THE REVIEW OF SYMBOLIC LOGIC
Volume 16, Number 1, March 2023

PREDICATIVISM AS A FORM OF POTENTIALISM

ØYSTEIN LINNEBO
University of Oslo
and

STEWART SHAPIRO
The Ohio State University

Abstract. In the literature, predicativism is connected not only with the Vicious Circle Principle but also with the idea that certain totalities are inherently potential. To explain the connection between these two aspects of predicativism, we explore some approaches to predicativity within the modal framework for potentiality developed in Linnebo (2013) and Linnebo and Shapiro (2019). This puts predicativism into a more general framework and helps to sharpen some of its key theses.

§1. A question about two aspects of predicativism.

It is a platitude that a definition must not be circular. A definition is said to be *impredicative* if it defines an entity by using quantifiers whose range includes the entity to be defined; otherwise, the definition is *predicative*. It is sometimes argued that impredicative definitions are circular because, by quantifying over the entity to be defined, they somehow presuppose this entity. Consider, for example, Russell's [38, p. 237] *Vicious Circle Principle*:

> No totality can contain members defined in terms of itself.

Or [39, p. 198]:

> ...whatever in any way concerns all or any or some of a class must not be itself one of the members of a class.

As is well known, Russell's attempt to observe the Vicious Circle Principle resulted in the *ramified hierarchy*, which involves a notational nightmare that we hope to avoid.

In his classic "Systems of predicative analysis" [8, p. 2], Feferman highlights another aspect of predicativism, namely the view that some totalities are inherently *potential*:

> ... we can never speak sensibly (in the predicative conception) of the "totality" of all sets as a "completed totality" but only as a *potential totality* whose full content is never fully grasped but only *realized in stages*.

As the context makes clear, the sets in question are sets of natural numbers.[1]

Received: May 21, 2020.
2020 *Mathematics Subject Classification*: 03A05, 03B20, 03F35.
Keywords and phrases: predicativity, impredicativity, Poincaré, Russell, Feferman.

[1] Feferman's claim that there is no "completed totality" of such sets is thus far more radical than the corresponding claim about Cantorian sets. This latter claim was made by Cantor

doi:10.1017/S1755020321000423

What is the connection between these two aspects of predicativism? That is, what is the connection between the Vicious Circle Principle and the thesis that there is no "completed totality" of all (predicatively defined) sets?

A similar question is prompted already by the work of Poincaré, the father of predicativism, who gives two *prima facie* different diagnoses of the logical and set-theoretic paradoxes. On the one hand, he blames the paradoxes on the use of viciously circular definitions (see, for example, [34, Section XI]). On the other hand, he also puts the blame on the Cantorian assumption that there are completed infinite collections: "*There is no actual* (*given complete*) *infinity*. The Cantorians have forgotten that, and they have fallen into contradiction" (ibid.)

Our primary claim is that these two aspects of predicativism—the Vicious Circle Principle and the view of certain totalities as inherently potential—are closely related and reinforce each other. To demonstrate this, we develop a predicativist approach to arithmetic within a modal framework for potentiality from Linnebo [27] and Linnebo and Shapiro [30]. The resulting analysis of predicativism is also illuminating, we believe, because it places predicativism in a more general framework of potentiality, alongside other forms of potentialism. This allows us to sharpen some of the key theses of predicativism. In particular, we will be able to make good sense of Feferman's more recent "semi-intuitionistic" approach to predicativity.[2] Moreover, our analysis allows us to address, in a sharp manner, the charge that predicativism is unstable because the notion of predicative definability is not itself predicatively definable. We show the extent to which the various predicativist theses can be stated in a manner that is acceptable to the predicativist herself.

§2. Our answer in a nutshell. Our answer to the important question of how the two aspects of predicativism are related can be found already in Poincaré's later work on the topic.[3] As we have seen, he rejects the completed infinite. He insists instead on a potentialist conception:

> when we speak of an infinite collection, we mean a collection to which we can add new elements unceasingly (similar to a subscription list which would never end, waiting for new subscribers). (Poincaré, [35], p. 463/47)

This much is similar to Aristotle's conception of infinity as merely potential. But Poincaré goes on and asks us to consider attempts to "classify" the elements of one of these potentially infinite collections. In some cases, this poses no problems. For

---

in his correspondence with Hilbert and Dedekind (see [7, pp. 926–940]), as well as by Zermelo [48].

2 This approach is not even mentioned in his survey article on predicativity [10], but is discussed in [12] and a number of more recent, though unpublished, manuscripts. See also Rathjen [36, 37].

3 Our analysis and interpretation of Poincaré is inspired by Crosilla [3], as well as by Kreisel [24]. The occasion of Feferman's paper quoted above was the concurrence of two approaches to predicative analysis (the second-order theory of sets of natural numbers): his own, from the point of view of the Poincaré analysis, and that of Schütte [40], from the point of view of Russell's Vicious Circle Principle. Both approaches arrived at essential the same theory of predicative analysis.

example, the classification as being among the first 1000 subscribers to a certain newspaper is undisturbed by later additions to the subscription list.

In other cases, however, "the principle of the classification rests on some relation of the elements to be classified with the entire collection" (ibid., pp. 46–47). An example is the classification as being the youngest subscriber to the newspaper. Although 10-year-old Billy may be thus classified today, tomorrow that honor may go to 9-year-old Sally, who may then just have started subscribing. From considerations like these, Poincaré concludes:

> From this we draw a distinction between two types of classifications applicable to the elements of infinite collections: the *predicative* classifications, which cannot be disordered [bouleversé] by the introduction of new elements; the *non predicative* classifications in which the introduction of new elements necessitates constant modification (ibid., p. 47).

That is, a classification of the elements of a potentially infinite collection is deemed predicative or not in accordance with whether or not the classification can be "disordered" by later additions to the collection.

Now, the easiest and most natural way to ensure this kind of stability of a classification is to require that all of its quantifiers are restricted to elements of the potentially infinite collection that have already been introduced. For example, let some objects *aa* be all the current subscribers to a newspaper. Then the classification as being the youngest subscriber among *aa* is predicative in Poincaré's sense of not being susceptible to being overturned as the collection of subscribers expands.[4] By restricting our attention to "old" elements of the collection, we ensure that the classification cannot be disrupted by any addition of "new" elements.

This brings us to the Vicious Circle Principle as a proposed requirement on permissible definitions. Consider the case of sets of natural numbers. Since (according to predicativists) there is no completed totality of such sets, this totality must be merely potential. In Poincaré's words, it must be a collection "to which we can add new elements unceasingly." Suppose now that we wish to use a formula $\varphi(n)$ to define a set $\{n : \varphi(n)\}$ of natural numbers. For this definition to be permissible, we must ensure that it "cannot be disordered by the introduction of new elements" to the domain of sets of numbers. Suppose that all atomic formulas are unaffected by the introduction of new objects. Then the easiest and most natural way to ensure that our attempted definition of a set is stable is by requiring that all of its quantifiers be restricted to old objects, that is, to sets of natural numbers already introduced. For again, when a condition restricts its attention entirely to the old elements, it cannot be disordered by the introduction of any new ones. But to require that a definition of a new element of a collection restricts all its quantifiers to old elements of the collection is precisely to impose the Vicious Circle Principle.

To sum up, the two aspects of predicativism are connected as follows. The belief that certain domains are merely potential poses a threat to our attempts to define new elements of such domains: will our definitions be stable as more and more elements are added to the domain or might the definitions instead be "disordered" by such

[4] We are here assuming that "pluralities" are rigid, in the sense that a "plurality" doesn't change its members from world to world. See Section 7 for discussion and references.

additions? A safe way to ensure the needed stability is to require that our definitions adhere to the Vicious Circle Principle.[5] Note, however, that on this analysis, the Vicious Circle Principle is merely a means to an end, namely definitional stability. The door is thus left open for other such means.

§3. Potential totalities: a modal analysis. A lot of work is needed to get our Poincaré-inspired analysis out of its nutshell. The most pressing concern is to clarify what it is for a domain to be merely potential and how at least some ordinary mathematics can be understood as concerned with such domains. To keep the paper self-contained, this section provides a brief overview of the general framework for potentiality developed in Linnebo [27] and Linnebo and Shapiro [30]. Readers familiar with the framework can skip ahead to Section 4.

Beginning with Aristotle, and until the nineteenth century, the majority of major philosophers and mathematicians rejected the notion of the actual infinite—a complete, existing entity with infinitely many members. They argued that the only sensible notion is that of potential infinity.

In *Physics* 3.6 (206a27-29), Aristotle wrote, “For generally the infinite is as follows: there is always another and another to be taken. And the thing taken will always be finite, but always different.” As Sorabji [42, pp. 322–323] put it, for Aristotle, “infinity is an extended finitude”(see also [25, 26]). The idea seems to be that the infinite is tied to certain *procedures* that can be repeated indefinitely.

This orientation toward the infinite was endorsed by mainstream mathematicians as late as Gauss [19], who in 1831 wrote: “I protest against the use of infinite magnitude as something completed, which is never permissible in mathematics. Infinity is merely a way of speaking.”

A nice example is provided by Aristotle’s claim, against the ancient atomists, that matter is infinitely divisible. Consider a stick. However many times one has divided the stick, it is always possible to divide it again (or so it is assumed). It is natural to explicate this temporal vocabulary in a modal way. This yields the following analysis of the infinite divisibility of a stick s:

$$\Box\forall x(Pxs \rightarrow \Diamond\exists y\, Pyx), \tag{1}$$

where Pxy means that x is a *proper* part of y. If, by contrast, the divisions of the stick formed an actual infinity, the following would hold:

$$\forall x(Pxs \rightarrow \exists y\, Pyx). \tag{2}$$

According to Aristotle, it is not even possible to complete infinitely many divisions of the stick, that is:

$$\neg\Diamond\forall x(Pxs \rightarrow \exists y\, Pyx). \tag{3}$$

By endorsing both (1) and (3), one is asserting that the divisions of the stick are *merely potentially infinite*.

[5] This tells against Gödel’s [20] widely discussed view that the Vicious Circle Principle is justified only if one is a constructivist: “...it seems that the vicious circle principle in its first form applies only if the entities are constructed by ourselves” (p. 136). On our account, the motivation derives from potentialism, not from a metaphysical view that mathematical objects are literally our constructions.

https://doi.org/10.1017/S1755020321000423 Published online by Cambridge University Press

According to Aristotle, the sequence of natural numbers too is merely potentially infinite. We can represent this view as the conjunction of the following theses:

$$\Box\forall m \Diamond \exists n\, \textsc{Succ}(m, n), \tag{4}$$

$$\neg\Diamond\forall m \exists n\, \textsc{Succ}(m, n), \tag{5}$$

where $\textsc{Succ}(m, n)$ states that n comes right after m. The modal language thus provides a nice way to distinguish the merely potential infinite from the actual infinite.

3.1. Three orientations toward the infinite. It is useful to distinguish orientations toward the infinite. *Actualism* unreservedly accepts actual infinities, and thus finds no use for modal notions in mathematics (or at least no use that is specific to the analysis of infinity). Actualists maintain that the non-modal language of ordinary mathematics is already fully explicit and thus deny that we need a translation into a modal language. Furthermore, actualists accept classical logic when reasoning about the infinite and typically also accept all of classical mathematics.

Potentialism stands opposed to actualism. According to this orientation, at least some of the objects with which mathematics is concerned are generated successively, and some of these generative processes cannot be completed. There is room for disagreement about *which* processes can be completed. An austere Aristotelian form of potentialism takes a very restrictive view, insisting that at any one stage, there are never more than finitely many such objects, but that we always (i.e., necessarily) have the ability to go on and generate more.[6] Generalized forms of potentialism take a more relaxed attitude. Potentialism about set theory provides an extreme example. According to this view, it is impossible to complete the process of forming sets from any objects that are available, but any generative process that is indexed by a set-theoretic ordinal can be completed.[7] As we shall see, Feferman himself falls in between the austere Aristotelian orientation and that of set-theoretic potentialism.

Potentialists can differ with respect to a *qualitative* matter as well. As characterized above, potentialism is the view that *the objects* with which mathematics is concerned are successively generated and that some of these generative processes cannot be completed. What about *the truths* of mathematics? As we see it, on any form of potentialism, these are modal truths concerned with certain generative processes. But how should these truths be understood?

Liberal potentialists regard the modal truths as unproblematic, adopting bivalence for the modal language. Consider Goldbach's conjecture. As potentialists interpret it, the conjecture says that necessarily, any even number greater than 2 that is generated can be written as a sum of two primes. Liberal potentialists maintain that this modal statement has an unproblematic truth-value—it is either true or false. Their approach to modal theorizing in mathematics is thus much like a realist approach to modal theorizing in general: there are objective truths about the relevant modal aspects of reality, and this objectivity warrants the use of some classical form of modal logic.

Strict potentialists differ from their liberal cousins by requiring not only that every object be generated at some stage of a process, but also that every truth be "made true" at some stage. Consider, again, the Goldbach conjecture. If there are counterexamples

6 Recall Sorabji's [42, pp. 322–323] suggestion that, for Aristotle, "infinity is an extended finitude."

7 See Linnebo [27] for an account along these Cantor–Zermelo lines.

to the conjecture, then its negation will presumably be "made true" at the stage where the first counterexample is generated. But suppose there are no counterexamples. Since the conjecture is concerned with *all* the natural numbers, it is far from obvious how it could be "made true" without completing the generation of natural numbers. But at least according to austere Aristotelian potentialism, there can be no such completion.

One response would be for strict potentialists to abandon any attempt at true universal generalizations over an incompletable domain. Indeed, the Vicious Circle Principle can be seen as an important step in that direction: at least when defining new members of an incompletable domain, we must refrain from quantifying over the entire domain and instead take care to restrict all quantifiers to the elements of the domain that *have* been generated. However, Linnebo and Shapiro [30] argues that strict potentialists have another, less restrictive option as well, namely to adopt a modal logic whose underlying logic is intuitionistic (or intermediate between classical and intuitionistic logic). In particular, strict potentialists should not accept every instance of the law of excluded middle in the background modal language. As we shall see toward the end of this paper, this dovetails with a view that Feferman and others adopt toward predicative mathematics, and it has ramifications for the articulation of predicativism and the extent of the mathematics that it captures.

3.2. The modal logic. Our next task is to identify the modal logic to be used in our modal explication of potential infinity. For the time being, we will be neutral on the liberal vs. strict divide and thus also on whether the non-modal part of the logic should be classical or intuitionistic.

It is often useful to invoke the contemporary heuristic of possible worlds, but we understand this as *only* heuristic, as a manner-of-speaking. Our official theory is formulated in the modal language, with (one or both of) the modal operators as primitive. The modal language is rock bottom, not defined in terms of anything else.

To invoke the heuristic, the idea is that a "possible world" has access to other possible worlds that contain objects that have been constructed or generated from those in the first world. From the perspective of the earlier world, the "new" objects in the second world exist only potentially. One sort of construction is geometric: the later world may contain, for example, a bisect of a line segment in the first. Or the later world might contain an extension of a line segment from the first world. Other sorts of constructions are arithmetic: the later world might contain more natural numbers than those of the first, say the successor of the largest natural number in the first world. Or, for a third kind of example, the later world may contain a set whose members are all the sets in the first world.

An Aristotelian assumes that every possible world is finite, in the sense that it contains only finitely many objects. This, of course, just is the rejection of the actually infinite. We make no such assumption here, however. Our goal is to *contrast* the actually infinite and the potentially infinite, so we need a framework where both can occur (to speak loosely). An actual infinity—or, to be precise, the possibility of an actual infinity—is realized at a possible world if it contains infinitely many objects. As will be shown, this is a fruitful background for Feferman's articulation of predicativism.

We also assume that objects are not destroyed in the process of construction or generation. This is in keeping with most ordinary mathematical talk about construction. We construct new objects but never destroy old ones. Suppose, for example, that a given line segment is bisected. Then the resulting "world"contains the two bisects, as well as the original line segment.

To continue the heuristic, it follows from the foregoing that the domains of the possible worlds are non-decreasing along the accessibility relation. So we assume:

$$w_1 \leq w_2 \to D(w_1) \subseteq D(w_2), \tag{6}$$

where '$w_1 \leq w_2$' says that w_2 is accessible from w_1, and for each world w, $D(w)$ is the domain of w. For present purposes, we can think of a possible world as determined completely by the mathematical objects—regions, numbers, sets, etc.—it contains. We will talk neutrally about the extra mathematical objects existing at a world w_2 but not at an "earlier"world w_1 which accesses w_2, as having been "constructed" or "generated."

This motivates the following principle:

> **Partial ordering**: The accessibility relation $\leq$ is a partial order. That is, it is reflexive, transitive, and anti-symmetric.

So the underlying logic is at least S4.

As is well-known, the conditional (6) entails that the converse Barcan formula is valid. That is,

$$\exists x \Diamond \varphi(x) \to \Diamond \exists x \varphi(x). \tag{CBF}$$

At any stage in the process of construction, we generally have a choice of which objects to generate. For example, we might have, at some stage, two intervals that don't yet have bisections. We can choose to bisect one or the other of them, or perhaps to bisect both simultaneously. Assume we are at a world w_0 where we can choose to generate objects, in different ways, so as to arrive at either w_1 or w_2. It makes sense to require that the license to generate a mathematical object is never revoked as our domain expands. The option to bisect a given line segment, for example, can always be exercised at a later stage.

This corresponds to a requirement that any two worlds w_1 and w_2 accessible from a common world have a common extension w_3. This is a directedness property known as *convergence* and formalized as follows:

$$\forall w_0 \forall w_1 \forall w_2 (w_0 \leq w_1 \land w_0 \leq w_2 \to \exists w_3 (w_1 \leq w_3 \,\land\, w_2 \leq w_3)).$$

We therefore adopt the following well-known principle:

> **Convergence**: The accessibility relation $\leq$ is convergent.

This principle ensures that, whenever we have a choice of mathematical objects to generate, the order in which we choose to proceed is irrelevant. Whichever object(s) we choose to generate first, the other(s) can always be generated later. Unless $\leq$ is convergent, our choice whether to extend the ontology of w_0 to that of w_1 or that of w_2 might have an enduring effect.

The mentioned properties of the accessibility relation $\leq$ allow us to identify a modal logic appropriate for studying the generation of mathematical objects. The modal logic S4 is sound with respect to our intended system of possible worlds. As is well known, the convergence of $\leq$ ensures the soundness of a modal principle known as 'G' (or sometimes '.2'):

$$\Diamond \Box \varphi \to \Box \Diamond \varphi. \tag{G}$$

The modal propositional logic that results from adding this principle to a complete axiomatization of S4 is known as S4.2.

https://doi.org/10.1017/S1755020321000423 Published online by Cambridge University Press

3.3. The logic of potential infinity. What is the correct logic when reasoning about potentially infinite collections? Informal glosses aside, the language of contemporary mathematics is strictly non-modal. We thus need a translation to serve as a bridge connecting the non-modal language in which mathematics is ordinarily formulated with the modal language in which our analysis of potential infinity is developed. Suppose we adopt a translation $*$ from a non-modal language $\mathcal{L}$ to a corresponding modal language $\mathcal{L}^{\Diamond}$. The question of the right logic of potential infinity is the question of which entailment relations obtain in $\mathcal{L}$.

To determine whether $\varphi_1, \dots, \varphi_n$ entail ψ, we need to (i) apply the translation and (ii) ask whether $\varphi_1^*, \dots, \varphi_n^*$ entail ψ^* in the modal system. This means that the right logic of potential infinity depends on two factors. First, the logic depends on the bridge that we choose to connect the non-modal language of ordinary mathematics with the modal language in which our analysis of potential infinity is given. Second, the logic obviously depends on our modal analysis of potential infinity; in particular, on the modal logic that is used in this analysis—including the underlying logic of the modal language, whether it is classical or intuitionistic. The second factor has been investigated above. Let us now turn to the first factor.

At the heart of potentialism lies the idea that the existential quantifier of ordinary non-modal mathematics has an implicit modal aspect. When an Aristotelian says that a number has a successor, she really means that it *potentially* has a successor—that it is *possible* to generate a successor. This suggests that the right translation of $\exists$ is $\Diamond\exists$. Since the universal quantifier can hardly be less inclusive in its range than the existential, this suggests that $\forall$ be translated as $\Box\forall$. We are not, however, proposing these translations as a contribution to empirical semantics; rather, they are proposed as a useful rational reconstruction, which explicates an implicit potentialist conception of the language of mathematics. Note also that we treat the two quantifiers separately, since we consider both classical and intuitionistic backgrounds.

Thus understood, the quantifiers of ordinary non-modal mathematics are not restricted to objects that have been generated at a given stage. Rather, they are understood as devices for generalizing over all objects that can be generated. In our modal language, these generalizations are effected by the strings $\Box\forall$ and $\Diamond\exists$. Strictly speaking, these strings are composites of a modal operator and a quantifier. Invoking the heuristic of possible worlds, however, these strings behave like quantifiers ranging over all entities at all (future) worlds. We will therefore refer to the strings as *modalized quantifiers*.

The proposal is thus that each quantifier of the non-modal language is translated as the corresponding modalized quantifier. Each connective is translated as itself. Let us call this the *potentialist translation*, and let $\varphi^{\Diamond}$ represent the translation of φ. We say that a formula is *fully modalized* just in case all of its quantifiers are modalized. Clearly, the potentialist translation of any non-modal formula is fully modalized.

Say that a formula φ is *stable* if the necessitations of the universal closures of the following two conditionals hold:[8]

$$\varphi \to \Box\varphi \qquad\qquad \neg\varphi \to \Box\neg\varphi.$$

[8] Unless stated otherwise, the formulas listed below can contain free variables, and so we assume their universal closures. Sometimes "stable" is called "persistent" or "absolute." This notion has been investigated by model-theoretic tools; see [9] or the earlier announcement [17].

https://doi.org/10.1017/S1755020321000423 Published online by Cambridge University Press

Invoking the heuristic, a formula is stable just in case it never “changes its mind,”in the sense that, if the formula is true (or false) of certain objects at some world, it remains true (or false) of these objects at all “later” worlds as well.

We are now ready to state two key results, which answer the question about the correct logic of potential infinity. Let $\vdash$ be the relation of classical deducibility in a non-modal first-order language $\mathcal{L}$. Let $\mathcal{L}^\Diamond$ be the corresponding modal language, and let $\vdash^\Diamond$ be deducibility in this language by $\vdash$, S4.2, and axioms asserting the stability of all atomic predicates of $\mathcal{L}$.

Theorem 1 (Classical potentialist mirroring). *For any formulas $\varphi_1, \dots, \varphi_n, \psi$ of $\mathcal{L}$, we have*:

$$\varphi_1, \dots, \varphi_n \vdash \psi \quad \textit{iff} \quad \varphi_1^\Diamond, \dots, \varphi_n^\Diamond \vdash^\Diamond \psi^\Diamond.$$

The proof goes by induction, on the complexity of the sentences in one direction and on the length of derivations in the other. A key lemma is that if φ is in the range of the translation, then φ, $\Diamond\varphi$, and $\Box\varphi$ are equivalent.[9]

The theorem has a simple moral. Suppose we are interested in logical relations between formulas in the range of the potentialist translation, in a classical modal theory that includes S4.2 and the mentioned stability axioms. Then we may delete all the modal operators and proceed by the ordinary non-modal logic underlying $\vdash$. This buttresses our choice of the potentialist translation as the bridge connecting actualist and potentialist theories. We will observe, as we go along, that the stability axioms on which the mirroring theorem relies are acceptable.

However, we will often use formulas outside of the range of the potentialist translation. Two examples are provided by the Aristotelian rejection of certain actual infinities, as expressed by (3) and (5). In this way, we use the extra expressive resources afforded by the modal language to engage in reasoning that cannot take place in the corresponding non-modal language, not even when all of its quantifiers are understood as implicitly modalized. The modal language thus allows us to look at the subject matter under a finer resolution, which the mirroring theorems enable us to turn on and off, according to our needs. In what follows, we will encounter many examples of reasoning that makes essential use of this finer resolution, but also ones where the finer resolution isn’t needed and can thus be turned off.

An important upshot of the theorem is that ordinary classical first-order logic is validated via this bridge. However, this response depends on the robustness of our grasp on the modality. We noted that our *liberal* potentialist accepts classical logic when it comes to the modality. Our first mirroring theorem fits in nicely with that perspective. As noted above, however, Linnebo and Shapiro [30] argue that a stricter form of potentialism pushes in the direction of intuitionistic logic. What to do then?

The answer is given by a second mirroring theorem, which we now explain. Say that a formula φ is *decidable* in a given theory if the universal closure of $\varphi \vee \neg\varphi$ is deducible in that theory. Let $\vdash_{\text{int}}$ be the relation of intuitionistic deducibility in a first-order language $\mathcal{L}$, plus the decidability of all atomic formulas of $\mathcal{L}$. Let $\vdash_{\text{int}}^\Diamond$ be deducibility

[9] See [27] for details. Recall that S4.2 has a rule of necessitation, such that if $\vdash P$, then $\vdash \Box P$.

https://doi.org/10.1017/S1755020321000423 Published online by Cambridge University Press

in the modal language corresponding to $\mathcal{L}$, by $\vdash_{\text{int}}$, S4.2, and the stability axioms for all atomic predicates of $\mathcal{L}$.

THEOREM 2 (Intuitionistic potentialist mirroring). *For any formulas $\varphi_1, \dots, \varphi_n, \psi$ of $\mathcal{L}$, we have:*[10]

$$\varphi_1, \dots, \varphi_n \vdash_{\text{int}} \psi \quad \text{iff} \quad \varphi_1^{\Diamond}, \dots, \varphi_n^{\Diamond} \vdash_{\text{int}}^{\Diamond} \psi^{\Diamond}.$$

Together, the two mirroring theorems show how our analysis of quantification over a potentially infinite domain can be separated from the question of whether the appropriate logic is classical or intuitionistic. Hold fixed our modal analysis of potential infinity, the propositional modal logic S4.2, and the potentialist bridge. Then the appropriate logic of potential infinity depends entirely on the first-order logic used in the modal system. Whichever first-order logic we plug in on the modal end—classical or intuitionistic—we also get out on the non-modal end. Since liberal potentialists see no reason to plug in anything other than *classical* first-order logic, they can reasonably regard this as the correct logic for potential infinity.

§4. Predicativism as a form of potentialism. Let us return to our contention that it is useful to understand predicativism as a form of potentialism. We start by explaining the predicativist conception of sets.

It is customary to distinguish between *combinatorial* and *logical* conceptions of sets (see, for example, [31]). According to combinatorial conceptions, a set is characterized by specifying each and every member, which can be chosen arbitrarily, not necessarily in accordance with any rule or principle. Here it is useful to invoke plural logic. A set can then be specified as $\{xx\}$ for some suitable plurality xx, which may be completely arbitrary. Thus, in any circumstance in which the set $\{xx\}$ exists, its members will be precisely xx. By contrast, on a logical conception, a class is characterized in terms of its membership criterion, say, as $\{x : \varphi(x)\}$ for some suitable condition $\varphi(x)$ (which may contain parameters). Thus, in any circumstance in which the class $\{x : \varphi(x)\}$ exists, its members will be precisely the objects satisfying the membership condition $\varphi(x)$, which may change from circumstance to circumstance.

Predicativists are typically hostile to the combinatorial conceptions, preferring instead to develop a version of the logical conception. The reason for this preference should be clear. By rejecting the combinatorial conceptions, predicativists avoid any reliance on the idea of completing an infinite specification of its members, one by one, and can thus endorse a more thoroughgoing potentialism. This line of reasoning is nicely illustrated by the following passage by Weyl:

> The notion of an infinite set as a "gathering" brought together by infinitely many individual arbitrary acts of selection, assembled and surveyed as a whole by consciousness, is nonsensical: "inexhaustibility" is essential to the infinite. [45, p. 23]

[10] The statement of this theorem in [30] is incorrect, but the proof there is correct for the theorem as stated here. Thanks to Michael Rathjen. The proof is similar to that in the classical case, but (as usual) more tedious.

https://doi.org/10.1017/S1755020321000423 Published online by Cambridge University Press

How, then, should predicativists develop the logical conception? Suppose they wish to use a formula $\varphi(x)$ to define a set X.[11] In an ordinary non-modal context, this would require:

$$\forall x(x \in X \leftrightarrow \varphi(x)). \tag{7}$$

But we are currently dealing with a domain of sets that is merely potential. We are explicating this potentiality by means of modal logic. Our explication applies the potentialist translation $\psi \mapsto \psi^\Diamond$ to every formula of the ordinary non-modal language. Thus, instead of requiring (7), we need to consider the strong requirement that a formula *necessarily* defines a set:[12]

$$\Box\forall x(x \in X \leftrightarrow \varphi(x)). \tag{8}$$

Our aim is therefore encapsulated by the following set comprehension principle:

$$\Diamond\exists X\Box\forall x(x \in X \leftrightarrow \varphi(x)). \tag{S-Comp}$$

For which formulas $\varphi(x)$ is this principle acceptable?

As a preliminary observation, we claim that membership has to be stable:

$$x \in X \to \Box(x \in X) \qquad\qquad x \notin X \to \Box x \notin X.$$

We impose this requirement for two reasons. First, this coheres well with our ordinary way of thinking about sets. Once a set is formed, it never "changes its mind" about whether an object that was available at the stage of formation is or is not a member of the set. Second, and more importantly, the assumption of stability of membership is needed for the mirroring theorems to be available, which in turn is responsible for connecting our modal analysis of potentiality with the non-modal language of ordinary mathematics.

Suppose, therefore, that we require membership to be stable. This requirement entails that the defining condition $\varphi(x)$ too has to be stable:

$$\varphi(x) \to \Box\varphi(x) \qquad\qquad \neg\varphi(x) \to \Box\neg\varphi(x).$$

To see this, recall that the condition $\varphi(x)$ must *necessarily* define the set X, as expressed by (8). Thus, if the set X cannot "change its mind" about whether or not an object x is a member, then nor can the condition $\varphi(x)$ "change its mind" about whether or not it applies to x. We therefore conclude that for (S-Comp) to be permissible, the condition $\varphi(x)$ needs to be stable. This yields an important necessary condition for permissibility. We will later articulate sufficient conditions as well.

Notice that our necessary condition echoes Poincaré's emphasis on the importance of ensuring that all definitions of mathematical objects be stable whenever we reason about a domain that is merely potential. Although Poincaré never developed a modal analysis, at least not explicitly, he anticipated one of its key conclusions.

An obvious choice that confronts any account of sets, whether predicativist or not, concerns what objects are eligible to be elements of the sets. A simple option, which will

[11] As noted, the formula φ may contain parameters for numbers and sets, and so we invoke the universal closure of (7).

[12] Of course, the potentialist translation of (7) would have $\varphi^\Diamond(x)$ instead of $\varphi(x)$. We choose to focus on the more general case where the attempted defining condition can be any formula of the modal language, not only formulas in the range of the potentialist translation.

https://doi.org/10.1017/S1755020321000423 Published online by Cambridge University Press

be our main theme here, is to restrict our attention to sets of natural numbers. It will then be natural to adopt a two-sorted formalism with distinct variables for numbers and for sets thereof. More generally, a two-sorted formalism will be appropriate for any approach on which sets are not themselves eligible to be elements of sets.[13]

How should the predicativist approach to sets be developed? We will now identify four such choices, which give rise to different varieties of predicativism. As we will see, our modal analysis sheds new light on these choices.

First, what is the starting point for the generation of sets? With our modal analysis, and engaging our heuristic of possible worlds, the question concerns the choice of a "base world," whose ontology can then be extended in various ways, as represented by all the other possible worlds. According to what we will call *absolute predicativity*, the base world is empty, with the result that the entire ontology must be generated.[14] More familiar is *relative predicativity*, which proceeds from a non-empty base world. Indeed, the most widely studied form of predicativity—which will also be our focus here—is predicativity relative to the natural numbers.[15] Here the base world consists of all the natural numbers.

Second, are sets allowed to "grow" as the generation of the merely potential domain unfolds? As argued above, every set must be stable, in the sense that it cannot "change its mind" about whether an object is or is not an element of the set. But this requirement applies only to objects that have been generated; it leaves open the possibility that as further objects are generated, some of these will "become" elements of the set. To illustrate, consider sets of natural numbers. Suppose that natural numbers too are successively generated, as an Aristotelian would have it. Consider the formulas '$n = n$' and 'n is prime', which seem entirely predicative. If these formulas are allowed to define sets, then the resulting sets would "grow" as new numbers are generated. Let us say that a set is *rigid* if it cannot exhibit this kind of growth. While the requirement that sets be stable is obligatory, predicativists have a choice whether or not to require additionally that sets be rigid.[16] Notice, however, that the choice evaporates when the domain from which the elements are drawn is fixed: for stability applied to a fixed domain implies rigidity.

The two remaining choices are more general and arise for any form of potentialism (cf. Section 3.1). The third choice concerns how thoroughgoing the potentialism is. What kinds of collections can be completed, so as to "fit into" a single possible world (which need not, of course, be the base world)? As observed, an austere Aristotelian

13 On the other hand, if we want to study sets that are permitted to have sets as elements, then a one-sorted formalism will be natural, with a single sort of variables ranging over sets and objects (including sets themselves) eligible to be elements of sets.

14 This sometimes goes by the name of "strict predicativity." We wish to avoid that term, however, because we have already used the word 'strict' as a label for one of the two main forms of potentialism.

15 Poincaré took this relative form of predicativity to be justified, despite his assertion (quoted above) that "[t]here is no actual infinity," because the natural numbers are given to us in intuition.

16 How might such a rigidity requirement be formulated? One option is to use plurals and let a plurality xx record all the potential elements available when a set X was generated. To require that X be rigid is to require that necessarily every element of X be drawn from xx. Another option is to adopt a version of the Barcan Formula for quantification restricted to X, namely $(\forall x \in X)\Box\theta \rightarrow \Box(\forall x \in X)\theta$. This can be shown to capture the rigidity requirement sufficiently at least for proof-theoretic purposes.

potentialist insists that no infinite collection can be completed. At the opposite end of the scale we find a relaxed set-theoretic potentialist, who allows any *set* to be completed, including very large transfinite sets, but insists that the *totality* of sets be regarded as merely potential. (Indeed, on pain of contradiction, our set-theoretic theorist must insist on this, since they hold that any given objects xx potentially define a set $\{xx\}$.)

Predicativism naturally suggests an intermediate option: while the collection of natural numbers can be completed, the collection of *sets of natural numbers* cannot. This view is particularly natural in connection with predicativity relative to the natural numbers, which regards the collection of natural numbers as completed already in the base world, while sets of natural numbers are generated without end and thus belong to an incompletable totality.[17] However, predicativism is interesting also in connection with the two extreme options. Aristotelian potentialists will, for example, have good reason to explore non-rigid (i.e., "growing") sets of numbers. Moreover, set-theoretic potentialists have the interesting option of combining a combinatorial conception of sets with a logical conception of predicative classes. Since the elements of these classes are drawn from a "growing" domain of sets, we will once again have good reason to allow the classes to be non-rigid.

Finally, there is the choice between liberal and strict potentialism. Is it sufficient to require that every object be generated at some stage of a process (as the liberals think), or should we additionally require that every truth be "made true" at some stage (as strict potentialists require)? We observed that the Vicious Circle Principle involves a step in the strict direction. When a definition of a new object is only allowed to quantify over old objects available at the stage where the definition is made, then at least any atomic statement about the new object will be "made true" by that stage. For example, suppose all the natural numbers are available in the base world and that a set Y of numbers is defined by a formula $\varphi(x)$ whose second-order quantifiers are restricted to some sets XX. Then the truth-value of any membership statement $n \in Y$ will be determined by how things stand with the numbers and the given sets XX.

As mentioned, there is also a more radical explication of strict potentialism based on the idea of letting the modal logic be intuitionistic (or semi-intuitionistic) rather than classical. We will have an occasion to consider this more radical explication toward the end of the paper, where we argue that this is in some respects superior to the classical rivals. For the time being, however, we will focus on systems using classical logic, as is fairly standard in discussions of predicativism.

To sum up: as advertised in the introduction, our modal analysis explicates predicativism as a form of potentialism and locates it in a general framework for potentiality, alongside other forms of potentialism. We find this explication illuminating, both in its own right and because of the light that it shines on the different varieties of predicativism.

§5. Some remarks on absolute predicativity. Absolute predicativity, in our sense, lets the base world be empty and thus requires that *all* objects be successively generated.

[17] A variant of predicativity relative to the natural numbers combines absolute predicativity (i.e., letting the base world be empty) with the possibility of completing the generation of the natural numbers. The resulting view is conceptually different from, but mathematically equivalent to, predicativity relative to the natural numbers as characterized here.

Since it is helpful to work with a specific example, let us consider the case of the natural numbers and sets thereof.

Suppose we stipulate that any definition of a set is only allowed to quantify over objects available at the relevant stage. Of course, it is commonplace for predicativists to prohibit quantification over *sets* other than those that have already been generated. But we are now taking a potentialist view of the natural numbers as well. So the mentioned requirement gives rise to an addition prohibition, namely against quantifying over *numbers* other than those that have already been generated. This additional prohibition is stricter than anything we find in, for example, Poincaré, Weyl and Feferman. In practice, the additional prohibition means that only bounded arithmetical formulas (possibly with set parameters) are allowed to define sets. A related approach has been developed and explored by Nelson [32]. It results in a weak form of second-order arithmetic in the vicinity of the system known as $\mathrm{I}\Delta_0$.[18]

Perhaps it is overkill, however, to require that all quantifiers occurring in a definition of a set be restricted to objects already generated. As our analysis in the previous section suggested, predicativity, at least as understood by Poincaré, is first and foremost a matter of definitional stability. By restricting all the quantifiers, even the arithmetical ones, we certainly ensure definitional stability. But perhaps there are other, less draconian ways to achieve the same result. We will now illustrate one way to develop this intriguing line of thought.

Consider liberal potentialist about the natural numbers, who take there to be robust facts about the non-actualized possibilities for generating natural numbers. Liberal potentialists are therefore prepared to engage freely in classical modal reasoning about such possibilities. Let $\varphi(x)$ be a formula of first-order arithmetic, and let $\varphi^\Diamond(x)$ be its potentialist translation. As is easily seen, this modal formula is guaranteed to be stable [27, Lemma 5.3]. Thus, as far as stability is concerned, absolute predicativists who are also liberal potentialists about the natural numbers can accept precisely as much set comprehension as can relative potentialists who accept a completed totality of natural numbers already in the base world. The only difference is that the former theorists construe this as set comprehension on a fully modalized formula $\varphi^\Diamond(x)$, typically resulting in a non-rigid set of numbers, whereas the latter theorists construe this as comprehension on a non-modal formula $\varphi(x)$ applied to a fixed domain of numbers, thus resulting in a rigid set. But in light of the classical mirroring theorem (Theorem 1), these two groups of theorists agree on everything expressed in the non-modal language of ordinary mathematics; the only difference is that the former theorists insist on this language being implicitly modal, whereas the latter theorists take the language at face value.

The upshot, then, is this. Consider our running example predicative sets of natural numbers. Suppose we regard predicativity as first and foremost a matter of definitional stability, rather than a doctrinal adherence to the Vicious Circle Principle. *Then it makes no difference in the ordinary language of non-modal second-order arithmetic whether one is an actualist or a liberal potentialist about the natural numbers.*

[18] This system is based on second-order Peano arithmetic with an axiomatic form of induction and only Δ_0-comprehension for sets of numbers. According to some commentators (e.g., Parsons [33, pp. 304–307]), absolute predicativists are entitled to a bit more, e.g., to exponentiation being total, resulting in a system known as $\mathrm{I}\Delta_0 + \exp$.

https://doi.org/10.1017/S1755020321000423 Published online by Cambridge University Press

As Crosilla and Linnebo argue in [4], this analysis provides an answer to a puzzle about what it means to be a predicativist relative to the natural numbers. Two of the historically most important advocates of this view, Weyl and Feferman, both use classical (non-modal) first-order arithmetic and then introduce predicative sets of numbers on top of that. But they advocate very different views of the natural numbers. As we have seen, Weyl insists that all infinities are merely potential, including the infinity of the natural numbers. By contrast, Feferman [10, p. 619] characterizes the natural numbers as a "completed totality." How can they disagree so fundamentally about the ontological status of the natural numbers while agreeing on the logical and mathematical explication of predicativism relative to the natural numbers? The answer, Crosilla and Linnebo argue, is that Weyl [45] is a liberal potentialist about the natural numbers, whereas Feferman is an actualist—of course, both are potentialists about *sets* of natural numbers. As we have seen, this philosophical difference has no effect on the resulting theory in the non-modal language of ordinary mathematics.

§6. The basics of relative predicativity. Our focus in what follows is on predicativity relative to the natural numbers. As just observed, there are two options. Following Feferman, we may regard the natural numbers as a given completed infinity, not constructed in stages. Alternatively, but equivalently for purely mathematical purposes, we may take a liberal potentialist view of the natural numbers. We will here adopt Feferman's option, since this is technically simpler and thus affords a better focus on what is now our primary concern, namely how *sets* of natural numbers are generated in stages in a way that can never be completed.

To engage our heuristic, we therefore assume that all of the natural numbers exist in each world. Each such world can be regarded as a Henkin model, with the (standard) natural numbers as the range of the first-order variables. The second-order variables range over sets of natural numbers, those sets that exist at that world. So, the domain of sets of natural numbers grows along the accessibility relation (although, again, the first-order domain, consisting of the natural numbers, does not).[19] We can also think of the background as a two-sorted first-order language and will sometimes do so without further comment.

There are three major phases to the development of a predicativist theory of sets of natural numbers. The first phase is where we think of a set of natural numbers as definable, in a given "world," in terms of what other sets exist at that world. The defined set exists in some later, accessible world.

The second phase is where one starts with a given world w and talks about *all* of the sets that are definable in w. That is, we envision another world w' that has all of the sets definable in terms of those that exist in w. We then envision a third world w'' that contains all of the sets definable in w'. Clearly, this iterates: the second-phase procedure results in an ω-sequence of worlds.

Finally, we wish to iterate applications of the second phase into the transfinite. This is the most challenging aspect of our project. If we try to iterate the procedure through *all ordinals* (or all countable ordinals, or all recursive ordinals), we will end up with something more like (a part of) Gödel's constructible hierarchy, and not a predicativist

[19] To avoid using a free logic, we assume that every "world" contains the empty set of numbers: $\Box\exists X\forall x(x \notin X)$. This is at most a small departure from the themes of predicativism.

framework. The underlying predicativist theme is that if, in a given world, we can *prove* (predicatively) that a certain relation is a (countable) well-ordering, then we can iterate the construction through the corresponding ordinal (taking unions at limit stages).

An (alleged) conceptual instability of predicativism comes to the fore in this last phase, since it does not seem possible to give a predicatively acceptable account of a predicatively acceptable proof, or a predicatively acceptable well-ordering, or, indeed, of a well-ordering. Stay tuned.

§7. Phase 1. This phase has its own challenges. Here is a first attempt at the relevant modal comprehension principle. It is a scheme:

$$\Box\forall\vec{y}\forall\vec{Y}\Diamond\exists X\Box\forall x(x \in X \leftrightarrow \varphi(x, \vec{y}, \vec{Y})), \qquad \text{(First Shot)}$$

one instance for each formula φ that does not have X free. It is also important that the embedded formula φ does not contain any modal operators. The idea is that at a given world w, we can define sets (of numbers) with reference to what exists at w (including which sets exist at w), but not with reference to what will be generated later.

This proposal faces two problems. First, as we observed, the comprehension condition $\varphi(x)$ needs to be stable, which isn't guaranteed on (First Shot). The second problem is more subtle. Start at a given world w. The $\Diamond$-operator takes us to an accessible world, say w'. The embedded formula—the instance of comprehension—is then evaluated at w', not at w, as intended. That is, the second-order existential quantifier $\exists X$ in (First Shot) ranges over the sets that exist at w'. So we define a set X by reference to a formula φ that may contain quantifiers whose range includes that very set X. So (First Shot) is impredicative! To solve the second problem, it is not sufficient to talk about a set existing in an accessible w' world: we also need the ability to talk, in that world (so to speak), about what exists only in the opening world w.

One option is to introduce one of the "backward looking" operators proposed (in a different context) by some philosophers (Fine [18] and Williamson [47]). We would like to avoid this conceptual and logical complication and stick to resources of standard modal logic. Another option is to introduce terminology for the levels in the ramified hierarchy. A standard way to do that is to attach a superscript to the set variables, to indicate their level. The comprehension principle would be something like this:

$$\forall\vec{y}\forall\vec{Y}\exists X^n\Box\forall x(x \in X^n \leftrightarrow \varphi(x, \vec{y}, \vec{Y})), \qquad \text{(RF)}$$

one instance for each formula φ that has no bound second-order variables of level n or higher, and no free second-order variables of level higher than n. This, of course, is the route of ramified type theory, dispensing with the modality altogether. As is familiar, it gets messy.

We submit that a better way to solve both of the mentioned problems is to invoke the resources of plural logic, applied to sets of numbers. These resources can either be adopted as a new primitive—which we will do at the outset, for ease of exposition—or else be "coded" by a clever use of sets. The latter option—developed in the next section—is obviously more economical and will eventually be our preferred choice.

We introduce a new style of variable XX, YY, ..., for plural reference to sets, and a relation $X \prec XX$, which is read "X is one of XX."[20] And we take $XX \preccurlyeq YY$ (to

[20] See [2] for a seminal discussion of plural logic and [28] for an overview.

https://doi.org/10.1017/S1755020321000423 Published online by Cambridge University Press

be read as "XX are among YY") as an abbreviation of $\forall X(X \prec XX \rightarrow X \prec YY)$. Continuing the heuristic, we follow [27, 30] in taking pluralities to be *rigid*, in the sense that some given sets XX comprise the very same sets in every world in which they exist. Consequently, the atomic formula $X \prec XX$ is rigid (and thus a fortiori stable). So this use of plural logic has some of the effect of a "backward looking" modal operator.

Of course, we do not want any impredicative comprehension principles for the plural variables. All we seem to need, for now, is the existence, in each world, of a "universal" plurality, one that "contains" every set in that world:[21]

$$\Box\exists XX\forall X(X \prec XX). \tag{Univ}$$

In words: necessarily there are some sets such that every set is one of them. Notice, incidentally, how this formula makes essential use of the "finer resolution" afforded by our modal framework (cf. Section 3.3). For this formula is not the potentialist translation of any non-modal formula.

We now come to the heart of our proposal. Let φ be a formula that does not contain the plural variable XX. Let $\varphi^{<XX}$ be the result of relativizing all second-order quantifiers to XX. That is, replace each occurrence of the form $\forall X \ldots$ or $\exists X \ldots$ with $\forall X(X \prec XX \rightarrow \ldots)$ or $\exists X(X \prec XX \wedge \ldots)$, respectively. Every formula that is relativized in this way is stable, in the usual sense that it never "changes its mind" about whether or not it applies to some numbers as the domain of sets expands. The reason for this stability is that every set quantifier is relativized to a rigid plurality of sets, which ensures that its range is unaffected by the domain expansions.[22] We therefore adopt the following Phase 1 principle:

$$\Box\forall\vec{y}\forall\vec{Y}\forall XX[\Diamond\exists Z\Box\forall x(x \in Z \leftrightarrow \varphi^{<XX}(x, \vec{y}, \vec{Y}))], \tag{P1}$$

one instance for each formula φ that does not contain XX. In words, (P1) says that, necessarily (i.e., in any world w), for any XX's (in w), there could be a set of numbers Z whose members are all and only the numbers that are φ at w.

In fact, every set generated by (P1) is not only stable but also rigid, in the sense that it has the same members at all worlds at which it exists. To verify this, observe first that there is a formula of the form $\varphi^{<XX}$ which necessarily characterizes membership in the set. Next, since this formula is stable and applied to a fixed domain, it is also rigid. Putting things together, it follows that membership in the set too is rigid. Notice that this result relies essentially on our assumption that the natural numbers form an

---

21 Recall that we assume that every world has at least the empty set of numbers. It is harmless to add some other comprehension principles, so long as they are not impredicative. The following scheme seems safe:

$$\Box\forall\vec{Y}\forall\vec{y}[(\exists X\varphi(\vec{Y}, \vec{y}, X) \rightarrow \exists XX\forall X(X \prec XX \leftrightarrow \varphi(\vec{Y}, \vec{y}, X)))],$$

one instance for each formula φ that does not contain any bound higher-order or plural variables. Note that instances of this scheme, including (Univ), are not in the range of the potentialist translation.

22 One of the main results of [9] is a partial converse of this: in a classical background, a formula is stable (in our terminology, or "invariant" in Feferman's) with respect to the class of all worlds (or models) that are "end-extensions" (in the usual model-theoretic sense) if and only if the formula is equivalent to one of the form $\varphi^{<XX}$.

https://doi.org/10.1017/S1755020321000423 Published online by Cambridge University Press

actual or completed infinity, available at each possible world. Without this assumption, membership in a set would be stable but not rigid; for example, the set of even numbers (which is obviously arithmetically definable) would acquire a new member whenever a new even number is generated.

We submit that (P1) gives us what we want. It asserts the possible existence of a set that is definable (via φ) from the sets that exist in a given world. So we see that, if the predicativist accepts the modal framework, then at least this principle can be formulated in an acceptable manner.

§8. Eliminating pluralities in favor of sets. As advertised, it is possible to eliminate pluralities of sets in favor of just sets. To do so, we use a pairing operation $\langle x, y\rangle$ on natural numbers. When we work over arithmetic, this operation is arithmetically definable, say, by letting $2^x \cdot 3^y$ represent $\langle x, y\rangle$. Upper-case variables can thus be used to code sets of ordered pairs as well as sets of numbers. And, of course, the pairing function enables us to handle n-tuples as well, for arbitrary finite n.

If X is a set of numbers and n is a number, we define *the n-section of X*, denoted X_n, as $\{x | \langle n, x\rangle \in X\}$. This definition enables us to use a single set to represent any countable plurality of sets. To see this, consider any countable plurality of sets, say, the X_n's for n any natural number. This plurality can be represented by the single set X whose n-section, for each n, is precisely X_n.

Of course, the big question is whether this strategy for representing pluralities of sets suffices for our purposes. We contend that it does. Notice, first, that our construction of sets can be carried so as to ensure that, at every world, there is a single set whose sections represent all of the other sets at that world:[23]

$$\Box\exists X\forall Y(Y \neq X \rightarrow \exists n(Y = X_n)). \tag{9}$$

If true, this contention means that every world contains at most a countable infinity of sets, which can be specified as a single set X and all of its sections X_n.

To defend our contention, the key is to observe that, whenever new sets are generated and added to the domain to obtain a more populous possible world, this extension is systematic in a way that enables us to represent all of the sets then available by means of a single set—provided that all of the sets previously available could be so represented. This holds for our principle (P1), as well as the relevant principles from Phase 2 and Phase 3, to be introduced below. Each new set is always defined by a specific formula. Since these are the only ways to add new sets, we know that our strategy for representing sets will always suffice.

Instead of relativizing a formula to an arbitrary plurality of sets, we can relativize it to the plurality of sets represented by a single set. Let φ be a formula without any occurrences of 'X'. We wish to let $\varphi^{<X}$ be the result of restricting the set-quantifiers in φ to the sets represented by X via its sections. We achieve this by translating $\forall Y\psi(Y)$ as

$$\forall Y\forall z(Y = X_z \rightarrow \psi^{<X}(Y)),$$

[23] Note that this formula is not in the image of the potentialist translation, as it is not fully modalized.

https://doi.org/10.1017/S1755020321000423 Published online by Cambridge University Press

where z is a new first-order variable, and by translating $\exists Y\ \psi(Y)$ in the obvious dual manner:[24]

$$\exists Y \exists z (Y = X_z \wedge \psi^{<X}(Y)).$$

We are now ready to formulate a purely set-theoretic formulation of our Phase 1 principle, namely:

$$\Box \forall \vec{y} \forall \vec{Y} \forall X \Diamond \exists Z \Box \forall x (x \in Z \leftrightarrow \varphi^{<X}(x, \vec{y}, \vec{Y})), \tag{P1$'$}$$

one instance for each formula φ that does not contain X. Notice that all the formulas we here use to define sets are stable, as required (cf. Section 4). Since sets are rigid, so are their sections. So restricting a formula to the sections of a set has the same effect as restricting it to a plurality. In particular, $\varphi^{<X}$ is stable, since its first-order variables range over the natural numbers, which exist at all worlds, and its second-order variables are restricted to sets that exist in all accessible worlds.

Before proceeding to our purely set-theoretic treatment of Phases 2 and 3 of predicativity relative to the natural numbers, we wish to make some remarks about the predicativist understanding of Cantor's theorem. From a classical point of view, the theorem is taken to show that the natural numbers are outnumbered by the sets of natural numbers: while there are only countably many such numbers, there are uncountably many such sets. As we will now see, although predicativists too can prove the theorem, they draw a rather different lesson from it, namely that the domain of sets is incapable of enumeration and "inherently potential" (as Feferman put it in the quote with which we began).[25]

Let us run through the reasoning behind Cantor's theorem. Suppose for *reductio* that there is an enumeration of all of the sets of natural numbers. That is, there is a surjective function f from the natural numbers to the sets of natural numbers. As usual, we define a diagonal set Δ by means of the formula '$n \notin f(n)$'. It follows that Δ cannot be a member of the enumeration effected by f. Since the formula '$n \notin f(n)$' is entirely predicative, predicativists too can conclude that there cannot be an enumeration of all the sets of natural numbers.

Let us make this reasoning fully explicit. First, we add the function symbol 'f' to our object language. It is natural also to add an axiom stating that for every set X, there is a function f that takes n to X_n. Second, we need to spell out the implicit modality. When we assume that f is surjective, we assume that every *possible* set is in its image:

$$\Box \forall X \exists n (f(n) = X).$$

We then invoke the fact that it is *possible* to use the formula '$n \notin f(n)$', which has no second-order quantifiers, to define a set:

$$\Diamond \exists Y \forall n (n \in Y \leftrightarrow n \notin f(n)).$$

24 We include a separate clause for the existential quantifier to keep the door open for an intuitionistic background language.

25 For the predicativist pioneers' own response to Cantor's theorem, see Poincaré [35, Section 6] and [34], as well as Weyl [45, Section 5] and [46, Section 2]. Thanks to Crosilla for helping us find these references.

By a modal variant of the standard Cantorian reasoning, we thus prove that there can be no function from the natural numbers onto all possible sets of such numbers. In short, the merely potential domain of all possible sets of numbers cannot be enumerated.

This object language statement must of course not be conflated with the metalinguistic claim that in any Kripke model of our theory, the possible sets of numbers cannot be enumerated. From the point of view of classical mathematics, that claim is false, since our theory has Kripke models with countably many worlds each of which has countably many sets. Thus, again using classical mathematics, there is an enumeration f of all the possible sets. The apparent conflict is removed by observing that this enumeration f is not predicatively acceptable; in particular, its existence cannot be established by the axiom mentioned above.

Let us return to the object language. We have shown that it is impossible to enumerate all possible sets of natural numbers. What *is* possible, however, is to enumerate *all the sets that are available at some given world*. To prove this, we first use (9) to find a single set X such that all the sets (at the given world) are X and its sections X_n. Next, using (P1′), it is possible to define a set Y such that $Y_0 = X$ and $Y_{i+i} = X_i$ for each i. We can thus obtain a single set Y that enumerates all the sets that are available at the given world, as desired. As usual, the talk about worlds is merely an eliminable heuristic. Our result can be stated in a purely modal language. If we retain plurals, the result admits of a particularly nice statement:

> Necessarily, given any sets XX, possibly there exists a single set Y whose sections are precisely XX.

Or, even more compactly:

> Necessarily, any given sets are possibly enumerated.

Even without plurals, the result can be expressed purely modally, albeit somewhat more long-winded, given (9):

> Necessarily for every X, if every set is either identical with X or one of X's sections, then possibly there exists a single set Y whose sections are precisely X and each of X's sections.

Let us take stock. We have shown, first, that it is impossible to enumerate *all possible* sets of numbers, and second, that it is possible to enumerate *any given* such sets.

To complete our account, it remains only to connect these two claims and draw the conclusion that all possible sets of numbers cannot be simultaneously given. That is, given any sets of numbers, it is possible for there to be yet more such sets:

$$\Box\forall XX \Diamond\exists Y(Y \not\prec XX). \tag{10}$$

We find this plural statement a particularly attractive way to express Feferman's claim that the domain of sets of natural numbers is incompletable or merely potential. In the presence of our other assumptions about plurals, (10) ensures that no Kripke model can have a single world containing all the possible (predicative) sets of numbers. It is also possible to express the incompletability claim without recourse to plurals:

$$\Box\forall X\forall Y(Y \neq X \rightarrow \exists n(Y = X_n)) \rightarrow \Diamond\exists Y(Y \neq X \wedge \forall n(Y \neq X_n)). \tag{11}$$

In words: given any set X such that any other set is one of its sections, it is possible for there to be a set Y that is distinct from X and each of its sections.[26]

Unsurprisingly, both these statements of the incompletability claim, the plural one and the singular one, admit of a direct proof as well. Consider any set X such that any other set is one of its sections. (By (9), any sets of numbers can be specified in this way.) As observed, it is possible for there to be a single set Y whose sections are precisely X and each of X's sections. We now diagonalize on this set Y. That is, we use (P1′) to establish the possible existence of a set Δ such that $\forall n(n \in \Delta \leftrightarrow n \notin Y_n)$. Since Δ cannot be identical with one of Y's sections, it follows that Δ is distinct from X and each of X's sections, precisely as desired.

§9. Phase 2. Now on to our second phase, where we want to gather together *all* of the sets that are definable in a given world into a single world. Informally, we begin with a world w that only contains the empty set. We want to assert the existence of a world w', accessible from w, that has all of the sets of numbers that are definable in w via something like (P1).

Start with a world w. Notice that, in light of (P1) and the convergence principle for the modal logic S4.2, we already have that for any *finite* collection of sets definable in w, there is a world w' that has all of the sets in this collection. So one way to proceed would be to add a generalization of the convergence principle to the system, namely that for any (countable) set W of worlds, all accessible from a single world, there is a world w' such that every member of W has access to w'. This, however, would make the possible worlds framework more than a mere heuristic. We would be formulating a principle with respect to the space of worlds, and using that to generate more sets of numbers.[27]

Our first suggestion is to introduce terminology for satisfaction. For convenience, we introduce a multi-arity predicate (or a series of predicates) $\mathrm{SAT}(m, \vec{y}, \vec{Y})$ which says, in effect, that $\vec{y}$ and $\vec{Y}$ satisfy the formula with code m (with the variables in a canonical order).[28] To stay (far) away from paradox, we restrict attention to those formulas that do not have the SAT predicate.

It seems that the predicativist herself does invoke something like satisfaction, at least informally, in the move to Phase 2. How else are we to understand phrases like "every set definable by such and such resources"? In *Principia*-style predicativism, for every n, we have an axiom *scheme* ensuring that every formula whose second-order quantifiers are

26 Notice that in the present modal framework, and assuming that the background logic is classical, we can sanction the potentiality of predicative sets only by invoking an axiom (like (9)) that is not in the range of the modal translation (see also note 21). Indeed, if every axiom is in the range of the modal translation, then the (first) mirroring theorem would entail that the theory is, in some strong sense, equivalent to the theory obtained by erasing all modal operators. And any model of that theory can be converted into a one-world Kripke model of the modalized theory. In effect, the single world would contain all sets of numbers that are generated. Thanks to Sam Roberts for pressing us to get clear on this.

27 It would be better if we had instead a modal principle, analogous to Principle G above (namely $\Diamond\Box\varphi \to \Box\Diamond\varphi$) which had that effect. We suspect there is no such principle. Even if we did manage to add such a principle, we would be more in the neighborhood of Gödelian constructibility, not predicativism.

28 We could, however, get by with a fixed arity by using pairing, as explained in the previous section.

restricted to order n defines a class of order $n+1$. Intuitively, there is a single thought behind all these instances of the axiom scheme. To make this single thought explicit, we need SAT or something like it (but see below). So if our satisfaction predicate is not acceptable to a predicativist, then we have an instance of the aforementioned instability in the position.

If z is the code of a formula φ that has x free, let $f(z)$ be the code of $\varphi^{<X}$, the result of restricting the higher-order (singular and plural) variables to the sections of X.

We are now prepared to formulate the relevant Phase 2 generation principle. It is a kind of quantified "semantic ascent" of (P1):[29]

$$\Box\forall\vec{y}\,\forall\vec{Y}\forall X\Diamond\exists W\forall z\forall x(\langle z,x\rangle\in W\leftrightarrow SAT(f(z),x,\vec{y},\vec{Y})). \tag{P2}$$

Call the resulting principle (Phase 2). It says that, in any world w, for any set X in w, there could be a set that has, as sections, all of the sets that are definable from the sections of X. Notice that this formula does not require the "finer resolution" afforded by our modal framework (cf. Section 2.3). For this formula is equivalent to the potentialist translation of a non-modal formula.[30]

The resources we have developed in Section 8 enable us to avoid the introduction of a satisfaction predicate, and this, as we will see, facilitates Phase 3. We characterize the *predicative jump* of a set by means of the following axiom scheme:

$$J(X,Y)\rightarrow\forall x\forall\vec{y}(\langle\ulcorner\varphi\urcorner,x,\vec{y}\rangle\in Y\leftrightarrow\varphi^{<X}(x,\vec{y})), \tag{Pred Jump}$$

where φ can be any formula without bound second-order variables and with no free variables other than $\vec{y}$ and X.[31] Thus, when Y is the predicative jump of X, then Y has sections corresponding to any set of numbers that is definable with a single second-order parameter X (and first-order parameters $\vec{y}$).

We contend that J can be assumed to be stable. Whether or not $J(X,Y)$ is true at some stage is a matter of whether or not, at this stage, Y codes all the sets that are predicatively definable from X. But since the statement that Y codes all the sets that are predicatively definable from X uses only bounded set quantification, it is independent of the particular stage at which it is evaluated.

---

29 As in note 28, this principle can be captured by a single axiom, rather than an axiom scheme. If one wants to maintain the use of plurals, our principle is this:

$$\Box\forall\vec{y}\,\forall\vec{Y}\forall XX\,[\forall Z(Z\prec XX)\rightarrow\Diamond\forall z\,\exists Z\,\forall x(x\in Z\leftrightarrow SAT(g(z),x,\vec{y},\vec{Y}))], \tag{P2$'$}$$

where, if z is the code of a formula φ that has x free, then $g(z)$ is the code of $\varphi^{<XX}$, the result of restricting the higher-order quantifiers to XX. In words, (P2) says that there is an accessible world w' with every set that is definable in w.

30 For this reason, and because the generation of sets, from Phase 2 onward, can be taken to proceed in a well-ordered manner, it might (at least for technical purposes) be possible to replace our modal explication with a stage-theoretic one where the stages are represented by ordinals. This might make for a connection between the present approach and that of Feferman [8] and Schütte [40]. We hope to explore this in future work.

31 By adding further axioms, we can obtain a *unique* characterization of Y relative to X, at least in ω-models. Conditional on $J(X,Y)$, we add, first, that Y consists solely of triples, and second, that any of these triples has as its first coordinate the Gödel number of a suitable formula. In the relevant cases, this can be used to provide a finite characterization of the predicative jump relative to the set whose jump it is. In fact, under some weak assumptions, the theory of predicative classes of some basic objects is known to admit of a finite axiomatization; see [23].

We can now further streamline our second phase with an axiom to the effect that predicative jumps potentially exist:

$$\Box\forall X \Diamond\exists Y\, J(X, Y). \tag{P2$''$}$$

§10. Phase 3. In the final, third phase we consider transfinite iterations of the predicative jump. This phase poses some entirely new problems.

As a warm-up, let us begin with finite jumps. For each natural number n, we wish to be able to talk about the nth predicative jump of a given set of numbers. Recall that Y_i is the ith section of Y, i.e., $Y_i = \{n \mid \langle i, n\rangle \in Y\}$. Let $JSeq(X, Y, n)$ abbreviate the formalization of the following:

$Y_0 = X$ and for each $i < n$, $J(Y_i, Y_{i+1})$.

We can formulate an axiom asserting the possible existence of n-fold iterations of the predicative jump:

$$\Box\forall X\forall n\Diamond\exists Y\, JSeq(X, Y, n).$$

Next, what we might call *stage* ω can be obtained as follows:

$$\Box\forall X\Diamond\exists Z(\forall n\exists Y\, JSeq(X, Y, n) \wedge Y_n = Z_n). \tag{Stage ω}$$

For each set X, (Stage ω) gives us a world with a set Z of numbers whose nth section is the nth jump of X, for each natural number n.

So far, so good. But this is surely not the limit of how "far" the predicativist can go. The purpose of the third phase of predicativism is to express the possibility of iterations of the jump operation along any (predicatively definable) *well-ordering*. Here we encounter one of the most interesting, and subtle, issues behind the enterprise—and we run up against the charge that predicativism is unstable. Recall that a relation is a well-ordering on, say, the natural numbers, just in case every non-empty set of numbers has a least element in that ordering. But, for the predicativist, there is no stage at which all of the sets of numbers exist. So the very notion of a well-ordering is impredicative! In a retrospective article, Feferman [10, p. 606] himself complains of some previous proposals that

> they ignored one crucial point if predicativity is only to take the natural numbers for granted as a completed totality: namely, that they involve in an essential way ... the impredicative notion of being a well-ordering relation.

Let us examine the complaint in the context of our modal analysis of predicativism as a form of potentialism. The present modal framework certainly seems to provide a way for the predicativist to express the notion of a well-ordering. The modalized quantifiers provide a way to talk of all sets of numbers, whenever they are generated. As we have seen, predicativists rely on this kind of talk in order to state their position, for example, to express that no matter how far the generation of sets has proceeded, it is possible to continue by further steps corresponding to Phase 1 or 2.

Using these expressive resources, let $WO(R)$ abbreviate the formalization of the following:

- R is stable.
- R is a set of ordered pairs of numbers.
- Necessarily for any X, if X is not empty, then there is a number n such that n is the R-least member of X.

We can now define the notion of a jump sequence along some well-ordering R, formalized as $JSeq'(X, Y, R)$, as follows:

- $WO(R)$,
- $Y_0 = X$ and, for each n,
  - if n is the R-successor of m, then $J(Y_m, Y_n)$,
  - if n occupies a limit position in R, then $(Y_n)_m = Y_m$ for each m that precedes n in the ordering R.

This, in turn, enables us to formulate an axiom asserting the possible existence of transfinite iterations of the predicative jump along any well-ordering R:

$$\Box\forall X \forall R(WO(R) \rightarrow \Diamond\exists Y\, JSeq'(X, Y, R)). \tag{P3}$$

In words, (P3) says that if R is a well-ordering, then there is an iteration of predicative jumps along R.[32]

However, as the passage by Feferman reminds us, an objection looms. Our principle (P3) avails itself of the impredicative notion of R being a well-ordering in order to characterize what sets can be generated by predicative means. As observed, *some* use of modalized quantifiers has to be acceptable in general in order to enable the predicativist to state her position, and to prove key theorems. The present concern is that reliance on modalized quantifiers *in axioms with existential import* might inadvertently admit impredicative reasoning.

It is instructive, in this connection, to compare our previous (P1) and (P2$''$) with the present (P3). The former two principles have the form "$\Box\forall\Diamond\exists$." Despite the (modalized) quantification over all sets, the predicativist is able to make good sense of these principles: both involve some predicatively acceptable operation which necessarily, given any numbers and sets as input, generates some further set that satisfies a certain condition. Predicativists can understand this as involving only a form of free-variable-based generalization over sets: given any numbers and sets that might become available, the operation can be applied to generate yet another set. Thus, no completed or determinate totality of sets of natural numbers is presupposed.

By contrast, (P3) has the form "$\Box\forall(WO(R) \rightarrow \Diamond\exists \ldots)$." This logical form is fundamentally different. (P3) says we have some predicatively acceptable operation which necessarily, given any input that *satisfies some impredicative condition*, generates something or other. This principle cannot be understood as involving only free-variable-based generalization over sets: for the italicized part of the antecedent makes essential use of (modalized) quantification over sets. Thus, there is a fundamental

32 Adding this axiom yields a modal analogue of a system known as ATR_0. Thanks to Stephen Simpson for pointing this out.

difference between (P1) and (P2″), on the one hand, and (P3), on the other. While the former two principles can be seen as involving only free-variable-based generality over the merely potential totality of sets of natural numbers, the third cannot. This difference means that (P3) is problematic, from a predicativist point of view, in a way that the first two principles are not.

How, then, can predicativists develop an acceptable version of Phase 3? We will consider two options: a more traditional one based on classical logic and a more radical one based on a semi-intuitionistic logic.

The first option is to try to find a version of Phase 3 that gets by with only the kind of free variable reasoning about sets that predicativists license anyway. To see how this might be possible, observe that predicativists *can* sometimes prove that a given relation R is a well-ordering. The proof can be given using only parameters, or free variables, ranging over sets. Consider the standard ordering $<$ on the natural numbers. Let X be any non-empty set of numbers (i.e., one that gets generated somehow—at some stage). Suppose that X has no least element. Consider the set Y of all numbers x such that x is not in X and neither is anything smaller than x. This definition is predicative because it is first-order (with X as a parameter). We now proceed by induction to show that every number is in Y. We establish the base case, 0 is in Y, by observing that 0 would otherwise be in X, which would thus have a least element. For the induction step, suppose n is in Y. Then n is not in X, and neither is anything smaller. If $n + 1$ were not in Y, either $n + 1$ is in X or something smaller is in X. The latter is ruled out by the induction hypothesis. So $n + 1$ is in X and is the smallest member of X. Contradiction. So, by induction, all numbers are in Y. So no number is in X, which contradicts our assumption that X is non-empty.

We contend that this proof does not presuppose the existence of a completed totality of all sets of natural numbers.[33] Using only free variable reasoning, we show that any set X, whenever it is generated, has a least element. In an analogous way, predicativists can show that $\omega + 1$ is a well-ordering, and so is $\omega + 2$, ..., $\omega \cdot 2$, and so on, for a while, anyway. Exactly how far is a delicate matter.

These considerations suggest the possibility of a rule-based version of Phase 3, along the following lines.

> Suppose we can prove, by predicatively acceptable means, that R is a well-ordering of the natural numbers. Then it is possible to iterate J along R.[34]

Of course, this suggestion presupposes that we can articulate precisely what counts as a proof by "predicatively acceptable means." We need to specify a deductive system, or a series of deductive systems, for the enterprise. Here it would be natural to follow

---

33 The proof is non-constructive, however, relying on the law of excluded middle (via the reductio reasoning). The least number principle fails in constructive mathematics. But the predicativist has no problem with classical logic, at least when it is restricted to first-order formulas (see below).

34 In fact, there are at least two ways of implementing this strategy. One is to restrict the background language to one that has only free second-order variables, or, equivalently, to restrict the language to Π^1_1-sentences (as in [41, Section 9.3], for example). The other way is to allow the full language, but to restrict axioms and rules to formulas with only free second-order variables (or to Π^1_1-sentences). In that framework, we could formulate our Phase 3 principle as a rule of inference, as above.

the lead of Feferman (and Schütte [40, Part C]), who develop a system of ramified type theory, with variables X^α indexed by predicatively acceptable ordinals. Feferman [10] writes:

> The crucial new point ... is that the predicative ordinals not only are those that can be *defined* by (what happen to be) well-ordering relations in the given systems, but also must previously be *proved* to be such relations. The problem is how to meet this requirement without unrestricted second-order quantification; the answer comes from the provability condition ...

As noted above, however, this would introduce variables indexed by predicatively acceptable ordinals, and this brings back some of the notational nightmare of ramified type theory.

A second, more radical option is to respond to the prevailing objection by turning to intuitionistic logic. As observed, predicativists are entitled to free-variable reasoning about sets. Perhaps they are also entitled to *intuitionistic* quantification over sets since—unlike classical quantification—this doesn't presuppose a completed or determinate totality of sets. Claims of this sort are familiar from Dummett [6, p. 319]. In more recent work, Feferman too expresses sympathy for the idea. In an unpublished, but widely circulated and discussed manuscript [13, p. 23], he writes:[35]

> One way of saying of a statement φ that it is definite is that it is true or false; ...that's the same as saying that the Law of Excluded Middle (LEM) holds of φ, i.e., one has $\varphi \vee \neg\varphi$. Since LEM is rejected in intuitionistic logic as a basic principle, that suggests the slogan, "What"s definite is the domain of classical logic, what's not is that of intuitionistic logic." In the case of predicativity, this would lead us to consider systems in which quantification over the natural numbers is governed by classical logic while only intuitionistic logic may be used to treat quantification over sets of natural numbers or sets more generally.

Based on these ideas, Feferman formulates a "semi-intuitionistic" system, whose axioms we will describe shortly.

First, however, we wish to reflect on his slogan that "What's definite is the domain of classical logic, what's not is that of intuitionistic logic." How might such a claim be substantiated? One option is to invoke the so-called Brouwer–Heyting–Kolmogorov (BHK) readings of the logical operators. (P3) would then be interpreted as something like "there is a way to transform any proof of $WO(R)$ into a proof that J can be iterated along R." We would thus avoid the need for a determinate totality of sets, as desired.[36] Another option is to develop a realizability interpretation, on which (P3) would say that there is an effective way to transform a realizer of $WO(R)$ into a realizer

35 See Rathjen [37] for an interesting discussion of Feferman's proposal.

36 It might be noted that the (BHK) semantics itself has an impredicative feel to it, since it talks about proofs about proofs (see [43]).

of iterability of J along R.[37] By contrast, since the natural numbers are assumed to be complete, quantification over them can still be assumed to behave classically.

More generally, what we need is a distinction between instance-based and non-instance-based explanations of universal generalizations. Let us explain. Suppose that all of your students are born on a Monday. The only explanation, we may assume, proceeds via each of your students. This explanation is highly instance-based. Other generalizations permit an entirely non-instance-based explanation. For example, to explain why every whale is a mammal, there is no need to invoke any particular whale; it suffices to observe that it is part of the nature of whales that they are mammals. When a domain is indeterminate, instance-based explanations won't always be available, since there may be no determinate totality of instances to consider. Linnebo [29] develops the intuitive distinction between instance-based and non-instance-based explanations and finds a connection with the logic that is validated. If every true universal generalization permits an instance-based explanation, then the logic is classical; but if not, then only intuitionistic logic (or semi-intuitionistic logic, of which more shortly) is validated. If correct, this supports the mentioned claim that intuitionistic, but not classical quantification is permitted over an indeterminate domain.

Our distinction between liberal and strict potentialism (Section 3.1) is also relevant in this connection. The former takes the modalized quantifiers to be unproblematic, and applies classical logic to the background modal language. By contrast, the latter requires that any truth about the system must somehow be "made true" at some stage. As observed, this requirement poses a challenge to any universal generalization over a merely potential domain. How can such a generalization be "made true" at some stage, given that many of its instances haven't even been generated? The idea of non-instance-based generality provides an attractive answer: the generalization may well be "made true" independently of all of its instances, thus freeing strict potentialists from a requirement they cannot satisfy, namely that all the instances be available to underwrite the generalization. As Feferman observes, however, this approach to generality naturally gives rise to intuitionistic logic (or some slight strengthening thereof). In our modal setting, this means that strict potentialism is naturally developed by adopting an intuitionistic modal logic.[38]

Returning to our main line of argument, the second option for responding to Feferman's objection to the naive version of Phase 3 is that predicativists should be *strict* potentialists about sets of numbers and accordingly adopt an intuitionistic modal logic. In this modified context, they can accept the axiomatic statement of Phase 3, namely (P3), precisely as stated.[39]

[37] See Kleene [22] for a classic exposition of a realizability interpretation of first-order arithmetic; see Linnebo and Shapiro [30] for an application to strict potentialism about first-order arithmetic. Of course, second-order arithmetic would require a different form of realizability.

[38] See [30] for a more detailed articulation of this claim in the context of arithmetical potentialism.

[39] As observed in note 31, adding (P3) in the context of classical logic yields a modal analogue of the system ATR_0. Following Feferman, we are proposing that strict potentialists are entitled to add (P3) in the context of their semi-intuitionistic logic. This yields a modal analogue of a semi-intuitionistic variant of ATR_0. If correct, this is important. For it is known that the fully intuitionistic variant of ATR_0 has the same proof-theoretic strength as ATR_0 with classical logic (see Crosilla and Rathjen [5], esp. Corollary 9.13), with the proof-theoretic

With this explication and defense of Feferman's slogan in place, we turn to the precise formulation of the relevant system. We begin with an intuitionistic modal logic S4.2, as above. As before, we use a two-sorted system, with one sort of variable ranging over natural numbers and the other ranging over sets of natural numbers. We add the closure of all instances of excluded middle in which the embedded formula contains no second-order quantifiers and is not modal. That is, we add each instance of:

$$\forall \vec{X} \forall \vec{x} (\varphi(\vec{X}, \vec{x}) \vee \neg\varphi(\vec{X}, \vec{x})),$$

where φ contains no bound set-variables.

Notice that this axiom schema allows the aforementioned proof that $<$ is a well-ordering, and so $WO(\omega)$, $WO(\omega + 1)$, ..., $WO(\omega \cdot 2)$, ... By contrast, it is well-known that one cannot prove the least-number principle in a completely intuitionistic setting. Moreover, the axiom schema ensures that an essential assumption of the intuitionistic mirroring theorem is satisfied, namely the decidability of each atomic predicate.

Following Feferman, we might also add further principles such as that of Bounded Omniscience:

$$\forall n(\varphi(n) \vee \neg\varphi(n)) \ \rightarrow\ \forall n\, \varphi(n) \vee \exists n \neg\varphi(n). \tag{BOM}$$

The idea is that quantification over the natural numbers behaves classically because the natural numbers form a "definite" domain. One might attempt to explicate this idea by means solely of the consequent of (BOM), namely $\forall n\, \varphi(n) \vee \exists n \neg\varphi(n)$. But that would be too quick: non-classical behavior might sneak in via the formula $\varphi(n)$ that is generalized. The correct explication is that, *conditional on the classical behavior of* $\varphi(n)$ *for each instance*—that is, conditional on the antecedent of (BOM)—the consequent holds.

To sum up our discussion thus far, when faced with the objection that Phase 3 illegitimately assumes the impredicative notion of a well-ordering, predicativists have two options. One is to adopt a rule-based version which handles the notion of well-ordering by means of free set variables only and stick with classical logic. Another option is to weaken the logic for quantification over sets to intuitionistic logic and permit (P3) as formulated.

We wish to end by observing how we are now in a position to acquit predicativism of a well-known charge that it is incoherent. Hellman [21, p. 299] expresses the charge well:

> the predicativist implicitly transcends predicativity him/herself in the very formulation of the limitative theses!

What are we to make of this?

Hellman is obviously right that it can be useful to "transcend" predicativism in favor of a stronger, classical approach in order to study the limits of predicative definability. Feferman and Schütte show this limit to be a certain countable ordinal Γ_0 (see [40,

ordinal Γ_0. Thus, if our Feferman-inspired analysis is right, strict potentialists are entitled to a form of predicativism relative to the natural numbers that takes us all the way to Γ_0 in a single deductive system. See Feferman [14, p. 87] as well as [11, Sections 4.1 and 4.3].

Part C]).[40] Predicativists cannot themselves recognize that result, since this would require a predicatively coherent definition of "predicatively definable."

Nevertheless, we submit that our two versions of Phase 3 can be used to show that predicativism is internally coherent, what Feferman and Hellman call "stable" [15, 16]. The modal framework can be used to state all the principles and closure conditions that the approach requires, in a way that predicativists themselves can endorse.

Notice, finally, that predicativism is different, in an interesting way, from many other forms of potentialism, especially of the liberal kind, such as Aristotelian or Cantorian. For most of these, we are able to state the generating principles once and for all. In most cases, the principles are just the modalized counterparts to the axioms of the given theory (arithmetic, set theory, etc.). In predicativism, by contrast, the generating principles are, in a sense, themselves generated. As we generate more and more sets of natural numbers, we also have more and more resources that can be used to prove that various numbers code recursive well-orderings, and thus we generate more and more instances of (P3) that can be then used to generate more sets of numbers.

§11. Concluding summary. Since we have come a long way, a brief summary may be useful. We opened by asking about the connection between two seemingly completely different aspects of predicativism: an adherence to the Vicious Circle Principle and the view that certain totalities, such as that of sets of natural numbers, are incompletable or merely potential. Taking a cue from Poincaré, we proposed a simple answer. Potentialism poses a challenge to the stability of our definitions. Might not a definition be disrupted by the stepwise generation of the domains with which they are concerned? We observed that adherence to the Vicious Circle Principle provides a simple and effective response: when all its quantifiers are restricted to objects already generated, a definition is immune to disruption by any future generation of further objects.

Inspired by this simple answer, we presented a general framework for analyzing potentiality in mathematics. We located predicativism in this general framework in a way we find illuminating, as it reveals various choices confronting predicativists and locates the resulting options in a larger landscape of different forms of potentialism.

Using our analysis of predicativism as a form of potentialism, we have also tried to shed light on some important questions concerning predicativity. First, we presented the predicativist take on the logical conception of set, highlighting a simple modal explication of the stability requirement: to define a set, a formula must never be allowed to "change its mind" on whether or not it applies to an object. Second, we solved a puzzle about predicativism relative to the natural numbers, explaining how, at least for logical and mathematical purposes, it doesn't matter whether one is an actualist or a liberal potentialist about the natural numbers. Third, we used the potentialism to explain the predicativist understanding of Cantor's theorem. Fourth, we presented Feferman's semi-intuitionistic approach to predicativity, arguing that this is usefully regarded as a form of strict potentialism. Finally, we observed that our explication of

40 Weaver [44] argues against this, claiming that ordinals much higher than Γ_0 are acceptable to a predicativist. We are neutral on that dispute. Weaver's claims are based on detailed analyses of no less than 10 rather complex deductive systems used in the literature, and we do not know how any of them relate to the present proposal. We do note that Weaver himself uses the language of potentiality, at least informally, and he argues that "intuitionistic logic [is] the appropriate tool for general predicative reasoning."

predicativism is acceptable to predicativists themselves, thus rebutting the charge that the view cannot coherently be stated.

Acknowledgments. Thanks to Laura Crosilla, Peter Fritz, Michael Rathjen, Sam Roberts, Wilfried Sieg, and three anonymous referees, as well as audiences in Leeds and Oslo for discussion and valuable comments on this material. We gratefully acknowledge funding from the Research Council of Norway (grant number 250654), which facilitated our work on this article.

BIBLIOGRAPHY

[1] Aristotle. (1941). *The Basic Works of Aristotle*. New York: Random House; edited by McKeon, R.

[2] Boolos, G. (1984). To be is to be a value of a variable (or to be some values of some variables). *Journal of Philosophy*, **81**, 430–450.

[3] Crosilla, L. M. (2016). Constructivity and Predicativity: Philosophical Foundations. Ph.D. Dissertation, University of Leeds.

[4] Crosilla, L. M., & Linnebo, Ø. (2020). *Two Kinds of Potential Domains*. Unpublished manuscript.

[5] Crosilla, L., & Rathjen, M. (2002). Inaccessible set axioms may have little consistency strength. *Annals of Pure and Applied Logic*, **115**, 33–70.

[6] Dummett, M. (1991). *Frege: Philosophy of Mathematics*. Cambridge: Harvard University Press.

[7] Ewald, W., editor. (1996). *From Kant to Hilbert: A Sourcebook in the Foundations of Mathematics*, Vol. 2. Oxford: Oxford University Press.

[8] Feferman, S. (1964). Systems of predicative analysis. *The Journal of Symbolic Logic*, **29**, 1–30.

[9] ——— (1968). Persistent and invariant formulas for outer extensions. *Compositio Mathematica*, **20**, 29–52.

[10] ——— (2006). Predicativity. In Shapiro, S., editor. *The Oxford Handbook of Philosophy of Mathematics and Logic*. Oxford: Oxford University Press, pp. 590–624.

[11] ——— (2009). Conceptions of the continuum. *Intellectica*, **51**, 169–189.

[12] ——— (2010). On the strength of some semi-constructive theories. In Feferman, S., and Sieg, W., editors. *Proofs, Categories and Computations: Essays in Honor of Grigori Mints*. London: College Publications, pp. 109–129.

[13] ——— (2012). *Is the Continuum Hypothesis a Definite Mathematical Problem*. Available from: http://logic.harvard.edu/EFI_Feferman_IsCHdefinite.pdf.

[14] ——— (2014). Logic, mathematics, and conceptual structuralism. In Rush, P., editor. *The Metaphysics of Logic*. Cambridge: Cambridge University Press, pp. 72–92.

[15] Feferman, S., & Hellman, G. (1995). Predicative foundations of arithmetic. *Journal of Philosophical Logic*, **24**, 1–17.

[16] ——— (2021). Challenges to predicative foundations of arithmetic. In Sher, G., and Tieszen, R., editors. *Between Logic and Intuition: Essays in Honor of Charles Parsons*. Cambridge: Cambridge University Press, pp. 117–138.

[17] Feferman, S., & Kreisel, G. (1966). Persistent and invariant formulas relative to theories of higher order (research announcement). *Bulletin of the American Mathematical Society*, **72**, 480–485.

[18] Fine, K. (1977). Prior on the construction of possible worlds and instants. In Prior, A., and Fine, K., editors. *Worlds, Times and Selves*. London: Duckworth, pp. 116–168.

[19] Gauss, K. F. (1831). Briefwechsel mit Schumacher. *Werke*, **8**, 216.

[20] Gödel, K. (1951). Russell's mathematical logic. In Schlipp, P. A., editor. *The Philosophy of Bertrand Russell* (third edition). Evanston: Northwestern University Press, pp. 123–153; (1983). In Gödel, K. (1990). *Collected Works*, Vol. 2. New York: Oxford University Press, pp. 119–143. Edited by Feferman, S., Dawson, J. W. Jr., Kleene, S. C., Moore, G. H., Solovay, R. M., and van Heijenoort, J.

[21] Hellman, G. (2004). Predicativism as a philosophical position. *Revue Internationale de Philosophie*, **3**, 295–312.

[22] Kleene, S. C. (1945). On the interpretation of intuitionistic number theory. *The Journal of Symbolic Logic*, **10**, 109–124.

[23] Krajewski, S. (1974). Predicative expansions of axiomatic theories. *Zeitschrift für mathematische Logic und Grundlagen der Mathematik*, **20**, 435–452.

[24] Kreisel, G. (1960). La prédicativité. *Bulletin de la Societé Mathématique de France*, **88**, 371–391.

[25] Lear, J. (1980). Aristotelian infinity. *Proceedings of the Aristotelian Society*, **80**, 187–210.

[26] ——— (1982). Aristotle's philosophy of mathematics. *Philosophical Review*, **41**, 161–192.

[27] Linnebo, Ø. (2013). The potential hierarchy of sets. *The Review of Symbolic Logic*, **6**, 205–228.

[28] ——— (2004). Plural quantification. *Stanford Internet Encyclopedia of Philosophy.* Available from: https://plato.stanford.edu/entries/plural-quant/.

[29] ——— (forthcoming). Generality explained. *Journal of Philosophy*.

[30] Linnebo, Ø., & Shapiro, S. (2019). Actual and potential infinity. *Noûs*, **53**, 160–191.

[31] Maddy, P. (1983). Proper classes. *The Journal of Symbolic Logic*, **48**, 113–139.

[32] Nelson, E. (1986). *Predicative Arithmetic*. Princeton: Princeton University Press.

[33] Parsons, C. (2008). *Mathematical Thought and Its Objects*. Cambridge: Cambridge University Press.

[34] Poincaré, H. (1906). Les mathématiques et la logique. *Revue de Métaphysique et de Morale*, **14**, 294–317; translated as Mathematics and logic III. In Ewald, W., editor (1996). *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, Vol. 2. Oxford: Oxford University Press, pp. 1052–1071.

[35] ——— (1908). La Logique de l'infini. *Revue de Métaphysique et Morale*, **17**, 461– 482; translated as The logic of infinity. In Bolduc, J.W., editor (1963). Henri Poincaré, *Mathematics and Science: Last Essays*, pp. 45 –64.

[36] Rathjen, M. (2016). Indefiniteness in semi-intuitionistic set theories: On a conjecture of Feferman. *The Journal of Symbolic Logic*, **81**(2), 742–754.

[37] ——— (2019). The scope of Feferman's semi-intuitionistic set theories and his second conjecture. *Indagationes Mathematicae*, **30**, 500–525.

[38] Russell, B. (1908). Mathematical logic as based on a theory of types. *American Journal of Mathematics*, **30**, 222–262.

[39] ——— (1973). *Essays in Analysis*. London: George Allen and Unwin.

[40] Schütte, K. (1977). *Proof Theory*. Berlin: Springer; translated by Crossley, J. N.

[41] Shapiro, S. (1991). *Foundations Without Foundationalism: A Case for Second-Order Logic*. Oxford: Oxford University Press.

[42] Sorabji, R. (2006). *Time, Creation and the Continuum: Theories in Antiquity and the Early Middle Ages*. Chicago: University of Chicago Press.

[43] Van Atten, M. (2017). *Predicativity and Parametric Polymorphism of Brouwerian Implication*. Available from: https://arxiv.org/pdf/1710.07704.pdf.

[44] Weaver, N. (2005). *Predicativity Beyond* Γ_0. Available from: https://www.math.wustl.edu/nweaver/gamma0.pdf.

[45] Weyl, H., Kontinuum, D., & Leipzig, V. (1918). *The Continuum: A Critical Examination of the Foundations of Analysis*. New York: Dover; translated by Pollard, S., and Boyle, T.

[46] Weyl, H. (1921). Über die neue Grundlagenkrise der Mathematik. *Mathematische Zeitschrift*, **10**, 39–79.

[47] Williamson, T. (2013). *Modal Logic as Metaphysics*. Oxford: Oxford University Press.

[48] Zermelo, E. (1930). Über Grenzzahlen und Mengenbereiche: Neue Untersuchungen über die Grundlagen der Mengenlehre. *Fundamenta Mathematicae*, **16**, 29–47; translated as On boundary numbers and domains of sets: New investigations in the foundations of set theory. In Ewald, W., editor. *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, Vol. 2. Oxford: Oxford University Press, pp. 1219–1233.

DEPARTMENT OF PHILOSOPHY, IFIKK
UNIVERSITY OF OSLO
OSLO, NORWAY
E-mail: oystein.linnebo@ifikk.uio.no

DEPARTMENT OF PHILOSOPHY
THE OHIO STATE UNIVERSITY
COLUMBUS, OH, USA
E-mail: shapiro.4@osu.edu

https://doi.org/10.1017/S1755020321000423 Published online by Cambridge University Press