Truth in Mathematics

Edited by

H. G. Dales

Department of Pure Mathematics
University of Leeds

and

G. Oliveri

Wolfson College
University of Oxford

CLARENDON PRESS • OXFORD

1998

17

Two conceptions of natural number

Alexander George and Daniel J. Velleman

The distinction between the completed and the potential infinite is well known. Less noted is a corresponding contrast between two different conceptions of natural number. It is only to be expected that there would be such a contrast, since the natural numbers form our most basic model of an infinite collection. In this note, we present these two distinct conceptions by articulating the philosophical visions that inspire them and the mathematical definitions that give them substance. We show how these analyses satisfy, in interestingly different ways, the basic demands that any such definition must meet. In keeping with the fundamental difference in perspective between these accounts of natural number, we should expect that those who advance the one definition will find the other wanting. We try to describe what form these respective criticisms will take and to say why they will appear misguided to proponents of the conception against which they are directed.

An intuitive way to try to characterize the natural numbers is to use the mathematical idea of the closure of a set under an operation. If A is a set and f is an operation, then A is said to be *closed* under f if, for every object a in A, the result of applying the operation f to a, denoted $f(a)$, is also in A. The *closure* of A under f is the smallest set containing A that is closed under f. For example, in this terminology the set $\mathbb{N}$ of natural numbers would be the closure of the set $\{0\}$ under the successor operation S.[1]

There are two ways that mathematicians commonly form the closure of a set A under an operation f. The first is to begin with the set A, and add additional elements to form the closure. For example, if a is in A, then $f(a)$ must be added to A if we are to obtain a set that is closed under f. But then $f(f(a))$ must also be added, and then $f(f(f(a)))$. In fact, anything that can be obtained by applying the operation f repeatedly to elements of A must be in the closure of A under f. Let A_* be the set of all elements of A, together with those objects obtainable by applying f repeatedly to elements of A. Then A_* is closed under f, and therefore it is the closure of A under f.

Another way to form the closure of A under f is to let A^* be the intersection of all sets that contain A and are closed under f. In other words, the elements of A^* are those objects that have the property of belonging to every set that contains A and is closed under f. It is not hard to see that A^* contains A and is closed under f, and as the intersection of all sets with this property it must be the smallest such set. Thus A^* is also the closure of A under f, and therefore $A^* = A_*$. (See, for example, (Enderton 1972, pp. 22–5).)

These two ways of forming the closure of a set under an operation suggest two ways of trying to characterize the natural numbers. If we let $A = \{0\}$ and f be the successor operation, then they correspond to the definitions we have given of A_* and A^*. The first characterization says that the natural numbers are just those objects obtainable from 0 by repeatedly applying the successor operation. We might regard this characterization as giving two rules for generating natural numbers:

(1) 0 is a natural number, and

(2) If n is a natural number, then so is $S(n)$.

The natural numbers, according to this characterization, are those, and only those, objects that are generated by these rules, so it is natural to call it the *build up* (BU) definition of "ℕ". The restriction that only objects generated by rules (1) and (2) are numbers is often referred to as the *extremal clause* of the definition.

The second characterization says that the natural numbers are precisely those objects that belong to every set that contains 0 and is closed under the successor operation. This characterization starts with sets that contain 0 and are closed under successor, most of which are larger than the set of natural numbers, and then eliminates the non-numbers from these sets by intersecting them. It is therefore appropriate to call it the *pare down* (PD) approach to defining the natural numbers.

The pare down definition of the natural numbers was first advanced independently by Richard Dedekind and Gottlob Frege in the nineteenth century. Note that the validity of mathematical induction is easily seen to follow from the PD definition. For if a predicate holds of 0, and it holds of $S(n)$ whenever it holds of n, then its extension is a set containing 0 and closed under S. Since any natural number, according to the PD definition, must belong to every set containing 0 and closed under S, it follows that the predicate in question must apply to every natural number.

It is also apparent that the PD definition captures all and only the natural numbers. "All" because each natural number belongs to every set that contains 0 and is closed under the successor operation; and "only" because anything that belongs to every such set will also belong to ℕ, since ℕ is itself such a set. But note that this reasoning will apply only if we reckon the set of natural numbers to be in the range of the second-order quantifier in the PD definition. This impredicativity may lead to concern. If one views the definition as offering a recipe for constructing the set of natural numbers, impredicativity is fatal: for it would then have us creating the collection of natural numbers through appeal to that very collection.

But this is clearly not the purpose of the definition according to those who offer it. Its object is not to create the collection of natural numbers, which, on the contrary, is viewed as already existing, but rather to identify which of the existing completed collections of objects ℕ is. Viewed as picking out the already existing collection ℕ, it is, as W.V. Quine has remarked, "not visibly more vicious than singling out an individual as the most typical Yale man on the basis of averages of Yale scores including his own" (1969, p. 243).

It should be clear that the PD definition of "ℕ" is most natural from a platonist perspective. For the PD approach is at home with a conception of the natural numbers as a completed infinite collection that exists, independently of our activity, amidst other likewise completed infinite sets. On this view, the task of an adequate definition is to pick out the set ℕ from this universe of entities.[2]

But now consider the matter from a constructivist standpoint. There is no circumscribed domain of sets "out there" in advance of our activity. The domain of sets is indefinitely extensible; that is, given any particular delimitation of the universe of sets, we can construct in terms of it another set not previously in the universe. In particular, the set of natural numbers will not exist until we construct it. Now the impredicativity of the PD definition is a serious problem. For given this definition, how can one show that, to cite a famous example, Julius Caesar is not a natural number? Only if there exists a set containing 0 and closed under successor that does not contain Caesar. The argument that ℕ itself is such a set will not satisfy the constructivist, since ℕ cannot be assumed to exist before it is constructed.

The impredicativity of the PD definition means that it cannot be viewed as a procedure for the construction of ℕ. The build up definition, however, can be. The BU approach coheres best with a constructivist stance, according to which a definition of "ℕ" should provide us with an account of how to generate all and only the natural numbers. For the constructivist, the only exclusionary clause that is required is one to the effect that only those objects generated by the two specified rules of construction are natural numbers. Once we know this, we can see, for example, that Julius Caesar is not a natural number, for he is not identical to the output of either rule.[3]

But just as the PD approach appears problematic from the constructivist perspective, so does the BU approach appear wanting to the platonist. Yes, the platonist will grant, the BU definition excludes the use of other than the intended rules in generating the members of ℕ, but it does not explicitly exclude unintended uses of those rules. In particular, there is nothing in the BU definition that bars non-finite iteration of the second generating rule. And such iteration must be ruled out, the objection continues, for otherwise there is no guarantee that the definition will capture only the natural numbers, and nothing else.[4] The complaint is not that those who offer the BU definition fail to realize that only finite iteration is permitted, but rather that this realization is no thanks to the definition.

For this reason, the BU definition will appear at best elliptical from the platonist perspective: it must be understood that the second rule of the definition permits only *finite* iteration of the successor operation to yield natural numbers. Of course, if "finite iteration" means "iteration n times, for some natural number n," then the definition is circular, as Dedekind himself objected:

> If one presupposes knowledge of the sequence N of natural numbers and, accordingly, allows himself the use of the language of arithmetic, then, of course, he has an easy time of it. He need only say: an element n belongs to the sequence N if and only if, starting with the element 1 and counting on and on steadfastly, that is, going through a finite number of iterations of the mapping φ [...], I actually reach the element n at some time; by this procedure, however, I shall never reach an element t outside of the sequence N. But this way of characterizing the distinction between those elements t that are to be ejected [...] and those elements n that alone are to remain is surely quite useless for our purpose; it would, after all, contain the most pernicious and obvious kind of vicious circle. The mere words "finally get there at some time," of course, will not do either; they would be of no more use than, say, the words "karam sipo tatura," which I invent at this instant without giving them any clearly defined meaning. (Dedekind 1967, pp. 100–1)[5]

Consequently, the platonist might offer the following as a friendly amendment to the BU proposal: delete the extremal clause, which perforce will be either inexplicit or circular, and secure its intent by specifying that induction is valid. For the platonist, the requirement that induction is a valid means of forming generalizations about the elements of some collection guarantees that non-numbers will be excluded from the collection. For the predicate "natural number" applies to 0, and applies to $S(n)$ if it applies to n, and therefore it must apply to every element of a collection for which induction is valid. Of course, this entailment presupposes that "natural number" (or a predicate of a piece with it) is taken to be well-defined: the validity of induction does not articulate the intentions of the extremal clause unless induction is understood impredicatively, as a generalization over a pre-existing domain of predicates that includes the very predicate being defined. As we have observed, this impredicativity is unacceptable from the constructivist point of view, and therefore the constructivist will not consider induction to be an adequate replacement for the extremal clause.[6, 7]

In fact, the constructivist will find the friendly amendment not only unhelpful, but unnecessary as well. For just as the constructivist's objection to the PD approach appears off the mark to the platonist, so too will the objection about BU's inexplicitness seem to the constructivist. From the constructivist viewpoint, no intelligible but unwanted possibility has yet been described that would require changes to, or replacement of, the extremal clause: since the constructivist accepts the potential infinite, but not the completed infinite, the idea of applying the generation rules an infinite number of times is unintelligible from the constructivist perspective, and so nothing need be said to rule out those alleged entities that would be constructed as a result of such an impossible application.[8, 9, 10]

If BU is not to postulate induction, how then is its validity to be secured?

The argument traditionally offered by constructivists is just this. Consider a predicate P for which the premises of induction hold and let n be any given natural number. The second premise of induction tells us that if P holds of 0 then P holds of $S(0)$. Taken together with the first, which states that P does hold of 0, we can conclude that P holds of $S(0)$, by modus ponens. Since we know, again by the second premise, that if P holds of $S(0)$ then it holds also of $S(S(0))$, we can likewise infer that P holds of $S(S(0))$. And so on. Thus, we see that at every stage of a construction that begins with 0 and proceeds by repeatedly applying the successor operation, P must hold of the object constructed. But on the BU view, n was obtained by precisely such a construction (this is what the extremal clause asserts), so we may conclude that P holds of n. Thus, induction is valid with respect to any well-defined predicate.[11]

Earlier, we noted that the impredicativity of the PD definition of natural number renders it unacceptable to the constructivist, who will turn instead to the BU account for an adequate analysis. This turn will clearly be rewarding for those who believe that in assessing a definition's impredicativity, it suffices to confine one's examination to that definition. For, as we have seen, the source of the impredicativity is the second-order quantifier in the principle of mathematical induction, and induction is not actually a part of the BU definition, but is rather a consequence of it.

But someone might think that the assessment of a definition's impredicativity cannot proceed through scrutiny of it alone, but also requires examination of conceptual truths about the defined notion, in this case the validity of induction.[12] Or someone might be convinced that a definition of "ℕ" cannot merely specify the extension of "natural number" but must also specify the grounds for generalizations about natural numbers; even if induction is not used for the first task (say, by appealing to the BU definition), it is needed for the second, and hence reference to it will have to be made in any adequate analysis of the meaning of "natural number."[13] While we do not here endorse these, or similar, proposals, they do make it worthwhile to inquire whether impredicativity lurks in the BU account of induction and, if it does, whether it is of the same nature as that encountered in the PD conception.

In fact, there is something in that account about which one might worry. According to the constructivist, the domain of predicates to which induction applies is indefinitely extensible. Indeed, the very act of defining "ℕ" extends this domain, by creating the new predicate "natural number," along with other predicates defined in terms of it. Induction ought to apply to these new predicates. We have argued that the BU definition implies the validity of induction, but might it imply it only for predicates previously defined? If so, then the BU definition would in a sense undermine itself: it would justify induction for all previously defined predicates, and it would simultaneously create new predicates, thus rendering obsolete the very version of induction that it implies.

This same self-sabotage is avoided in the PD definition at precisely the cost of impredicativity. For this definition picks out a set ℕ, which then gives us the new predicate "natural number" and other predicates defined in terms of it, and these new predicates have the potential to undermine the work done previously by the PD definition, since their extensions should have been included among the collections that were intersected to produce $\mathbb{N}$ in the first place. Of course, from the platonist perspective, they were indeed included, since they existed all along, and so all is well. But a constructivist cannot likewise argue that the BU justification of induction goes through because all well-defined predicates already exist quite independently of mathematical activity, for this is precisely what a constructivist denies. How, then, to respond to this concern about BU's justification of induction?

We believe that Charles Parsons suggests the answer when he notes that:

> the principle [of induction] refers to arbitrary predicates, without any assumptions having been made about what counts as a predicate. Like the principles of predicate logic itself, we have a purely formal generalization about predicates, which is not a generalization over a given *domain* of entities and could not be, since it is not determined what predicates will or can be constructed and understood. (Parsons 1992, p. 143)

Our understanding of Parsons' insight is as follows. Often when one makes a generalization—say, that everything in some domain D has property P—one justifies it by examining the objects in D. The most straightforward case would be when one examines the elements of D one by one to see if they have property P (as may happen when the domain D is finite). If we take this as our model, we might be inclined to say that one cannot arrive at a generalization about the elements of a domain D until one knows what is in the domain. But this is not always true. The intuitionistic understanding of "every" illustrates another possibility: even if D is indefinitely extensible, one might arrive at the generalization that everything in D has property P by examining P, not the elements of D, and realizing that something about P makes it true of anything that we would allow into D, even if we do not yet know what is in D. This is precisely what happens with induction, on the BU view. According to the latter, we believe induction applies to all predicates, not because we have surveyed the available predicates and noted that the induction principle applies to all of them (this procedure would indeed lead to the feared self-undermining), but rather, as we saw above, because of the extremal clause, which implies that induction will apply to any predicate, even predicates not yet constructed. Thus, one should not suppose that a grasp of the whole domain of predicates is needed in order to understand why induction holds for all predicates. If this justification still involves impredicativity, then it is an attenuated impredicativity that should be distinguished from that present in the PD approach.

It may well be that an analysis of natural number that succeeds in justifying the principle of induction will have either to be impredicative or to interpret constructively the principle's second-order quantifier. Someone who rejects these two options will find it difficult not to reject the induction principle in its full generality.[14] In any event, one advantage of sharply distinguishing, as we have, between the two accounts of natural number is that doing so enables one to see that the oft-repeated claim that all definitions of natural number are impredicative elides interesting differences.[15]

We would like to conclude, however, by stressing an important affinity between these two approaches, one that has perhaps not always been recognized. In the case of both the PD and the BU definitions of natural number, something has to be in place in order for them to be taken as intended by someone trying to learn the meaning of "natural number." In the first case, the learner must understand the second-order quantifier as ranging over a pre-existing completed totality of sets, including the set $\mathbb{N}$ itself. And in the second, the learner must grasp the concept of finite iteration (perhaps because, as we saw constructivists would insist, this is the only kind of iteration that is intelligible to him). Without this stage-setting, the definitions cannot be understood as intended by their respective proponents.

It is important to notice that what must be in place is, in each case, akin to a grasp of the very notion being defined. These definitions are not circular, but taking them in the intended ways does presuppose some understanding of the very concepts being defined. Let us call such definitions *elucidations*.[16]

We do not want to say that elucidations must fail to convey the target concepts to someone who does not already possess the relevant understanding. After all, the BU and PD elucidations, as a matter of fact, often do help students to acquire the defined notions. Rather, our point is that such elucidations cannot convey these notions to someone who lacks them *through being understood as intended*, for these definitions cannot be so understood except by someone who already grasps in essence the notions being defined.

Elucidations that succeed in conveying an understanding of a term are comparable to speech to infants that facilitates acquisition of language. For such talk likewise does not convey knowledge of language through its being understood as intended—if it were so understood, there would be no need to convey this knowledge.

We will not speculate here regarding how such learning is accomplished. We do not know of any reason, though, for distinguishing between the BU and PD conceptions as regards their conveyability. In the first place, the conceptions behind the BU and the PD definitions are, of course, both infinitistic, and as such neither can be exhaustively displayed by any observable stretch of human behavior.

Secondly, if a learner lacks the relevant conceptions (say because they are not given as part of his innate conceptual endowment), then, as just noted, he cannot gain either of them by taking in their definitions as intended, for so understanding them requires a grasp of something akin to those very conceptions. And if under these circumstances a learner can nevertheless somehow work his way to the target conceptions by taking in their definitions as other than intended, then, pending further information regarding how this takes place, we cannot say that the one conception is any more easily arrived at than the other.

Finally, if appeal to innate notions is made, there is no *a priori* reason why the one conception should be natively given to us and not the other. It might be tempting to argue that there is such a reason, namely that the PD conception is not intelligible and so not there to be given to us, whereas the BU conception is. But if the argument for unintelligibility is ultimately grounded (as it often is) on considerations of acquisition, we are plainly moving in a circle.

In conclusion, though we have been at pains to show that the above two conceptions of natural number do indeed differ in significant ways, we cannot say with any confidence that they do so in point of conveyability.[17]

Notes

1. We are not concerned here about exactly how 0 and the successor operation are defined. One could, for example, use von Neumann's set-theoretic definitions $0 = \emptyset$ and $S(x) = x \cup \{x\}$, or one could regard numbers as strings of strokes and take 0 to be the empty string and the successor operation to be the operation of adding one more stroke to a string. The discussion in the rest of the paper would apply equally to either definition.

2. Henri Poincaré seems to have been the first to note the link between a PD approach to the natural numbers and a commitment to the completed infinite; see sections VIII and XI of "The Last Efforts of the Logisticians" in (Poincaré 1952). Because Poincaré held that "*There is no actual infinity*" (p. 195, original italics), he also rejected PD-type definitions of the natural numbers.

3. Although we have emphasized the strong connection between, on the one hand, PD and platonism, and, on the other, BU and constructivism, note that we have taken no position here regarding whether PD requires a platonist perspective, or BU a constructivist one.

4. For example, if numbers are taken to be strings of strokes, we must ensure that infinitely long strings of strokes are excluded from $\mathbb{N}$.

5. The circularity becomes even more apparent if we try to formalize in set theory the build-up method of forming the closure of a set A under an operation f. The usual approach is to define recursively a sequence of sets $A_0, A_1, A_2, \ldots$ by letting $A_0 = A$ and, for each n, $A_{n+1} = \{ f(x) : x \in A_n \}$. The set A_* can then be defined to be the union of all sets A_n. Of course, our sequence of sets is indexed by the natural numbers, so it would be circular to use this method (with $A = \{0\}$ and $f = S$) to define the natural numbers.

6. Poincaré may be the earliest proponent of the BU approach conscious of the distinction between it and the PD perspective. According to one definition, Poincaré says, "*a finite whole number is that which can be obtained by successive additions, and which is such that n is not equal to n − 1*," while the other holds that, as he puts it, "*a whole number is that about which we can reason by recurrence*." Poincaré continues: "The two definitions are not identical. They are equivalent, no doubt, but they are so by virtue of an *a priori* synthetic judgment: we cannot pass from one to the other by purely logical processes" ("The New Logics," reprinted in (Poincaré 1952, p. 173), original italics). The "*a priori*

synthetic judgment" here is the validity of mathematical induction, for, as we shall see shortly, inferring that induction holds from the BU definition requires the use of induction itself. Because the second definition of natural number (essentially, the PD account) is unacceptable to Poincaré (see note 2 above), he concludes that the validity of mathematical induction cannot be established by purely logical means from any adequate account of natural number and, hence, that the logicist reduction of arithmetic fails.

7. Someone disturbed by a perceived inexplicitness in the BU definition might alternatively offer the following amendment: specify that when the second generation rule is iterated, the set of steps in the iteration must be Dedekind finite, where a set is said to be *Dedekind finite* just in case it cannot be mapped one to one onto any proper subset of itself. The amended definition is not circular, for it does not employ the notion of "finite" or "natural number," and it avoids the use of induction in securing the effect of BU's extremal clause. Yet, this proposal would likewise be rejected by a proponent of BU for it continues to involve an impredicativity. To say that a set is Dedekind finite is to say that there does not exist a function of a certain kind. This claim therefore involves quantification over all functions, including those defined in terms of the natural numbers. (For example, this is the reason why Solomon Feferman and Geoffrey Hellman (1995) chose not to take this approach; see their note 3 and page 15.)

8. Michael Dummett, for example, seems to suggest that no replacement for the extremal clause is needed:

> Even if we can give no formal characterisation which will definitely exclude all such elements, it is evident that there is not in fact any possibility of anyone's taking any object, not described (directly or indirectly) as attainable from 0 by iteration of the successor operation, to be a natural number. (Dummett 1978, p. 193)

9. Even those who accept the completed infinite can defend the extremal clause of the BU definition against the criticisms of the platonist, if they are willing to accept the concept of finiteness as being understood in advance of the characterization of the natural numbers, and then to use this concept to express the extremal clause. This seems to be the standpoint taken by Feferman and Hellman (1995). Their approach, in effect, is to prove the existence of structures satisfying Peano's axioms by constructing an example of one. The universe of their example is defined to be the set of all those objects x such that there exists a finite set containing precisely the predecessors of x under iteration of the operation S, with 0 being the only one of these predecessors that does not itself have a predecessor. This finite set could be thought of as recording the process of constructing x by a finite iteration of the successor operation, beginning with 0, and thus this definition could be thought of as a version of the BU definition. Note that it is important that the recording set be finite so as to ensure that the iteration is finite. (For their analysis, see the first line of their proof of Theorem 7 on p. 10, and their definition of "Fin" on p. 4. The requirement that the recording set be finite is enforced in their formal system by the axiom

(Card), which guarantees that the set is Dedekind finite. This axiom is used to prove that the induction axiom holds in their example. For some discussion of the history of this proposal, and further elaboration on the relationship between their definition and the BU definition, see the following note.)

Feferman and Hellman call their approach "predicativism", or "predicative logicism", and contrast it with classical logicism as follows:

> Classical logicism provides a complete analysis of the concepts "finite", "infinite", and "cardinal number", but at the price of *impredicative comprehension* with all of its attendant "metaphysical" commitments. Predicativism avoids the latter but must presuppose the concept of "finite" in some form or other. However, [...] it can do this in a natural way *without thereby taking the natural number system as given.* (p. 15)

The fact that predicativism must presuppose the concept "finite" will make it unacceptable to anyone who believes that this concept is as much in need of analysis as the concept "natural number". As Daniel Isaacson (1987) suggests, the predicativist definition will be successful only if (i) the second-order quantifier in the definition ranges over a domain that includes all finite initial segments of $\mathbb{N}$, and (ii) the domain contains no infinite sets. He concludes that the definition therefore "does not fare significantly better on the score of avoiding impredicativity than the one based on full second-order logic" (p. 156). Feferman and Hellman (1995, note 5, p. 16) argue in response that the existence of the required finite initial segments can be justified predicatively, but it seems to us that they have failed to answer part (ii) of Isaacson's objection, namely that infinite sets must be excluded from the domain of quantification. As we saw earlier, it is this exclusion of infinite sets from the second-order domain that guarantees that Feferman and Hellman's definition will capture *only* natural numbers. In fact the difficulty here is in effect the same as the difficulty that the platonist finds with the BU definition; it is not the inclusion of the desired elements in the domain that causes problems, but rather the exclusion of unwanted elements.

Charles Parsons also considers a similar definition and finds it wanting for the same reason:

> As a defense of the claim that induction on natural numbers is after all predicative, this exercise is hardly impressive. What has been assumed about finite sets will just reinforce the reply that although perhaps one can escape the impredicativity of induction on natural numbers, one merely throws the matter back to the notion of finite set, where the same problems will arise. (Parsons 1992, p. 148)

10. Feferman and Hellman (1995, note 5), say that their approach "realizes in effect a suggestion attributed to Michael Dummett". This appears to raise a problem for our analysis, according to which Feferman and Hellman's proposal is to be reckoned a BU approach to the natural numbers, for Dummett's suggestion was originally attributed to him by Hao Wang, who claims that it "is more

closely related to the Frege–Dedekind definition [than to the approach of Zermelo, Grelling, and Bernays, who manage without the axiom of infinity]" (Wang 1963, p. 52).

Illumination of this apparent conflict is not furthered by the variation one finds in descriptions of Dummett's suggestion. Wang attributes to Dummett a definition of "ℕ" according to which $k \in \mathbb{N}$ just in case

(i) $(\forall X)((0 \in X \;\&\; (\forall y)((y \in X \;\&\; y \neq k) \rightarrow S(y) \in X)) \rightarrow k \in X)$ and
(ii) $(\exists X)(0 \in X \;\&\; (\forall y)((y \in X \;\&\; y \neq k) \rightarrow S(y) \in X))$.

Parsons, referring to Wang's attribution to Dummett, offers a definition similar to (i) and (ii) and traces the idea back to Zermelo and Grelling (Parsons 1987, p. 206). Isaacson, by contrast, though likewise referring to Wang's attribution to Dummett, offers only clause (i). He adds, however, that in order for (i) "to define anything" (ii) must also obtain for every k (Isaacson 1987, p. 155). Feferman and Hellman (1995, note 5) apparently following Isaacson, also give only clause (i) when describing Dummett's definition. In spite of this, their actual definition is closer to (ii) than to (i), being essentially existential rather than universal.

This confusing variation might be due to different assumptions about the range of the definition's second-order quantifier. If it is assumed to include infinite sets, then (ii) alone will not suffice to exclude all non-numbers (since each such will render it true, for X includes ℕ in its range), but (i) will. Hence, under this assumption, (ii) is superfluous and (i) will do by itself. On the other hand, if the range of the second-order quantifier is taken to consist only of finite sets of all sizes, then (i) will not suffice to exclude all non-numbers (since the antecedent of its instances will be false, if k is not a natural number), whereas (ii) will. Hence, in this second situation, (ii) by itself suffices and (i) is not needed. Offering both clauses, as Wang does, will inevitably be redundant, but may be appropriate if one wishes to provide a definition that works whether or not infinite sets are included in the range of the second-order quantifier.

We are now in a position to resolve the conflict presented in the first paragraph of this note. If one is imagining that second-order quantifiers range over infinite sets, then Dummett's definition effectively consists in clause (i) and Wang is correct to assimilate it to Frege–Dedekind's PD definition: for the elimination of non-natural numbers will require ℕ itself to be in the domain of second-order quantification. On the other hand, if one takes these quantifiers to range only over all finite sets, as Feferman and Hellman do, then Dummett's definition in effect amounts to (ii) and therefore, for reasons given in the previous note, should be likened rather to the BU definition.

In this context, it is worth mentioning another definition of the natural numbers, this time first offered by Quine (1961); see also (Quine 1969, pp. 75ff.). According to this definition, $k \in \mathbb{N}$ just in case

(iii) $(\forall X)((k \in X \;\&\; (\forall y)(S(y) \in X \rightarrow y \in X)) \rightarrow 0 \in X)$.

If we assume that the second-order quantifier ranges over infinite sets, then this definition captures all the natural numbers (since if k is a natural number then every set containing k and closed under predecessor will contain 0) and only them (for if k is not a natural number, say an entity with infinitely many predecessors, then the complement of $\mathbb{N}$ contains k and is closed under predecessor, but does not contain 0, and hence the closure of $\{k\}$ under predecessor will not contain 0).

However, contrary to Quine's suggestion, the definition does not work to exclude non-numbers if the range of the second-order quantifier is restricted to finite sets of all sizes: for if k has infinitely many predecessors, then (iii) will be vacuously true. If the domain is so restricted, then (iii) could be supplemented by (iv):

(iv) $(\exists X)(k \in X \ \&\ (\forall y)(S(y) \in X \to y \in X))$.

It is certainly true that if k is a natural number, then there exists a finite set containing k and closed under predecessor. Furthermore, if k has infinitely many predecessors, then it fails to satisfy (iv), for there will not exist a finite set of the requisite kind. Can one, in this context, make do with (iv) alone? No, for (iv) fails to rule out Caesar as a natural number, because there does exist a finite set containing Caesar and closed under predecessor, namely {Caesar}. But, taking X to be this set, we see that Caesar does not satisfy (iii). In sum, if the second-order quantifier ranges only over all finite sets, then one emendation of Quine's definition consists of the conjunction of (iii) and (iv). (See (George 1987) and (Parsons 1987, pp. 210–1).)

It is not easy to say where this particular emendation of Quine's definition falls in our classificatory scheme, for it contains elements of both the PD and the BU approaches. There is, however, another way of supplementing (iii) that leads to a more straightforward outcome. Consider:

(iv′) $(\exists X)(k \in X \ \&\ (\forall y)(S(y) \in X \to y \in X) \ \&$
$(\forall y)((y \in X \ \&\ y \neq 0) \to (\exists z)(S(z) = y)))$.

Clearly, if k is a natural number, it satisfies (iv′). Also, (iv′) rules out all non-natural numbers, including Caesar, since he is unequal to 0 and has no predecessor. Hence, when the second-order quantifier ranges only over all finite sets, (iii) is superfluous and can be replaced by (iv′). Furthermore, (iv′) is plainly in the spirit of a BU approach to the natural numbers.

Although Feferman and Hellman say that their own approach realizes Dummett's definition, it is in fact closer to (iv′) than it is to (ii), for their definition employs closure under predecessor rather than closure under successor. (Their definition differs from (iv′) in only one respect: they require that the set X be the *smallest* set containing k and closed under predecessor. However, an examination of the proof of their Theorem 7 shows that this additional requirement plays no role in the proof, and therefore could have been eliminated. Thus, (iv′) captures the essence of Feferman and Hellman's definition.)

Recently, Peter Aczel has shown that the existence of a structure satisfying Peano's axioms can be proven in Feferman and Hellman's system *without* using their axiom (Card), which restricts the range of the second-order quantifiers to

Dedekind finite sets (Feferman, personal communication). Aczel's approach is, roughly, to let the universe of the structure be defined by the conjunction of (iii) and (iv), rather than (iv′). This conjunction characterizes the natural numbers whether or not the range of the second-order quantifiers includes infinite sets. However, if infinite sets are included then, as we observed earlier in this note, the exclusion of non-numbers by (iii) requires reference to a collection, the complement of ℕ, that is defined in terms of ℕ, the very set being defined. Thus, despite Feferman and Hellman's description of their formal system as "predicatively justified," it seems questionable to us whether Aczel's definition should be called predicative. Aczel's theorem can also be proven using the conjunction of (i) and (ii), rather than the conjunction of (iii) and (iv). However, if infinite sets are not excluded from the range of the second-order quantifiers, then the use of (i) will once again render the definition impredicative.

11. There is a circularity in this argument, for mathematical induction will be needed in order to show that P holds at every stage of the construction. This is not a circularity in the BU definition of natural number; rather, it is a circularity that appears in the justification of induction on the basis of that definition. (It is reminiscent of the circularity Hume discovered in attempting to justify empirical induction.) Yet the argument is, as Parsons has noted, "no worse than arguments for the validity of elementary logical rules" (1992, p. 143).

One way of summarizing both this argument for induction from the BU definition and the argument for why the PD definition captures only the natural numbers (see above) is to say that the extremal clause and the principle of mathematical induction are interderivable. For example, this is essentially what S. C. Kleene says (1952, p. 22). While correct as far as it goes, we prefer not to put the matter this way, for it obscures the distinctive approaches to the natural numbers that we believe animate the two definitions.

Even more obscuring is to construe the extremal clause as *saying* that induction is valid. Parsons, for example, at one time articulated the view that the principle of induction "could be regarded simply as an interpretation of" the extremal clause. Yet, he did not then advance the position and in fact also mentioned the possibility that "the induction principle [...] will be in some way a consequence of" the extremal clause (Parsons 1967, p. 194). More recently, however, he appears to endorse this view, as when he describes induction as "a principle cashing in our intention that the numbers should be what is obtained by the introduction rules and those alone" (Parsons 1992, p. 143). This way of viewing the matter no doubt contributes to his analysis of possible options:

> The readily available alternatives to something like the induction-definition model of the concept of natural number [...] would be to give it an explanation that is blatantly circular, such as, the natural numbers are what is obtained by beginning with 0 and iterating the successor operation an *arbitrary finite number* of times, or to take the concept of natural number as given and the principle of induction as evident without any explication connecting it with the concept of

> natural number. Either alternative seems to me a counsel of philosophical despair that leaves us with no motivation for the principle of induction. (Parsons 1992, p. 143)

Yet, as we have seen, these alternatives are not exhaustive, for the proponents of the BU definition view it as neither "the induction-definition model of the concept of natural number" (i.e., the PD definition), nor as circular, nor as failing to provide a justification of induction.

Again, we are not arguing in favor of one or the other definition, but rather attempting to delineate two philosophically distinct approaches to the nature of number.

12. For example, this thought may be behind Parsons' claim that "If one explains the notion of natural number in such a way that induction falls out of the explanation, then one will be left with a similar impredicativity"—similar, that is, to the impredicativity of the PD definition (Parsons 1992, p. 141). For a discussion, see (George 1987).

13. Parsons gives voice to this view as well when he suggests that "induction is constitutive of the meaning of the term 'natural number'" (Parsons 1992, p. 155). Dummett, also, believes that "the meaning of the expression 'natural number' involve[s], not only the criterion for recognising a term as standing for a natural number, but also the criterion for asserting something about all natural numbers" (Dummett 1978, p. 194).

14. Edward Nelson, for example, adopts a principle of induction restricted to those predicates involving bounded quantification (Nelson 1986, p. 2).

Another example is (Feferman and Hellman 1995), in which the authors establish induction only for formulas containing no quantification over the collection of all classes. In fact, it can be shown that induction for all formulas is not provable in their system EFSC*. The reason is that, if full induction were added to EFSC* as a new axiom, it would be possible to define a satisfaction relation and use it to prove that all theorems of PA are true in $\mathbb{N}$, and therefore that PA is consistent. But as Feferman and Hellman show in their Metatheorem 9, the consistency of PA is not provable in EFSC*.

It might be helpful to spell out a few more of the details of this proof. The satisfaction relation for formulas in the language of PA can be represented as a function assigning to each pair (φ, s), where φ is a formula and s is an assignment of values to the free variables of φ, one of the values 1 or 0, representing true and false. By assigning Gödel numbers to both formulas and assignments of values to variables, we can think of this function as mapping $\mathbb{N} \times \mathbb{N}$ to $\{0, 1\}$. Let us say that a function from $\{0, 1, 2, \ldots, n\} \times \mathbb{N}$ to $\{0, 1\}$ is an *n-satisfaction function* if it satisfies the usual recursive definition of satisfaction for formulas with Gödel numbers up to n. Then we can prove by induction that $(\forall n \in \mathbb{N})(\exists F)(F$ is an n-satisfaction function). Note that the formula being proven by induction contains the quantifier "$(\exists F)$", where "F" stands for an infinite class, so the weak form of induction proven by Feferman and Hellman would not be sufficient for this proof. Our satisfaction relation $\mathrm{Sat}(x, y)$ can now be defined to be the formula: $(\exists F)(F$ is an x-satisfaction function & $F(x, y) = 1)$.

It is interesting that full induction is provable in a slight strengthening of Feferman and Hellman's system. Let EFSC$^+$ be the same as EFSC*, except with no restrictions on the formula φ in (Sep), the separation axiom for finite sets. Then we can prove full induction in EFSC$^+$ by imitating the proof of Feferman and Hellman's Theorem 8. Let $\varphi(n)$ be any formula, and suppose we have both $\varphi(0)$ and $(\forall n)(\varphi(n) \rightarrow \varphi(n+1))$. If $\neg\varphi(m)$ for some natural number m, then let $B = \{n : n \leq m\}$, a finite set since m is a natural number, and let $Y = \{n \in B : \neg\varphi(n)\}$. Then, as in Feferman and Hellman's proof of their Theorem 8, it can be shown that Y is both finite and Dedekind-infinite, contradicting the cardinality axiom of EFSC$^+$. Note that the definition of Y requires the strengthened version of (Sep), since φ might involve quantification over the collection of all classes.

It is difficult to say whether or not EFSC$^+$ should be considered predicative. The strengthened version of (Sep) allows one to quantify over the collection of all classes when defining a subset of a finite set, and such a definition would appear to be impredicative. But even the original version of (Sep) allows one to define a subset of a finite set by quantifying over the collection of all finite sets, a collection that includes the very set being defined, and Feferman and Hellman do not consider this to be impredicative. Their reason is that they "assume that the notion of *finite set* is predicatively understood, governed by some elementary closure conditions" (p. 2). Feferman and Hellman appear to take this to mean that a definition of a set that involves quantification over the collection of all finite sets is predicative, but one involving quantification over the collection of all classes is not. Thus, they consider their version of (Sep) to be predicatively acceptable, but would presumably reject the strengthened version of (Sep) as impredicative.

Some might question whether such closure conditions for the collection of finite sets should be considered to be predicatively acceptable. But if closure conditions are to be accepted, we believe an argument can be made for a closure condition that implies the strengthened version of (Sep). One way of justifying such a closure condition would be to argue that given a finite set, it is possible to make a finite list of all subsets of that set. Each of these subsets is definable by a predicative definition that simply lists its elements. Any other definition of a subset of the original finite set, even an impredicative one, must pick out one of these subsets, whose existence has already been established by a predicative definition. Thus, for any definition that specifies unambiguously which elements of the finite set are to be included in a subset, there must exist a subset containing precisely the elements specified by that definition. This reasoning would apply whether the definition involves quantification over only the collection of all finite sets or quantification over the collection of all classes, so it would justify not only the original version of (Sep), but also the strengthened version.

15. For example, see the quotation from Parsons in note 12; see also (Nelson

1986, pp. 1-2).

16. This term is suggested by Ludwig Wittgenstein's (1921, 3.263).

17. Thus, we dissent from what seems to be Michael Dummett's position in (Dummett 1993*b*); see p. 443. For some further discussion, see (George 1994).

Bibliography

Dedekind, R. (1967). Letter to Keferstein. In *From Frege to Gödel: a source book in mathematical logic, 1879–1931* (ed. J. van Heijenoort), pp. 98–103. Harvard University Press, Cambridge, Massachusetts.

Dummett, M. A. E. (1978). The philosophical significance of Gödel's theorem. Reprinted in his *Truth and other enigmas*, pp. 186–201. Harvard University Press, Cambridge, Massachusetts.

Dummett, M. A. E. (1993*b*). What is mathematics about? Reprinted in his *The seas of language*, pp. 429–45. Clarendon Press, Oxford

Enderton, H. (1972). *A mathematical introduction to logic.* Academic Press, New York.

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

George, A. (1987). The imprecision of impredicativity. *Mind*, **96**, 514–18.

George, A. (1994). Intuitionism and the poverty of the inference argument. *Topoi*, **13**, 79–82.

Isaacson, D. (1987). Arithmetical truth and hidden higher-order concepts. In *Logic Colloquium '85* (ed. Paris Logic Group), pp. 147–69. North-Holland, Amsterdam.

Kleene, S. C. (1952). *An introduction to metamathematics.* North-Holland, Amsterdam and von Nostrand, Princeton.

Nelson, E. (1986). *Predicative arithmetic.* Mathematical Notes, Vol. 32, Princeton University Press.

Parsons, C. (1967). Mathematics, foundations of. In *The encyclopedia of philosophy*, Vol. 5 (ed. P. Edwards), pp. 188–213. Macmillan, New York.

Parsons, C. (1987). Developing arithmetic in set theory without infinity: some historical remarks. *History and Philosophy of Logic*, **8**, 201–13.

Parsons, C. (1992). The impredicativity of induction. In *Proof, logic and formalization* (ed. M. Detlefsen), pp. 139–61. Routledge, London.

Poincaré, H. (1952). *Science and method* (trans. F. Maitland). Dover, New York.

Quine, W. V. (1961). A basis for number theory in finite classes. *Bull. American Math. Soc.*, **67**, 391–92.

Quine, W. V. (1969). *Set theory and its logic.* Harvard University Press, Cambridge, Massachusetts.

Wang, H. (1963). Eighty years of foundational studies. Reprinted in his *A survey of mathematical logic*, pp. 34–56. North-Holland, Amsterdam.

Wittgenstein, L. (1981). *Tractatus logico-philosophicus*. Routledge, London.

Department of Philosophy
Amherst College
Amherst, MA 01002
USA
email: AGEORGE@amherst.edu

Department of Mathematics
and Computer Science
Amherst College
Amherst, MA 01002
USA
email: DJVELLEMAN@amherst.edu