. What kind of proof might that be?

Insofar as we wish to say that (1) is true because of what its semantically significant sub-sentential parts stand for, it is most natural to suppose that a

[4] In a formal semantic setting we would speak of what 'Amy' and 'Mary' denote, and then of what 'loves Mary' denotes. I set aside such complications in everything that follows. For a discussion of the basic role of currying in semantics, see [Helm and Kratzer, 1997, §§2.3, 2.4]. See also the discussion in [Simchen, 2017, Ch. 3].

[5] I set aside for now whether such an explanation begs questions against metasemantic interpretationism. I return to the topic in Section 5.

proof of (1), as a route to its truth, should emanate from what one, zero, and immediate succession are.[6] Think of (2) again. Here we wanted to say that the truth or falsity of (2) should emanate from its subject matter, namely, the particular individuals involved and the relation at issue. In the case of (1), too, we would like to say that its truth proceeds bottom-up from what those numbers are and what this relation is, except that here the establishment of truth is via proof. The proof in question should proceed likewise from what those numbers are and what this relation is. This is a non-trivial yet natural requirement on the proof in question; and by extension, I want to suggest, a requirement on how claims such as (1) come to mean what they do. We may think of it as a metasemantic criterion of adequacy on the proof of such claims.[7]

As I said, the bottom-up requirement on the proof of (1) is not trivial. Consider, for example, von Neumann's set-theoretical construction of the natural numbers as structure $\mathbb{N}$ defined on the universe of pure sets by operation S that yields for any set the union of that set with its singleton set, starting with the empty set. If we stipulate that zero is to be the empty set and one is to be its singleton set, and stipulate further that immediate succession is to be the operation S structuring $\mathbb{N}$, the proof that one directly follows zero becomes an immediate consequence of the structure $\mathbb{N}$ being what it is under those stipulations.[8] But we wouldn't say that the proof here emanates from what one and zero and immediate succession *are*. Our stipulations have the requisite consequence, sure, but those are stipulative restrictions on the general structure $\mathbb{N}$. That one directly follows zero is determined here by what the general structure $\mathbb{N}$ is under the relevant stipulations. It isn't determined by what the items constituting the pre-theoretical subject matter of (1) are, namely, one, zero, and immediate succession.

[6] Speaking of proof as a route to truth is meant to track how working mathematicians and lay people think of proving theorems as means for establishing their truth. It is a well known result in logic that for any formalization of basic arithmetic, there are true sentences of the language of the formal system that are neither provable nor refutable in the system (assuming the system's soundness with respect to the intended interpretation). But this general limitative result doesn't take away from the basic point that there are significant differences between kinds of proof for arithmetical truths from the standpoint of our common metasemantic lore, which is the focus of the present discussion.

[7] It is not meant as a criterion of adequacy on the proof of *every* numerical claim. The numerical equivalent of

> there is a sentence of formalized type theory which is neither provable nor refutable in the system on the condition of the system's ω-consistency

isn't such a claim. But for metasemantic purposes, the apparent similarity between the likes of (1) and (2) does suggest the bottom-up constraint suggested in the text. I do not have a criterion for deciding *when* this constraint on proofs should come into play as a general matter.

[8] Letting the directly-follows relation be the operation S structuring $\mathbb{N}$, for any $X, Y \in \mathbb{N}$, Y directly follows X just in case $Y = X \cup \{X\}$. Given that $\emptyset \cup \{\emptyset\} = \{\emptyset\}$, $\{\emptyset\}$ directly follows $\emptyset$. Therefore, insofar as one is chosen to be $\{\emptyset\}$ and zero is chosen to be $\emptyset$, one directly follows zero.

To see this, consider alternative stipulations according to which one is to be the empty set, zero is to be the singleton set of the empty set, and immediate succession is to be S in all cases but for the following two: the empty set directly following the singleton set of the empty set, and the set with only the empty set and its singleton set as members directly following the empty set. (The empty set and its singleton "switch places" on this alternative scheme.) (1) would still come out true under the alternative stipulations, and its proof would emanate from what $\mathbb{N}$ is under the alternative stipulations.[9] In this respect, the proof of (1) under the von Neumann construction is top-down rather than bottom-up. First, a structure $\mathbb{N}$ is defined on the universe of pure sets. Officially, this has nothing to do with the numbers or relations among them. Next, certain stipulations are entered as to what the directly-follows relation, zero, and one are to be in that structure. Finally, that one directly follows zero falls out of the construction together with those stipulations. The proof here doesn't take into account the pre-theoretical subject matter — zero, one, and the relation in question — beyond the stipulations. If we were to consider the parallel situation with (2) again — despite there being no question of proof here — such a determination of truth would construe the truth of (2) as determined by a restriction on a certain general structure of entities standing in some relation stipulated to be the loving relation under further stipulations regarding who Amy and Mary are among the entities so related. It is clearly a significant departure from how we naturally and ordinarily think of the determinants of truth for the likes of (2).

In the next two sections we will witness how Frege's 1884 construction of number in *The Foundations of Arithmetic* [1953] delivers an understanding of the truth of (1) that meets the natural and intuitive requirement that proof should emanate bottom-up from what the numbers zero and one and the relation of immediate succession are. Alternative construals, notably neo-logicist ones, fail to deliver this sort of understanding.[10] The immediate moral drawn from all this is that we still lack a satisfying account of what it takes for whole numerical sentences to mean what they do beyond Frege's original indications, a framework famously mathematically untenable. But on the positive side, at least we see here the metasemantic direction one should pursue for a satisfying account of whole sentential significance for numerical discourse. We have a guideline of sorts for what to expect when it comes to the significance of numerical discourse. And on a larger methodological scale, a theory which is

---

[9] We let S^* be $S\setminus\{\langle\{\emptyset\},\emptyset\rangle,\langle\{\emptyset,\{\emptyset\}\},\{\emptyset\}\rangle\}$ and let the directly-follows relation be $S^*\cup\{\langle\emptyset,\{\emptyset\}\rangle,\langle\{\emptyset,\{\emptyset\}\},\emptyset\rangle\}$. Insofar as the directly-follows relation includes $\langle\emptyset,\{\emptyset\}\rangle$, and insofar as one is chosen to be $\emptyset$ and zero is chosen to be $\{\emptyset\}$, one directly follows zero.

[10] The top-down character of proving the likes of (1) is a feature of both structuralist and ordinal conceptions of the natural numbers. For a recent comparative discussion, see [Snyder *et al.*, 2018]. Consideration of space precludes me from discussing these alternatives in detail, but suffice it to say that according to both kinds of approach, the proof of (1) invariably proceeds from the definition of a certain structure and stipulations regarding the occupants of nodes in the structure. In what follows I focus instead on two varieties of cardinalism, Frege's original logicist project, and neo-logicism.

so spectacularly untenable for one theoretical purpose — mathematical foundations — is revealed as more attractive than extant alternatives for another theoretical purpose — metasemantics (at least on one common construal of this explanatory endeavor).

3. FREGE'S 1884 CONSTRUCTION

To get a concrete sense of what a bottom-up proof of the likes of (1) looks like, we turn to Frege's construction of number in *The Foundations of Arithmetic* [1953]. The details are everything here and do not require extensive mathematical sophistication. In general, Frege's work in the foundations of mathematics can hardly be appreciated without attention to those details, but in our case such attention is critical for a proper appreciation of the structure of the relevant proofs.

Frege's construction of number contains various contributions of lasting importance to philosophy. Among them, and most fundamentally, Frege distinguishes sharply between concepts and objects. Neither category is definable — a matter to which we return briefly in the next section — but concepts are characterized as functions of a certain kind, functions to truth values, while the two truth values are but two among the objects. Fregean concepts come in different "levels": a first-level concept maps objects to truth values, a second-level concept maps first-level concepts to truth values, a third-level concept maps second-level concepts to truth values, and so on. Fregean concepts also come in different arities: monadic, dyadic, triadic, and so on, making them akin to traditional properties and relations. The two dimensions of variability interact: the relation between a thing and its properties, for example, is a dyadic Fregean concept mapping objects and first-level concepts to truth values. (Such a concept is neither first- nor second-level but rather cross-level.) Concepts are associated with their extensions, which are objects. Those extensions consist of all and only the arguments for which the associated concepts yield truth.[11]

Against this background, Frege defines the natural numbers as extensions of certain second-level concepts.[12] He begins by defining the dyadic second-level concept of equinumerosity, or "concept equality": for first-level concept F to be equinumerous with first-level concept G is for there to be a first-level dyadic concept ϕ such that everything that is F bears ϕ to a unique thing that is G and for everything that is G, there is unique thing that is F that bears ϕ to it. In other words, ϕ is a one-to-one correspondence between the Fs and the Gs. Equinumerosity between first-level concepts is a matter of there being such a correspondence between their extensions. A number n is then construed as

[11] Reassuringly, Frege writes: "I assume it is known what the extension of a concept is" [1953, §69 n.1]. Ironically, it is precisely the association of concepts with their extensions, and more generally of functions with their courses-of-values, which is captured by the notorious Basic Law V that infects Frege's foundations with the inconsistency discovered by Russell.

[12] We follow the usual practice of assuming concepts without arity specification to be monadic.

the extension of the second-level concept of being equinumerous with F, where F is a first-level concept with an extension consisting of exactly n objects. Specifically, the number zero is defined as the extension of the second-level concept of being equinumerous with the first-level concept of being self-distinct, a concept with a null extension (for nothing is self-distinct and provably so under Frege's Leibnizian definition of objectual identity[13]). The number one is then defined as the extension of the second-level concept of being equinumerous with the first-level concept of being identical with zero, a concept with an extension consisting of exactly a single object, namely, zero. Two can then be defined as the extension of the second-level concept of being equinumerous with the first-level concept of being either identical with zero or identical with one, a concept with an extension consisting of exactly two objects, namely, zero and one. And so on. Number n being an immediate successor of number m is defined as there being a first-level concept F and an object y such that y falls under F and n is the extension of being equinumerous with F while m is the extension of being equinumerous with the first-level concept F-other-than-y.

The proof that one directly follows zero then proceeds via the lemma that the extension of the second-level concept of being equinumerous with the first-level concept of being self-distinct — namely, zero — is identical with the extension of the second-level concept of being equinumerous with the first-level concept of being both identical with zero and not identical with zero. The lemma is proved by the fact that the first-level concept of being self-distinct is equinumerous with the first-level concept of being both identical with zero and not identical with zero — no object falls under either concept — together with the auxiliary fact, known as Hume's Principle, that the extension of a second-level concept of being equinumerous with a first-level concept F is identical with the extension of a second-level concept of being equinumerous with a first-level concept G just in case F and G are equinumerous. Now, the concept of being identical with zero and the object zero are indeed such that this object falls under this concept (zero being, like anything else, self-identical); and one is the extension of the second-level concept of being equinumerous with this concept of being identical with zero by definition; and zero is the extension of the second-level concept of being equinumerous with this concept of being identical with zero while being other than this object, zero, by the definition of zero and the lemma. And so, one directly follows zero.

Zooming out, we observe that the arithmetical truth of one's directly following zero is established via a proof that essentially depends on the definitions of the number zero, the number one, and the relation of immediate succession. This demands a brief comment. Officially, Fregean definitions assign significance to symbols. But Frege's understanding of semantic significance in 1884 is

---

[13] See [Frege, 1953, §§65 and 74]. Object x being identical with object y is defined along Leibnizian lines as x being F if and only if y is F for every first-level concept F. Under this definition, the assumption that some object a is self-distinct amounts to its not being the case that a is F if and only if a is F for some F, which entails a contradiction. Hence no object is such.

undifferentiated as per the later distinction between sense and reference — it is a notion of undifferentiated semantic content (*Inhalt*). So when '1' is definitionally assigned the extension of the second-level concept of being equinumerous with the first-level concept of being identical with zero as its content, this assignment provides *both* for the intuitive meaning of the sign *and* for what the sign stands for, namely, the number one itself. Small wonder that the opening sentence of the Introduction to [Frege, 1953] treats the question what the number one is as interchangeable with the question what the numeral '1' means.[14] The definitions Frege offers are content-givers in this undifferentiated sense of 'content'. This means that Frege's proof that one directly follows zero emanates from what the numbers zero and one *are*, and what the relation of immediate succession *is*, as given by their respective definitions.

Frege offers the following general and rather picturesque characterization of his logicism:

> The truth is that [the conclusions drawn] are contained in the definitions, but as plants are contained in their seeds, not as beams are contained in a house. Often we need several definitions for the proof of some proposition, which consequently is not contained in any one of them alone, yet does follow purely logically from all of them together. [Frege, 1953, §88]

Conclusions are contained in definitions as plants are contained in seeds. The bottom-up growth of those conclusions is traced by proofs that establish them. Frege's method of proof for the arithmetical fact that one directly follows zero is unusual in comparison with later approaches. And yet, as discussed in the previous section, when it comes to the metasemantics of the claim that one directly follows zero, Frege's method nicely comports with what we would like to say about the significance of claims of the form aRb quite generally. It is a mathematical oddity, perhaps, but one with a certain unmistakable metasemantic appeal. The Fregean proof that one directly follows zero proceeds bottom-up from what one and zero and immediate succession are, as opposed to a top-down approach such as the proof that proceeds from von Neumann's definition of the structure $\mathbb{N}$ together with the attendant stipulations.

Frege's bottom-up method of proof also highlights an easily neglected aspect of his logicist project. Classical logicism is standardly thought of as a two-pronged reductionist project of showing that (i) mathematical truths — particularly arithmetical ones — are logical truths, and that (ii) mathematical objects — particularly arithmetical ones, numbers — are logical objects. But in Frege's hands we get as a further requirement a certain connection between (i) and (ii). Arithmetical truths about numbers are shown to follow logically from what the numbers themselves — together with their properties and relations — are, as given by their definitions. Mathematical truths are logical truths because mathematical objects are logical objects, as revealed by their definitions.

[14] "When we ask someone what the number one is, or what the symbol 1 means, we get as a rule the answer 'Why, a thing"' [Frege, 1953, p. xiii].

Here is a further illustration of this last point. In [Frege, 1953, §73] we get a proof of the right-to-left direction of Hume's Principle, the claim that the extension of the second-level concept of being equinumerous with F — "the number belonging to F" — is identical with the extension of the second-level concept of being equinumerous with G — "the number belonging to G" — just in case F and G are equinumerous.[15] Assuming F and G are equinumerous, there is a one-to-one correspondence between the Fs and the Gs. And for any first-level concept H equinumerous with F, there is a one-to-one correspondence between the Hs and the Fs. So the composition of the two one-to-one correspondences — between the Fs and the Gs and between the Hs and the Fs — is itself a one-to-one correspondence between the Hs and the Gs. So H and G are equinumerous. And so, any first-level concept equinumerous with F is equinumerous with G. An analogous argument shows that any first-level concept equinumerous with G is equinumerous with F, which completes the proof that the extension of being equinumerous with F is identical with the extension of being equinumerous with G on the condition of the equinumerosity of F and G. This proof capitalizes on the kind of object numbers belonging to concepts are. Insofar as they are extensions, and insofar as the identity of extensions is a matter of anything belonging to the one belonging to the other and vice versa, from the equinumerosity of first-level concepts we get the identity of the numbers belonging to them. This is a *substantive* grounding of Hume's Principle. It is a truth about numbers that admits of proof, a proof generated by what the numbers are as given by their explicit definitions. I return to this particular aspect of Frege's logicist project in the next section.

As a final illustration of the requirement that proofs emanate bottom-up from their subject matter as given by relevant definitions, consider Frege's proof [1953, §75] that every concept under which no object falls is equinumerous with every concept under which no object falls and no other. In this case, the proof issues directly from the definition of equinumerosity. Letting F and G be first-level concepts under which no object falls, every dyadic relation ϕ is such that for any object that is F there is an object that is G to which it is ϕ related (since nothing is F) and for any object that is G there is an object that is F ϕ related to it (since nothing is G). But insofar as this holds for any dyadic relation, it holds in particular for objectual identity, which is one-to-one. It follows that F and G are equinumerous by the definition of that relation. On the other hand, letting F be a first-level concept under which no object falls as before and letting H be a concept under which some object a falls, for every dyadic relation ϕ there is no object that is F and bears ϕ to a because there is no object that is F in the first place. So for every such relation ϕ, it isn't the case that for every object that is H there is an object that is F which is ϕ

[15] The left-to-right direction is trivial. If the extension of being equinumerous with F is identical with the extension of being equinumerous with G, any first-level concept belonging to the one belongs to the other and vice versa; so any first-level concept equinumerous with F is equinumerous with G. But F itself is equinumerous with F — objectual identity being a one-to-one correspondence — and so, F is equinumerous with G.

related to it. So there is no dyadic relation such that for every object that is F there is an object that is H to which it is ϕ related and such that for every object that is H there is an object that is F which is ϕ related to it. And so, F and H are not equinumerous. The proof here clearly emanates from what the relation of equinumerosity is as given by its definition.

4. NEO-LOGICISM AND METASEMANTICS

As is widely known, Frege's logicist project came to grief upon Russell's discovery of an inconsistency in its foundations. Frege's system includes Basic Law V: the course-of-values (graph) of function f is identical with the course-of-values of function g just in case f and g yield the same value for each argument. Regarding first-level concepts, we get the extension of F being identical with the extension of G just in case an object falls under F just in case it falls under G. The contradiction Russell discovered emerges directly from this principle. Assuming the version of Basic Law V regarding first-level concepts, the following is a special case: the extension of F is identical with the extension of F just in case an object falls under F just in case it falls under F. The right-hand side of the biconditional is a logical truth; so we conclude that the extension of F is identical with the extension of F. Existentially generalizing on one side of the identity, and universally generalizing on F, yields a principle of unrestricted comprehension for first-level concepts: for any such concept there is an object identical with the concept's extension. Defining the membership of x in y as there being a first-level concept H such that x falls under H and y is identical with the extension of H, and selecting the first-level concept of being non-self-membered, it is easily shown that this concept's extension is a member of itself just in case it isn't.

We now know that the standard Dedekind–Peano axiomatization of arithmetic can be derived from Hume's Principle in second-order logic without reliance on Basic Law V.[16] The neo-logicist idea is to discard Basic Law V and treat the following version of Hume's Principle as an axiom: the number of Fs is identical with the number of Gs just in case F and G are equinumerous. But we are not construing the left-hand side of this biconditional as an identity of extensions of second-level concepts. The meaning of 'the number of Xs' (or 'the number belonging to X') is given contextually by instances of the new axiom. Under this proposal, we identify zero as the number belonging to the first-level concept of being self-distinct, one as the number belonging to the first-level concept of being identical with zero, and so on, without defining numbers as objects of any sort beyond such characterizations. We can define n's being an immediate successor of m along the original Fregean lines as there being a first-level concept F and an object y falling under F such that n is identical with the number belonging to F and m is identical with the number belonging to F-other-than-y. The proof of (1) can then proceed exactly

[16] Basic Law V is only needed for the proof of Hume's Principle itself. For a comprehensive history of work on this issue, see [Heck, 2011, Ch. 1]. A principal inspiration here is [Geach, 1955]. The idea was subsequently developed in [Parsons, 1965; Wright, 1983], and elsewhere.

as before, noting that no mention need be made of extensions of the relevant concepts.[17] At the end of the day, all we need for the proof of (1) beyond the modified definition of immediate succession is the identification of zero as the number belonging to the first-level concept of being self-distinct, and of one as the number belonging to the first-level concept of being identical with zero, *whatever else those happen to be*. From the neo-logicist perspective, the whole point of the exercise is to prove (1) and its ilk without identifying numbers as any particular kind of object beyond belonging to certain concepts and whatever that entails. We are in effect proving the likes of (1) without our proof's emanating from what zero and one are *specifically*, what they are specifically beyond belonging to this or that concept. And if we think of what the relation of immediate succession is as depending on what its relata are, we are also proving the likes of (1) without our proof depending on what the relation of immediate succession is *specifically*.

One way to appreciate what is at stake here is to consider the situation again as it pertains to the likes of (2), just as we did earlier regarding the von Neumann construction of $\mathbb{N}$. Even if there is no question of proving (2), an analog of the neo-logicist strategy in this case would be a construal of the truth of (2) as depending not on who Amy and Mary are specifically, specified in any which way, but on the general satisfaction of certain canonical descriptions of individuals, say descriptions of the form 'the person with center of mass at $\langle x, y, z, t\rangle$' used attributively.[18] In a similar vein, the analog of the neo-logicist strategy would construe the truth of (2) as depending not on what the relation of loving is directly, but rather on the relation of loving as holding between satisfiers of such canonical descriptions, whoever those satisfiers happen to be as specified independently of the satisfaction of those descriptions. It is certainly a significant departure from how we normally think of the determinants of truth for the likes of (2). The natural thought is that the relation of loving itself determines directly the truth of (2) *inter alia*, rather than the relation under a rendering that is essentially beholden to its relata being described in a particular way. Similarly, it is natural to think of the relation of immediate succession itself as determining directly the truth of (1) *inter alia* along traditional Fregean lines, rather than the relation under a rendering that is essentially beholden to the relation's relata specified as numbers belonging to certain concepts along neo-logicist lines.

Against the present background emphasis on proofs that issue bottom-up from definitions of their subject matter, there is also an important contrast between Hume's Principle and Basic Law V. The former admits of a proof that issues from what numbers belonging to concepts are, as given by their explicit

[17] For example, the proof of the lemma that the concept of being self-distinct is equinumerous with the concept of being both identical with zero and not identical with zero nowhere requires mentioning extensions.

[18] Or perhaps 'the *actual* person with center of mass at $\langle x, y, z, t\rangle$', or 'dthat[the person with center of mass at $\langle x, y, z, t\rangle$]' if modal profile is a concern, but I set such distracting complications aside.

definitions. Basic Law V, on the other hand, admits of no such proof because its subject matter — concepts and their extensions in full generality — is too basic. The indefinability of concepts and their extensions precludes the possibility of proving Basic Law V in the way that would be required by Frege from relevant definitions of subject matter. This goes some way to answering an important question raised by Heck [2011, p. 11]: "Why didn't Frege himself consider retreating from Law V to HP when confronted with Russell's Paradox?". The answer proposed by Heck is that Frege thought Hume's Principle was vulnerable to the Julius Caesar problem and mistakenly thought Basic Law V wasn't.[19] The answer proposed here is that Hume's Principle admits of the kind of proof Frege demands, one that proceeds from definitions of its subject matter, while Basic Law V doesn't. Concepts and their extensions in full generality are just too basic to admit of definition, in a way that numbers belonging to concepts are not. So Basic Law V has what it takes as a candidate for being a basic truth, while Hume's Principle doesn't. In the different context of defending the general deductive character of his investigation, Frege writes: "[I]t is in the nature of mathematics always to prefer proof, where proof is possible, to any confirmation by induction" [1953, §2]. In this spirit, we can also say that it is in the nature of Frege's foundational efforts to prefer proof where proof is possible to positing a truth as primitive. And proof here is understood as bottom-up from definitions of subject matter.

5. CONCLUSION: METASEMANTIC OPTIONS

In 'What numbers could not be' Benacerraf writes:

> [F]or Frege a number was an equivalence class ... Although an appealing notion, there seems little to recommend it over, say, [von Neumann's]. [1965, p. 59]

The sentiment expressed in this passage neglects the metasemantic full-sentential aspect of Frege's construction, which has been our main focus here.[20] What the numbers are is only one facet of Frege's logicist aspirations. Surface similarities between the likes of (1) and (2) suggest a certain metasemantic parity at the full-sentential level, which Frege's logicism respects while von Neumann's construction does not. In such light, the contrast between the accounts seems rather stark. I end this discussion with a brief overview of a menu of

---

[19] Heck writes:

> Frege's 'solution' of the Caesar problem in *Die Grundlagen* consists in the identification of numbers with extensions of certain concepts. As has often been remarked, however, this solution works only if we assume that we know how to resolve the Caesar problem for extensions themselves ... The Caesar problem is, therefore, not so much solved by Frege's identification of numbers with extensions as it is relocated by it. [Heck, 2011, p. 16]

[20] In this connection, see also fn. 3.

metasemantic options and its implications for the metasemantics of numerical discourse.

The primary concern in metasemantics is how semantically significant items, be they linguistic expressions, signposts, bits of mentality, or anything else of semantic significance, gain their significance. To pick a paradigm case, when the metasemantic query is raised with respect to the name 'Amy', a metasemantic answer will seek to explain how the name ends up naming Amy. There are of course refinements and qualifications that would need to be entered at this point. One such is that the English name 'Amy' is a name borne by many individuals and should really be thought of as what Kaplan [1990] calls a *generic* name. And while there are surely interesting facts about how 'Amy' got to be included in English as a generic name, the metasemantic issues here are underexplored.[21] But when it comes to a specific name 'Amy', a name purporting to name someone in particular, the metasemantic question targets how the name comes to name the particular individual named. And here there is a metasemantic fault line that distinguishes two types of approach to the subject.

According to the first type of approach, the name comes to name the individual named in virtue of certain conditions surrounding the name's production or employment. It is sometimes claimed, for example, that referential occurrences of the likes of 'Amy' lie in the causal wake of the individual being so named, however remotely, through the presence of referential intentions governing certain uses of the name in context. This is supposed to explain how it is that 'Amy' names Amy in particular, rather than anyone else or no one.[22] On this type of approach, endowment with sub-sentential significance is prior in the order of metasemantic explanation to endowment with full-sentential significance. Sentences in context mean what they do because certain sub-sentential expressions mean what they do and are put together in a certain way within the sentence; those expressions mean what they do because of certain conditions surrounding their production or employment, such as the presence of referential intentions that depend on causal rapport of the appropriate type with the named individuals.[23]

The second approach to answering the question how 'Amy' comes to name Amy prioritizes the interpretive situation. A sub-sentential expression meaning what it does is a matter of belonging to a larger swath of discourse interpretable in a certain way under various constraints. On such an approach, what makes it the case that 'Amy' means what it does, ultimately, is that the likes of (2), various 'Amy'-containing sentences, mean what they do *inter alia*. 'Amy' standing for Amy is a downstream effect of (2) and its likes meaning what they do,

[21] It is tempting to suppose that the generic name 'Amy' doesn't have semantic significance and therefore that metasemantics need not concern itself with it. This seems hasty. The generic name is conventionally gendered, which suggests that it does carry, *qua* generic, some semantic information.

[22] In this connection, see also the [Geach, 1969] passage cited in Section 1.

[23] The explanatory priority of sub-sentential significance over sentential significance on such views isn't accidental, but a full discussion of the issue would take us too far afield. See [Simchen, 2017, §4.3] for discussion.

and is secondary in the order of metasemantic explanation. Under this second approach, factors such as causal connectedness between the name and the named might still come into play in the explanation of how it is that the name names the named, but only indirectly via matters of interpretability for whole sentences as further parameters in the selection of the best overall interpretation of the surrounding discourse. Under the first approach, by contrast, such causal connectedness would figure directly in the explanation of sub-sentential semantic facts and not in a way that is beholden to the interpretability of larger swaths of discourse.

There are compelling reasons for skepticism about the viability of the second approach — known as *interpretationism* — as a general approach to metasemantics. A full discussion of those reasons would take us too far afield, but a quick survey of a couple of them is in order. The first reason has to do with the phenomenology of sub-sentential significance. A speaker can utter 'President Biden ... never mind' without ever completing her sentence and without generating full-sentential significance in any other way while still managing to refer successfully.[24] Interpretationism invariably considers the semantic interpretation of 'President Biden' in the speaker's mouth as beholden to a larger swath of discourse that is stipulated to be missing in this case, prompting the interpretationist to locate the broader sentential context outside the speaker's immediate purview. And yet the speaker may be using the name to refer, for the very first time, to her newly acquired cat. The interpretationist would counter with the interpretability of the speaker's accompanying referential intention *vis-à-vis* the cat. Discussion of the latter move is for another day, but suffice it to say that the construal of referential intentions as themselves beholden to matters of interpretability has little to recommend it on general independent grounds.

A second reason for skepticism about the viability of interpretationism as a general approach to metasemantics has to do with matters of semantic indeterminacy and with the inability of interpretationism to rule out outlandish possibilities of reference being hopelessly garbled, even when the interpretationist position is amplified with so-called "reference magnetism".[25] This can be illustrated both with respect to singular terms and with respect to predicates, but in the interest of brevity we focus on the former. Briefly, it is a familiar feature of semantic interpretation that truth conditions for a swath of discourse can remain invariant under variability of assignments of semantic values to singular terms as per a given domain of individuals. This means that the bare interpretationist position fails to provide resources that would guarantee semantic determinacy for singular terms, something we are otherwise deeply invested in. We care a great deal about our demonstrative pronouns in context, for example, referring to certain particular objects rather than to some proxy objects lying outside our light cone. The interpretationist would then attempt to rule out deviant interpretations of singular terms by insisting

[24] For further discussion of the utility of such examples, see [Simchen, 2017, §4.3].

[25] See [Lewis, 1983; 1984].

on a certain fixity in the interpretation of predicates, most commonly via the Lewisian notion of naturalness. We are not free to vary the interpretation of predicates in any which way to compensate for deviance in the interpretation of singular terms, this interpretationist would argue. Some interpretations of predicates are just objectively better than others, better from the point of view of the world. Setting aside questions concerning what it is for one interpretation to be objectively better than another, this last move by the interpretationist can be countered by introducing a deviant semantic clause into the semantic apparatus of truth-in-a-model that would induce a permutation of assignments to the extra-logical vocabulary for atomic sentences, resulting in a deviant definition of what it is for a sentence to be true under an interpretation. This would allow indeterminacy in the interpretation of singular terms to remain despite magnetic fixity in the interpretation of predicates.[26]

It is sometimes supposed that the threat of semantic indeterminacy doesn't target interpretationism alone — that anti-interpretationist metasemantic views are just as vulnerable to it.[27] This, however, is a mistake borne by the idea that semantic indeterminacy must affect the language in which the anti-interpretationist position is articulated.[28] Endowment with semantic significance, according to the anti-interpretationist, is directly determined by conditions of production or employment of the items semantically endowed. Semantic indeterminacy under such an approach would be akin to genetic indeterminacy for a baby, or indeterminacy in which of two identical twins actually sat for a particular passport photo. It isn't easy to view such matters as credible worries for the anti-interpretationist, certainly not at the level at which interpretationism is afflicted by them.[29] But for all its faults, the interpretationist orientation can go some way towards explaining how mathematical discourse gains its semantic significance. Recall again the von Neumann construction. We might say that 'one', 'zero', and 'directly follows' mean what they do in virtue of the likes of (1) being susceptible to such proofs as the one sketched in Section 2, a top-down proof that emanates from what $\mathbb{N}$ is under suitable stipulations.[30] Here we are clearly prioritizing the meaning of (1) over the meanings of its

---

[26] This is easily shown by the deviant definition of truth-in-a-model compensating for the deviance in the interpretation of singular terms. I omit the technical details here and refer the interested reader to [Simchen, 2017, Ch. 2; Simchen, 2020].

[27] See, *e.g.*, [Sider, 2011, §3.2].

[28] For an early discussion of this "just more theory" move in a related context, see [Devitt, 1983]. For a more recent discussion, see [Simchen, 2017, Ch. 1].

[29] In this connection, see also the last sentence in the cited passage from [Geach, 1969, p. 289] in Section 1: "It is not our knowledge of this chain that validates our use, but the existence of such a chain; just as according to Catholic doctrine a man is a true bishop if there is in fact a chain of consecrations going back to the Apostles, not if we know that there is".

[30] Frege would strongly disagree here. In his polemic against formal theories of arithmetic, he writes:

> Since it is arbitrary [according to the formal mathematician] what reference one wants to give to a sign, it follows that the content of the sign will have these or

sub-sentential constituents in the order of explaining how numerical discourse generally gains its semantic significance. Given the truth of the entire claim, we are asking how the numerals and the relational predicate are to be interpreted to facilitate its proof. Unfortunately, interpretationism isn't plausibly extendable to non-mathematical discourse for at least the reasons outlined above. And the combination of interpretationism for mathematical discourse and anti-interpretationism for non-mathemantical discourse is unattractive both because it seems to undermine unmistakable pre-theoretical affinities between the likes of (1) and (2), and because it renders the discourse of applied mathematics, consisting of such sentences as 'Jupiter has ninety-five moons', metasemantically problematic.

If we think of mathematical discourse as continuous with other forms of discourse, it is difficult to avoid the conclusion that we don't have a workable idea of how to proceed theoretically when it comes to the metasemantics of even the simplest swaths of arithmetical language. (In the interest of caution we must add: assuming that interpretationism isn't a live option as a general metasemantics.) Familiarly, we lack a plausible story of how semantic significance for sub-sentential numerical expressions gets determined. And yet an important theme that emerges from Frege's construction of number, one that seems no less relevant to thinking about how numerical discourse gains its semantic significance, is the determination of full-sentential significance beyond the familiar perplexities surrounding the metaphysical status of the subject matter of numerical discourse. And here, Frege's construction proves superior to the extant alternatives.

Put another way: if numerical discourse is just an extension of general discourse, familiar options in the philosophy of mathematics become less promising from a metasemantic standpoint than is often assumed. This may be difficult to appreciate if sub-sentential semantic significance is the focus of our metasemantic attention, which leads very naturally to the familiar problem area of the metaphysics of number. Focusing on the metasemantics of 'zero' and 'one' leads very quickly to worries about the metaphysical status of zero and one. But if we shift our focus from sub-sentential significance to full-sentential significance and its determination, as we've done here, the contest between major competitors in the philosophy of mathematics becomes salient in a very different way. We can then appreciate the metasemantic superiority of Frege's original theory over those later alternatives, despite its foundational untenability, and also appreciate how little metasemantic advance has been gained beyond Frege's pioneering efforts. That, by itself, is an advance, if only a limited one.

> those properties, depending on the particular choice made. Therefore it in part depends on my will, which properties the content of the sign has. [1984, p. 115]

The complaint is that such arithmetical facts as that one directly follows zero — or in the context of the original discussion, that $\frac{1}{2} + \frac{1}{2} = 1$ — cannot issue from our acts of will, from our stipulations, but should emanate, rather, from what the numbers themselves and their properties and relations *are*.

ACKNOWLEDGEMENTS

Many thanks to members of my UBC/SFU seminar 'Talk About Numbers' during the spring term of 2023, and to two anonymous referees for *Philosophia Mathematica*, for discussion of this material.

REFERENCES

Benacerraf, P. [1965]: 'What numbers could not be', *Philosophical Review* **74**, 47–73. DOI: https://doi.org/10.2307/2183530.

——— [1973]: 'Mathematical truth', *Journal of Philosophy* **70**, 661–679. DOI: https://doi.org/10.2307/2025075.

Devitt, M. [1983]: 'Realism and the renegade Putnam: A critical study of *Meaning and the Moral Sciences*', *Noûs* **17**, 291–301. DOI: https://doi.org/10.2307/2215148.

Frege, G. [1953]: *The Foundations of Arithmetic.* Oxford: Basil Blackwell.

——— [1984]: *Collected Papers on Mathematics, Logic, and Philosophy.* Oxford: Basil Backwell.

Geach, P. [1955]: 'Class and concept', *Philosophical Review* **64**, 561–570. DOI: https://doi.org/10.2307/2182633.

——— [1969]: 'The perils of Pauline', *Review of Metaphysics* **23**, 287–300.

Gómez-Torrente, M. [2020]: *Roads to Reference.* Oxford University Press. DOI: https://doi.org/10.1093/oso/9780198846277.001.0001.

Heck, R.K. [2011]: *Frege's Theorem.* Oxford University Press. (Originally published under the name 'Richard G. Heck, Jr'.)

Helm, I., and A. Kratzer [1997]: *Semantics in Generative Grammar.* Malden, Mass.: Blackwell.

Hodes, H. [1984]: 'Logicism and the ontological commitments of arithmetic', *Journal of Philosophy* **81**, 123–149. DOI: https://doi.org/10.2307/2026440.

Kaplan, D. [1990]: 'Words', *Proceedings of the* Aristotelian Society, Supplementary *Volume* **64**, 93–119. DOI: https://doi.org/10.1093/aristoteliansupp/64.1.93.

Kripke, S. [1980]: *Naming and Necessity.* Cambridge, Mass.: Harvard University Press.

——— [1992]: '*Logicism, Wittgenstein, and de re beliefs about numbers'*, Whitehead Lectures, Harvard University. Unpublished.

Lewis, D. [1983]: 'New work for a theory of universals', *Australasian Journal of Philosophy* **61**, 343–377. DOI: https://doi.org/10.1080/00048408312341131.

——— [1984]: 'Putnam's paradox', *Australasian Journal of Philosophy* **62**, 221–236. DOI: https://doi.org/10.1080/00048408412340013.

Parsons, C. [1965]: 'Frege's theory of number', in M. Black, ed., *Philosophy in America*, pp. 180–203. Ithaca, N.Y.: Cornell University Press. DOI: https://doi.org/10.4324/9781315830636-16.

Sider, T. [2011]: *Writing the Book of the World.* Oxford Univeresity Press. DOI: https://doi.org/10.1093/acprof:oso/9780199697908.001.0001.

Simchen, O. [2017]: *Semantics, Metasemantics, Aboutness.* Oxford Unversity Press. DOI: https://doi.org/10.1093/acprof:oso/9780198792147.001.0001.

——— [2020]: 'Modeling truth for semantics', *Analytic Philosophy* **61**, 28–36. DOI: https://doi.org/10.1111/phib.12168.

Snyder, E., S. Shapiro, and R. Samuels [2018]: 'Cardinals, ordinals, and the prospects for a Fregean foundation', *Royal Institute of Philosophy Supplement* **82**, 77–107. DOI: https://doi.org/10.1017/s1358246118000176.

Wright, C. [1983]: *Frege's Conception of Numbers as Objects.* Aberdeen University Press.