

Intensions, Church's Thesis, and the Formalization of Mathematics

NICOLAS D. GOODMAN*

I Recently a number of authors have proposed including intensional notions, notably epistemic notions, in the underlying logic of mathematics. (See Goodman [6] as well as several of the essays in Shapiro [15]; and for a discussion very close in spirit to the present essay, see Shapiro [14].) These authors have made explicit just what intensional notions are needed for the analysis of mathematical language and how they should be formalized. The present essay is concerned not with those questions, but rather with trying to make the prior case that intensional notions are genuinely needed for the formalization of classical mathematics.

Let us begin with an example of the use of intensional notions in informal mathematical exposition:

The proof by Gerd Faltings [2] of the Mordell conjecture implies that, for any fixed exponent, there are at most a finite number of counterexamples to the Fermat conjecture for that exponent. That is, let k be some fixed large integer, say $k = 9,437,512,798$. Then there is some finite integer n such that there are exactly n pairs of rational numbers p and q such that

$$p^k + q^k = 1.$$

If the Fermat conjecture for k is true, then $n = 0$. At the moment, however, the number n is not known. In fact, we do not even know how to bound n . Nevertheless, there is some hope that bounds may be obtained. For the first time, we seem to be in a position to make systematic progress on this 350-year-old problem.

It is a remarkable fact that the main logical tradition, stemming from Frege, holds that the last four sentences of the above discussion are not part of

*Preparation of this paper was supported in part by grant number SES8411917 from the National Science Foundation.

mathematics. Within the standard formalizations we can assert that, for any exponent, there are at most a finite number of counterexamples to the Fermat conjecture for that exponent. We can define the number k , and, from it, the number n of counterexamples for k . We cannot, however, assert that n is unknown. We cannot formulate the problem of finding a bound for n , although we can say that there is such a bound, and, if we knew a specific bound, we could say that it was a bound. We cannot express what is scientifically the most important point of the discussion—namely, the present problem situation. We cannot distinguish what we know from what we do not know. We cannot say what we take to be the problem needing to be solved. In general, we cannot formulate those statements mathematicians make that have epistemic content.

Perhaps this is not a defect in the standard formalizations. Perhaps a “canonical notation” for science (Quine [10]) is intended only to allow us to express what we know, not to say what we want to find out. But then it clearly will not suffice for scientific communication, even in principle. As Quine himself says, it is not “that the idioms thus renounced are supposed to be unneeded in the market place or in the laboratory. . . . The doctrine is only that such a canonical idiom can be abstracted and then adhered to in the statement of one’s scientific theory. The doctrine is that all traits of reality worthy of the name can be set down in an idiom of this austere form if in any idiom” ([10], p. 228).

On its face, this is a reasonable doctrine. It has often been held that the principal aim of science is the description of those features of reality which have objective validity, independent of any particular observer or even of our particular scientific community. But it is no longer obvious that this goal makes sense. Much of the effect of the new developments in physics at the beginning of the present century was to undermine just this notion of objectivity. Not many physicists today would hold that it is possible to describe the physical world as it is, independently of any observer. It seems likely, therefore, that a “canonical notation” for a science rich enough to include contemporary physics would have to contain modal notions, including epistemic ones. Just what those notions would look like, however, is far from clear.

Whatever the language needed for formalizing physics, it is widely held that no nonextensional locutions are needed for formalizing mathematics. Mathematics is the most austere of all the sciences. Moreover, “it is an empirical fact that all of mathematics as presently known can be formalized within the **ZFC** system” (Henson [7], p. 130). This system (Zermelo–Fraenkel set theory including the axiom of choice) is formalized in first-order predicate calculus and is a paradigm case of the sort of theory Quine approves of. From a logical point of view, the theory is as austere as anyone could desire. It contains no modal vocabulary to enable us to discuss anyone’s epistemic state. The only possible conclusion, then, is that the discussion above of the current state of the Fermat problem, while perhaps interesting to a mathematician, is not part of mathematics.

No one should be surprised that there are discussions interesting to mathematicians as mathematicians which are not part of mathematics. Consider, for example, a discussion of salaries of mathematicians or of job opportunities for mathematicians. It is the mathematician’s business to study abstract structures. Perhaps the term ‘mathematics’ should be restricted to those things

mathematicians say which are intended as objective descriptions of abstract structures. Such a restriction on our use of the term 'mathematics', however, would at least appear to exclude some propositions which everyone agrees are strictly mathematical. I have in mind the constructive or algorithmic parts of mathematics. It is one thing to assert that a continuous real-valued function on a closed interval which is negative at one end and positive at the other must take on the value zero somewhere in the interior. This is a nonconstructive assertion which can be formalized in purely extensional terms. It is quite another to assert that we have an algorithm, such as bisection or Newton's method, which, in favorable cases, can actually be used to find a point in the interval where the value of the function is as close as we like to zero. The latter assertion seems to contain modal elements. It asserts something about what we can do. As such, it seems to have intensional content which transcends the expressive power of extensional theories like **ZFC**. It seems likely, nevertheless, that even Quine would want to include discussion of such algorithms within the confines of mathematics, strictly so-called.

2 It is difficult to say in full generality what it is for a language to be *extensional*. Using the conceptual apparatus in the philosophy of language derived from Frege, we may say that a language is extensional if the truth-values of its sentences are determined only by the denotations or referents of their terms, and not by their connotations or senses. In practice, as Quine has emphasized, this comes down to whether terms which have the same referent can be freely substituted for each other in all contexts of the language without changing the truth value of any sentence. This will be true of **ZFC**, or of any theory whose underlying logic is the first-order predicate calculus. To show that mathematics cannot be formalized in such a theory, then, it will suffice to show that the language of mathematics contains *oblique* contexts which do not admit the free substitutivity of coreferential terms. (For excellent discussions of the issues involving this notion of oblique contexts, see Linsky [8] and [9].)

If we are willing to include the formulation of mathematical problems within the bounds of strictly mathematical language, then it is easy to find oblique mathematical contexts. Suppose, for example, that the Fermat conjecture is as a matter of fact true. Then the number of counterexamples to the Fermat conjecture for k is, as a matter of fact, zero. It is an open problem to prove that the number of counterexamples to the Fermat conjecture for k is zero. It is not an open problem to prove that zero is zero, even though this problem may be obtained from the previous one by replacing the singular term 'the number of counterexamples to the Fermat conjecture for k ' by the coreferential term 'zero'. On the same assumption, the number of counterexamples to the Fermat conjecture for k is equal to the smallest root of the equation $x^2 - x = 0$, even though we do not know how to find the number of counterexamples to the Fermat conjecture for k but do know how to find the smallest root of the equation $x^2 - x = 0$.

For an example of a rather different character, let f be the function given by $f(x) = x^3 - 3$, and let g be the function such that $g(x) = 1$ if x is rational, $g(x) = 2$ if x is irrational but not the cube root of 3, and $g(x) = 0$ if x is the cube root of 3. Then the unique zero of f is equal to the unique zero of g . Nev-

ertheless, Newton's method can be used to find the zero of f but not to find the zero of g . This example is interesting in that the obliqueness of the context is strongly dependent on the formulation in terms of the problem of finding the zero of the function. After all, Newton's method applied to f will generate a sequence which converges to the zero of g . Perhaps we are wrong in thinking that the sentence 'Newton's method can be used to find the zero of f ' should be analyzed as containing the singular term 'the zero of f '. Even in that case, though, it surely contains the singular term ' f '. Let h be f if the Fermat conjecture is true and g otherwise. Suppose that the Fermat conjecture is true. Extensionally, then, f is h . Nevertheless, we cannot use Newton's method to find the zero of h since we cannot compute very many of its values or any of the values of its derivative.

3 One interesting approach to handling intensionality in mathematical language which has been advocated recently by Feferman (see [3] and [4]) is to think of mathematical language as elliptic. Mathematical objects should be thought of as equipped with additional information that can be used to mitigate the non-constructivities inherent in our more usual ways of talking. For example, a continuous real-valued function of a real variable should, in some contexts, be thought of as being given together with a specific modulus of continuity. Again, a computable function should, in some contexts, be thought of as being given together with a Turing machine program to compute it. In our particular example above, we may think of the sentence 'Newton's method can be used to find the zero of f ' as asserting that Newton's method applied to an algorithm for computing f will give an algorithm for computing the zero of f . On this view, then, we should think of Newton's method as an effective operation which takes one program (the program for computing a function) to another program (a program for computing a zero of the function). This characterization can be stated in extensional language if we can express in such a language what a program is *for*. Certainly we can say in set theory that if a program E computes a function f , and if certain other conditions are satisfied, then the sequence of real numbers generated by Newton's method from f using E will converge to a zero of f . This is, as it were, the extensional content of the assertion that Newton's method can be used to find a zero of f . That is not all the content of the assertion, nevertheless, since it also says something about the utility of this extensional behavior of the algorithm. We may imagine a rather dull student who knows the facts about the convergence of the sequences generated by Newton's method but to whom it has never occurred that these facts can be exploited to find the zeros of functions whose zeros we do not already explicitly know. I do not see how, on a view like Feferman's, this additional information can be thought of as part of mathematics. Nevertheless, it clearly is part of mathematics.

I think it is not possible to talk about mathematical problems in an extensional language. But without talk about problems, it is hard to see how we can capture what is essential to the constructive or algorithmic aspect of mathematics. The point of discussing a particular algorithm is usually that it solves some problem that interests us. I conclude that extensional languages, such as the language of set theory, are incapable of formalizing the whole of mathematics.

4 The question of how to parse the constructive aspect of mathematics is quite old. The first logician in the modern sense, Gottlob Frege, wrote as follows about Euclid's distinction between axioms and postulates:

Postulates seem at first sight to be essentially different from axioms. In Euclid we have the postulate 'Let it be postulated that a straight line may be drawn from any point to any other.' This is obviously introduced with a view to making constructions. The postulates, so it seems, present the simplest procedures for making every construction, and postulate their possibility. At first sight we might perhaps think that none of this is of any help in providing proofs, but only for solving problems. But this would be a mistake, for sometimes an auxiliary line is needed for a proof, and sometimes an auxiliary point, an auxiliary number – an auxiliary object of some kind. In the proof of a theorem an auxiliary object is one of which nothing is said in the theorem, but which is required for the proof, so that this would collapse if there were no such object. And if there is no such object, it seems that we must be able to create one and we need a postulate to ensure that this is possible. But what in actual fact is this drawing a line? It is not, at any rate, a line in the geometrical sense that we are creating when we make a stroke with a pencil. And how in this way are we to connect a point in the interior of Sirius with a point in Rigel? Our postulate cannot refer to any such external procedure. It refers rather to something conceptual. But what is here in question is not a subjective, psychological possibility, but an objective one. Surely the truth of a theorem cannot really depend on something we do, when it holds quite independently of us. So the only way of regarding the matter is that by drawing a straight line we merely become ourselves aware of what obtains independently of us. So the content of our postulate is essentially this, that given any two points there is a straight line connecting them. So a postulate is a truth as is an axiom, its only peculiarity being that it asserts the existence of something with certain properties. From this it follows that there is no real need to distinguish axioms and postulates. A postulate can be regarded as a special case of an axiom. ([5], pp. 206–207)

I agree that mathematics has nothing to do with any “subjective, psychological possibility,” and that “the truth of a theorem cannot really depend on something we do”. What is speaking here is Frege's platonism. He holds, as do I, that mathematics is about abstract objects which are neither physical nor mental and that it enunciates truths about those objects whose truth in no way depends on our activity or on our psychological state. It follows then, as Frege asserts, that the possibility referred to by the postulate must be objective. It nevertheless remains a possibility, not an actuality. When Hilbert gave a modern axiomatization of geometry at the beginning of the present century, he asserted the bald existence of the line. Euclid, however, also asserted that it can be constructed. The modality is not in any way refuted by Frege's argument. It is merely denied. In essence Frege argues for an extensional reading of Euclid by denying the possibility of objective modalities. Such arguments are still common today. Nevertheless, the natural way to read Euclid here is not extensional, although more needs to be said about what kind of intensionality is involved and about what kind of objective modal notions are required.

We may look at Frege's argument somewhat differently. Frege holds that the actual existence of the line is the ground for the possibility of our drawing

the line. We cannot draw a line unless, conceptually, there is already a line there to be drawn. But then Frege holds that the existence of the line is all that is needed for the cogency of the proof in which that line is referred to. We may as well drop the possibility of drawing the line from our considerations. Thus on Frege's view we may replace the constructive postulate by the nonconstructive existential axiom.

John Myhill has pointed out to me that some of Euclid's constructions contain what are *prima facie* nonconstructive steps. For example, the treatment of the Euclidean algorithm in Book X, Proposition 3 begins with an unjustified use of excluded middle. In this case, a justification is not very hard to supply: specifically, it suffices to prove Proposition 5 before Proposition 3. Nevertheless, Euclid's presentation indicates either that he is not aware of this nonconstructivity or that he is not concerned about it. In the latter case, Frege is probably right that we should view Euclid's constructions as mere existence proofs, even though this conflicts with a literal reading of Euclid's language.

Frege's position concerning Euclid's postulates is reductionistic. He claims that we can reduce the intensional, modal, constructive aspect of Greek mathematics to its extensional, nonmodal, nonconstructive aspect. He can only do this, however, if he is willing to give up all discussion of problems. He must reduce every problem of the form "to construct a such-and-so" to the propositional form "to show that there is a such-and-so".

5 Geometrical constructions are not central to contemporary mathematics. Even geometry as practiced today is thoroughly algebraic or analytic. It treats of actual space, if at all, only by way of the familiar correspondence between points and triples of real numbers. Thus the problem Frege found in interpreting Euclid is not urgent for the interpreter of modern mathematics. Nevertheless, the difficulty remains that mathematics appears to have a modal component in its constructive aspect. Large parts of mathematics, like numerical analysis, are centrally concerned with the construction and analysis of algorithms for the effective solution of problems whose nonconstructive solutions are well known. To refer back to our example above, the nonconstructive mathematician may be able to prove that a function has a zero, but the numerical analyst wants to know how to find it. The nonconstructive mathematician may have a proof that a particular matrix is nonsingular, but the numerical analyst wants to know how to invert it. Every junior mathematics major can show that all matrices have eigenvalues, but the literature contains many papers and entire books devoted to the problem of how to find the eigenvalues.

Now it is true that the value of the zero of the function, the components of the inverse matrix, or the eigenvalues of the matrix are rarely relevant to proofs that the pure mathematician wants to construct. Their bare existence is generally enough. The extent to which this is true may be seen from the fact that the algorithms found in elementary textbooks are often not feasible in practice. Some books on linear algebra, for example, suggest computing the eigenvalues of a matrix by finding the characteristic polynomial directly from its determinant and then factoring it. Others suggest finding the minimal polynomial and factoring it. Neither of these approaches makes numerical sense for matrices of

order larger than, say, four or five. Indeed, it is not possible in general to compute the minimal polynomial even of a three by three matrix in a numerically stable way. Presumably the authors of these textbooks have never had occasion actually to compute eigenvalues in nontrivial cases. That is not their field.

Even if most pure mathematicians do not need to find the numbers they discuss, the study of how to find them is surely still part of mathematics. Journals like *The SIAM Journal of Numerical Analysis*, *Numerische Mathematik*, or *The Mathematics of Computation* find space on the shelves of our mathematics libraries. Of course, much of this material is thought of as applied mathematics. In some circles there are still doubts as to the respectability or centrality of such work. That we think of it as applied mathematics, however, is a sociological fact rather than a fact about the nature of the work. Applied mathematics, strictly so called, is mathematical reasoning about nonmathematical objects employing, in part, nonmathematical premises. Numerical analysis is applied in that sense only when it descends to the nitty-gritty of how to implement the algorithms being discussed on a specific machine. This occurs rarely in the literature. Most of the content of journals like those mentioned above is quite pure. Nor is the level of rigor in such journals lower than that in other mathematical journals. There are examples of computational mathematics, moreover, which are not at all applied even in the weak sense that they might be of use to an engineer. We may think, for example, of the recent use of computers to help solve the four-color problem or to help in the classification of finite groups. It does not seem plausible to exclude all such algorithmic work from mathematics.

If the search for better algorithms to find numbers we already know to exist is part of mathematics, then a logician interested in the formalization of mathematics must make allowances for this work in his account of the foundations of mathematics. If he holds, like Frege and almost every philosopher of mathematics since Frege, that mathematical language can be given a purely extensional interpretation, then he must have some doctrine about how the apparently intensional aspects of such talk about algorithms can be given an extensional analysis. The core of such an analysis must presumably be a reductionistic thesis which tells him how to reduce talk about algorithms and computation to talk about numbers, functions, and sets. It is widely believed that Church's thesis is such a reductionistic thesis and that, since Church's thesis is true, it solves the problem of giving an extensional account of the constructive aspect of mathematics. I harbor no doubts about the truth of Church's thesis. Nevertheless, it is one of the main points of the present essay to argue that Church's thesis is not a reductionistic thesis and that it does nothing to support an extensional reading of the logic of mathematics.

6 The practicing mathematician appears to deal with a wide variety of objects. There are sets and functions, numbers and points, ordered pairs and logical formulas. The set-theoretic reductionist proposes to replace this motley by the monochrome universe of sets. That this can actually be carried out in any reasonable sense is a tribute to the ingenuity of the set theorist and to the expressive power and proof-theoretic strength of **ZFC**. Of course, the reduction has its problems. No one would believe that the ordered pair of x and y is really the

unordered pair, or doubleton, consisting of the singleton of x and the doubleton of x and y . The latter construct, called the Wiener–Kuratowski ordered pair, may be an acceptable substitute for the ordinary ordered pair in many contexts, but it in no sense constitutes an analysis of the notion of ordered pair as it exists in informal mathematical reasoning. There are many other constructs which would serve just as well. As Quine puts it ([11], p. 58), “The fundamental law demanded of ordered pairs is that $\langle x, y \rangle = \langle z, w \rangle$ not whenever $\{x, y\} = \{z, w\}$, but only when $x = z$ and $y = w$. Any definition of ‘ $\langle x, y \rangle$ ’, however arbitrary and artificial, is to the purpose if it fulfills this fundamental law”.

It is worth emphasizing that in this respect set-theoretic reductionism differs fundamentally from other scientific reductionisms. For example, consider the claim that heat is really mean molecular kinetic energy. Within the context of molecular physics and statistical mechanics, there is nothing arbitrary about this choice. There is no other construct one could form within the framework of that theory which would serve as well. Moreover, although mean molecular kinetic energy is not a conceptual analysis of the concept of heat, it is a physical analysis of heat as it actually is. To doubt that heat really is mean molecular kinetic energy is to doubt the underlying theory. Any property of mean molecular kinetic energy derivable from the relevant physics is ipso facto a property of heat. If there were properties of mean molecular kinetic energy which were plainly not properties of heat, that would constitute evidence disconfirming the theory, or at least disconfirming the alleged reduction. On the other hand, no one believes that x is really an element of an element of the ordered pair of x and y . That is just an arbitrary feature of the reduction. To that extent, then, the set-theoretic reductionist need not hold that ordered pairs are really Wiener–Kuratowski ordered pairs, merely that they may be replaced by Wiener–Kuratowski ordered pairs for all ordinary mathematical purposes. As Quine characteristically puts it, ordered pairs may be dispensed with in favor of Wiener–Kuratowski ordered pairs (see, for example, [12], p. 55).

The advantage of so dispensing with ordered pairs is the conceptual simplicity and economy thereby effected. On one level, we might respond that anyone who would be willing to countenance the universe of set theory, with its ontological principle of plenitude, ought not to stick at anything. Having accepted more infinite cardinals than there are elements in any set, why should he cavil at a few ordered pairs? From a technical point of view, however, there certainly is a genuine economy to be achieved by restricting the number of fundamental kinds. In the case of ordered pairs, in fact, there may be no offsetting loss. In other cases, however, the situation is far worse.

Consider, for example, any of the reductions of arithmetic to set theory. Numbers are identified with certain rather complicated sets. According to John von Neumann, the number zero is the empty set. The number one is the singleton of the empty set. The number two is the doubleton of the numbers zero and one. In general, each natural number is the set of its predecessors. Thus each natural number is a set of its own cardinality. From a technical point of view, this construction is very elegant because it makes the natural numbers fit together smoothly with the infinite ordinals. From the point of view of pure mathematics, moreover, this reduction probably does no harm. Its worst defect is its arbitrariness.

Natural numbers, however, also play a role in the lives of people who are not pure mathematicians. Von Neumann's finite ordinals cannot support these other roles comfortably. For example, consider the child who learns that two plus three make five by counting on her fingers. After making this experiment, she knows that two plus three make five. That is, she knows a proposition about natural numbers. Any account of what natural numbers are must ground her knowledge. That is, it must make it comprehensible how she came to have this knowledge by counting on her fingers.

The standard story about the connection between her fingers and the fact that two plus three makes five is as follows. There is a one-to-one correspondence between the set of her first two fingers and the von Neumann ordinal two. Moreover, there is a one-to-one correspondence between the last three fingers on her hand and the von Neumann ordinal three. These two sets of fingers are disjoint. Hence the cardinality of their union must be the sum of their cardinalities. When she counts them all, however, she establishes a one-to-one correspondence between the set of all of the fingers on her hand and the elements of the von Neumann ordinal five. The result that two plus three makes five follows.

The trouble with this account is that it says nothing about the child's knowledge. If 'five' is the name of the von Neumann ordinal, then to know something about five it must be necessary somehow to have access to that von Neumann ordinal. We would not ordinarily say that the small child knows that the positron has the same mass as the electron, even if she has been taught to repeat this formula, because the child knows nothing about the complex theory that gives this statement its context, and even less about the experimental procedures and equipment that give it its operational meaning. Nevertheless, I think the child does know that two plus three make five, and that her experiment with her fingers is an adequate proof of that fact. If I am right, then the standard story above is irrelevant to the grounds of her knowledge, and therefore also irrelevant to the grounds of my knowledge that two plus three make five. From this, in turn, it follows that 'five' is not the name of the von Neumann ordinal.

Difficulties of this general kind become even more striking when we consider the case of geometry. The set-theoretic reductionist must claim that spatial points are ordered triples of real numbers. (For a discussion of this view and some references for the identification of points with triples of real numbers in mathematical textbooks, see Goodman [6].) It is obvious that many people know geometric facts who have never had a course in analytic geometry or calculus, and who therefore do not know anything about the "geometric" properties of triples of real numbers. In fact, when we teach analytic geometry, we usually rely on the prior geometric knowledge of the students. The first thing taught in such a course is the correspondence between points in the plane and ordered pairs of real numbers. We are assuming that some of the students did not previously know that connection. Hence their previous geometric knowledge, if well grounded, could not have been knowledge about triples of real numbers.

Again let us contrast this kind of case with that of the reduction of heat to mean molecular kinetic energy. Children learn about the properties of heat empirically, being burned by fires, warmed by sunshine, and chilled by cold winds. Assuming that the kinetic theory is correct, these are all encounters with mean molecular kinetic energy. Children also learn the properties of natural

numbers empirically, but it is implausible to hold that counting on one's fingers is an encounter with the von Neumann numbers. Children also learn geometric facts empirically, as when they fold paper and see that a diagonal fold divides the square sheet into congruent triangles. Again, it is implausible to hold that this is an encounter with triples of real numbers, or sets of triples of real numbers.

When we teach analytic geometry, as a matter of fact, we describe ourselves as coordinatizing space—that is, as giving a correspondence between points and triples of real numbers. We give no argument to the effect that points really are triples of real numbers. At this stage of the student's education, then, he is faced with two sorts of things which are related. Later in his mathematical education, one of these kinds of things drops out. We might hold, of course, that physical space and mathematical space are distinct, and that what we are relating are these two kinds of space. But in that case the study of physical space, which was surely Euclid's subject, has somehow ceased being mathematics and become physics. This divorce between pure and applied mathematics, which is forced on us by set-theoretic reductionism, is very widely deplored just now. On its face, nevertheless, this is not an implausible move in the geometric case. Are we willing, however, to make the same move in the arithmetic case? Are there physical natural numbers and mathematical natural numbers which are put in one-to-one correspondence in the explanation of the von Neumann reduction? This line of argument seems to lead far. In pure set theory one only considers sets of sets of sets of sets of. . . . No individuals, or urelements, are countenanced. Perhaps this is merely a matter of technical convenience, and not a matter of principle. Nevertheless, we might take it seriously and hold that the real numbers and functions of the physicist, which are related to objects which are not sets, are physical real numbers and functions, not mathematical real numbers and functions. In that case, there would be a whole universe of physical mathematical objects in addition to the universe of set theory. But then I urge that we forget about the universe of set theory and go back to studying the physical mathematical objects. Those were the ones we were originally interested in anyway.

Set-theoretic reductionism characteristically seeks to replace mathematical objects which we do not naturally think of as sets by suitably chosen set-theoretic constructs. These constructs are generally rather arbitrary, although there is often considerable ingenuity involved in selecting them in such a way that they will fulfill the purely mathematical uses of the objects they are intended to replace. The question I now want to raise is whether Church's thesis is in that sense an example of mathematical reductionism. More specifically, the question is whether we should think of Church's thesis as grounding the replacement of the informal mathematical notion of algorithm with the formal set-theoretic notion of a Turing machine program.

7 An *algorithm*, in the informal mathematical sense, is a specific procedure for solving a particular kind of mathematical *problem*. Examples are Euclid's algorithm for finding the greatest common divisor of two integers, Newton's method for finding the zeros of a differentiable function of a real variable, or

Simpson's rule for approximating the Riemann integral of a continuous function of a real variable. The specification of the algorithm is complete only if it includes a statement of the problem it is intended to solve. That statement, however, is inherently intensional, as we saw above. In each case, moreover, the algorithm is susceptible of many different implementations on many different machines. To think of such an algorithm as a Turing machine program, or even as a set of Turing machine programs, is to ignore much of its structure. For example, the Turing machine program does not tell you what the program is for. A Turing machine program which implemented Newton's method with respect to some representation of the real numbers and of functions on the real numbers would be enormously complicated and contain a great deal of irrelevant arbitrariness. More important, it might be extremely difficult to determine from the code that this was an implementation of Newton's method. It is for this reason that it is universally agreed that good programming practice requires adequate documentation to explain what the program is for, how it is to be used, how it works, what the different variables represent, and so forth. This documentation contains the intensional content which is missing from the bare machine code (or even the bare higher level language code) and brings the program closer to the algorithm which it is intended to implement.

Suppose a computing machine is running a program. There is no reductionism involved in describing that situation as consisting of this machine running this program. In fact, if we have access to the machine code, then there is no reductionism involved in describing what is going on in terms of what bits are going where. The system is already reduced. If we have access only to the machine code and to the actual behavior of the machine, then there will generally be no other way to describe the process we are observing. What algorithm the machine should be described as following or what data the machine should be described as operating on depend on the coding of abstract objects as bit patterns in the machine. This coding is arbitrary. It depends on the intentions or interpretations of the user of the machine. The detailed behavior of the machine can almost always be interpreted in more than one way if we are willing to change our interpretation both of the data and of the algorithm. Thus the actual behavior of the machine does not uniquely determine what problem the machine is working on or even what algorithm, in the usual mathematical sense, the machine is following.

This point is worth emphasizing. The fact that this computer is using Gaussian elimination with partial pivoting to solve that system of linear equations is a fact about the relationship between this piece of hardware and the human being who is using this piece of hardware for his mathematical purposes. It is intrinsic neither to the hardware nor to the machine code that the computer is executing. Against this it is often urged that there may be no other way to make sense of the behavior of the computer other than by understanding that it is using Gaussian elimination with partial pivoting to solve that system of linear equations. That may be so. But if it is so, then it is a fact about human beings that we are able to comprehend very complex behavior only by imposing this sort of purposive structure on it. We may imagine an extraterrestrial intelligence with sufficient powers of concentration and sufficient combinatorial acumen that it can directly see where all the bits are going and in this way predict the future

behavior of the machine, based on the fact that it is executing this sequence of machine instructions, without ever understanding that these bit sequences code floating point numbers, let alone that the machine is solving a system of linear equations. Such an intelligence might not feel any need for an additional, goal-oriented account of the machine's behavior, since for it there would be no remaining mystery to explain. If we tell this intelligence that the machine is solving a system of linear equations, it may respond, "Yes, that is an amusing way to look at it, but what is really going on, of course, is that . . ." followed by a description of where all the bits are going.

The view I am opposing is sometimes put by saying that mental qualities are *emergent* from the mindless bit-shuffling of the computer. Computers running sufficiently complex programs are said to know, to understand, and even to intend. This kind of emergentism is closely related to the sort of reductionism which holds that human beings who actually do know, understand, and intend are, in some important sense, merely shuffling bits. More specifically, if the computer as it stands and independently of any relation to its human programmer can be said to be applying Gaussian elimination with partial pivoting to solve a system of linear equations, then that makes it far more plausible that talk about Gaussian elimination is just talk about the extensional behavior of physical or abstract machines. Conversely, if the machine in and of itself cannot be said to be carrying out an algorithm in the informal mathematical sense, then talk about the algorithm is not merely talk about the behavior of a machine.

Except as a psychological observation about human beings, I think it is not plausible to assert that the only way to understand the machine's behavior is by understanding that it is carrying out this or that algorithm. There is, however, another argument offered in favor of the emergentist view that the machine is intrinsically carrying out some algorithm. That is namely the question as to what else could be going on when a human being consciously executes the algorithm which is not going on when the machine carries out the same steps. To return for a moment to Frege's formulation of the problem, the question is, what else is involved in Euclid's assertion that "a straight line may be drawn" which is not captured by the nonconstructive fact of the conceptual existence of the line.

Suppose that a student is successfully doing an exercise in a recursive function theory course which consists in implementing a certain Turing machine program. There is then no reductionism involved in saying that he is carrying out a Turing machine program. He intends to be carrying out a Turing machine program. As I said above, the system is already reduced. Now suppose that, unbeknownst to the student, the Turing machine program he is carrying out is an implementation of the Euclidean algorithm. His instructor, looking at the pages of more or less meaningless computations handed in by the student, can tell from them that the greatest common divisor of 24 and 56 is 8. The student, not knowing the purpose of the machine instructions he is carrying out, cannot draw the same conclusion from his own work. I suggest that the instructor, but not the student, should be described as carrying out the Euclidean algorithm. (This is a version, adapted to my purposes, of Searle's Chinese room argument. See [13].) It seems to me, therefore, that we may answer the question as to what else is involved by saying that it is the understanding of the purpose of the com-

putations. This is, however, an intensional notion which is not expressible in an extensional language like that of **ZFC**.

On the other hand, if a human being is carrying out some standard procedure in numerical analysis rather than consciously intending to follow a Turing machine program, then there is a genuinely illicit reductionism involved in describing his behavior as consisting in following a Turing machine program. If the day comes when we can analyze human thought in neuronal terms, without residue, then it may be that at that time we will be able to say unambiguously what machine program (though not Turing machine program) this student is carrying out when, this afternoon, he solves a system of three simultaneous linear equations. The evidence available today, however, allows us to doubt that such a complete reduction is possible. Even if such a complete reduction is performed, moreover, it will not show that the carrying out of the algorithm consists merely in the execution of some brain program. For, as I argued just now, carrying out the algorithm has an additional intensional component which is not captured by any analysis, no matter how complete, of the individual computational steps.

Let us concede for a moment that a complete reduction of mental behavior to neurology is possible, and let us, for that same moment, ignore the intensional component of the assertion that someone is following an algorithm. Even with these concessions, it does not follow that mathematical talk about algorithms can be replaced by talk about machine programs—even about human brain programs. For consider the case in which an experienced teacher, an inexperienced student, and a suitably programmed microcomputer are all solving the same system of three simultaneous equations by Gaussian elimination. We would certainly normally say that they are all applying the same algorithm in the mathematical sense, since we just agreed to ignore the problems involved in saying that the microcomputer is applying any algorithm at all. Nevertheless, there is no reason to think that they are following similar machine programs. For example, the student, unlike the teacher, frequently consults the textbook to make sure of the rules of the procedure. The student and the teacher, unlike the microcomputer, use paper and pencil to store intermediate results. This list of differences could obviously be extended considerably. Thus even on these premises, talk about mathematical algorithms in an extensional context like set theory will require a notion of two machine programs being instances of the same algorithm. Of course, this notion as it ordinarily occurs is quite vague. That is not the main objection, however, since the mathematization of informal theories always involves the replacement of vague notions by precise ones which can only be approximations to the notions they replace. The main objection, rather, is that, as I argued above, the question of whether two machine programs embody the same algorithm, or of what algorithm a particular machine program implements, depends on the context. That is, we can know what algorithm a particular machine program implements only if we know the intentions and interpretations of the programmer. If that is right, then there is no hope of giving an extensional account of the relation which holds between two machine programs when they implement the same algorithm. Indeed, there is no such relation. Two machine programs may implement the same algorithm in one context and different algorithms in another context. (Here I think of a machine program as a type, not as a token. Perhaps we should say that a specific token of

a machine program will always embody only the algorithm which its programmer had in mind. But such tokens are not discussible in the usual abstract mathematical formalisms anyway.)

Earlier I argued that natural numbers cannot be von Neumann ordinals because children learn propositions about natural numbers without ever encountering von Neumann ordinals and that points in space cannot be triples of real numbers because children learn geometric facts without ever encountering triples of real numbers. It is interesting to ask whether we can make an analogous argument here. Certainly grade school children have traditionally learned the long division algorithm without being taught anything about Turing machines or any other kind of computers. My opponent may argue, however, that the child's carrying out of the long division algorithm is precisely the implementation of some suitable program by the computer which is his brain. The child learning the long division algorithm does encounter a machine program in the sense that he is trying to get his own brain to carry out a suitable program. I agree with this analysis, although I deny that that is all that is going on. Thus it seems to me that, unlike the situation in most set-theoretic reductions, we are dealing here with a genuine scientific reductionism. My objection, then, is only that I think the proposed reduction of algorithms to Turing machine programs is incorrect because it fails to capture the content of the mathematical talk it is intended to reduce. It must fail because it is extensional and the mathematical talk it attempts to analyze is intensional. The logic of the two discourses is irreducibly different.

In particular, consider a physicalistic philosopher who holds that all of our mentalistic vocabulary is either meaningless or else can be explained in extensional terms by referring to the structural features of the programs our brains implement. Such a philosopher must hold with Quine that the intensional use of our verbs of propositional attitude, like 'know', is meaningless, at least in the sense that it will not form part of the "canonical notation" of science. My claim here, however, is that such a philosopher would have to exclude significant parts of mathematics from science as well. That seems to me a strong argument against such a physicalism.

8 I have argued that informal talk of mathematical algorithms and mathematical problem-solving cannot be replaced by more formal talk about machine programs without essential loss. Church's thesis, on the other hand, asserts that talk about the existence of algorithms or the solvability of mathematical problems can be replaced by talk about the existence of machine programs. There is no conflict here. Let us recall the precise wording of Church's thesis.

In 1936 Alonzo Church wrote as follows:

We now define the notion, already discussed, of an *effectively calculable* function of positive integers by identifying it with the notion of a recursive function of positive integers (or of a λ -definable function of positive integers). This definition is thought to be justified by the considerations which follow, so far as positive justification can ever be obtained for the selection of a formal definition to correspond to an intuitive notion. ([1], p. 100)

This is Church's thesis as originally formulated. In the terms we used above, it asserts that any effectively calculable function is extensionally the same as the function computed by some Turing machine program. Of course, it does not assert that we can actually find this Turing machine program, or that the Turing machine program is uniquely determined.

It is incorrect to think of Church's thesis as providing an analysis of the informal concept of algorithm. It at most provides a necessary condition for the existence of an algorithm. That is, a problem which no Turing machine can solve cannot be solved algorithmically. However, a Turing machine program without additional explanation is not an algorithm, and an algorithm is not as it stands a Turing machine program. Someone who wanted to argue against the truth of Church's thesis would presumably have to find some mechanism, or some effective procedure carried out by human beings, whose extensional behavior could not be simulated by any Turing machine. I think that is impossible. My contention is rather that not all of the content of our informal intensional talk about algorithms is captured by extensional talk about Turing machine programs.

From this point of view, then, there is nothing reductionistic about Church's thesis, since it does not assert that any previously known objects can be eliminated in favor of any new objects. All of the usual apparatus of constructive mathematics is left intact by Church's thesis. A numerical analyst, for example, who fervently believed in Church's thesis would still have to talk about informal algorithms and their various machine implementations just as before. As a matter of fact, such a numerical analyst would be almost indistinguishable from the more usual kind who has never heard of Church's thesis. Of course, a number theorist who believes that natural numbers are really von Neumann ordinals might also be indistinguishable from the more usual sort of number theorist. This is one of the peculiarities of some mathematical reductions—they make no difference to anyone other than the logician or the philosopher. This is in distinction to successful reductions in other parts of science. A thermal physicist who comes to believe that heat is mean molecular kinetic energy may well be led to new conclusions that he could not have arrived at before. The reduction of genes to DNA has led to a revolution in genetics. Church's thesis, on the other hand, has had no effect on branches of mathematics like numerical analysis which are concerned with particular algorithms, rather than with the possibility of algorithms. In general, mathematical reductions usually at least bring about a change of language. Often they bring about changes in proof procedures. For example, the reduction by Cauchy and Weierstrass of talk about infinitesimals to talk about epsilons and deltas produced a profound change in the appearance of rigorous analysis. Even a number theorist who was committed to the identity of natural numbers and von Neumann ordinals might at least change his language. He might say, for example, that three is a member of five where most of us would say that three is less than five. It is not evident just what change even in language would come about for the numerical analyst who accepted Church's thesis.

The general adoption of Church's thesis in the later 1930s did bring about important mathematical consequences. The development of recursive function theory, one of the richest and most highly developed parts of mathematical logic, would have been quite unmotivated without Church's thesis. But this is a the-

ory of computability, not a theory of mathematical algorithms in the informal sense. It replaces no intensional objects. It just gives new insight into what intensional objects are possible.

It might appear, nevertheless, that Church's thesis does perform a reduction along the lines of Frege's proposed elimination of the modality involved in Euclid's constructive language. That is, someone might hold that the existence of a Turing machine program for computing the values of a certain function f can be used to explicate the possibility in principle of computing f . In many contexts that is correct. In particular, Church's thesis is generally used to underpin proofs that some function cannot be computed by showing that there is no Turing machine program for computing it. There are contexts, nevertheless, in which modalities very close to this one cannot be eliminated by an appeal to Church's thesis. For example, suppose I assert that I know how to compute some specific function f in whose values you are interested. Suppose you say, "Oh good. I need to know the value of $f(3)$. Please compute it for me". And now suppose that I reply, "Oh, I can't actually compute any of the values. But I have an elegant nonconstructive proof that f is recursive. Would you like to see it?" Then it seems to me that you are justified in feeling misled. I do not really know how to compute f .

To say that we can calculate the function f is to say more than that there is an algorithm for computing f . It is also to say that we know such an algorithm, call it E , and that we know that E actually computes f . But such epistemic modalities are irreducibly intensional. They introduce oblique contexts that cannot be eliminated. Thus the intensional epistemic component of a claim to be able to compute a particular function cannot be eliminated by an appeal to Church's thesis. Even if we assume that the numerical analyst means Turing-machine computability when he speaks of computability in principle, this assumption will not enable us to express in extensional terms his claims to be able to compute particular functions.

9 Mathematics is that human activity in which we come closest to perfect objectivity. Here there are the fewest problems of bad communication, of personal bias, or of irrational belief. Here, of all the sciences, it is clearest what constitutes correct argument, adequate evidence, precise definition. It is for these reasons, I think, that we have come closest in mathematics to writing down a Quinean "canonical notation" for the truths of the science. There is a sense, indeed, in which it is plausible to hold that everything which is objectively true about pure mathematical objects can be expressed in the language of set theory and that everything we now know to be true about those objects can be proved in **ZFC**. That is the sense in which Henson is right that all known mathematics can be formalized in **ZFC**. Even in mathematics, however, our objectivity is not perfect. Mathematics is not just an abstract doctrine that is true. Mathematics is an activity that we engage in. That engagement brings with it an irreducible intensionality in the language we use to talk about our activity. As the geometer Felix Klein is said to have remarked, it is impossible to do mathematics without a problem. But a problem cannot be formulated in purely extensional terms, because it only exists as a problem in a certain epistemic situation. Thus

there is also a sense in which Henson is wrong. Not all known mathematics can be formalized in **ZFC**. Frege is right that “the truth of a theorem cannot really depend on something we do”. But the essence of mathematics is solving mathematical problems, and that just is “something we do”.

REFERENCES

- [1] Church, A., “An unsolvable problem of elementary number theory,” *American Journal of Mathematics*, vol. 58 (1936), pp. 345–363. Reprinted in M. Davis, ed., *The Undecidable*, Raven Press, Hewlett, NY, 1965, pp. 88–107.
- [2] Faltings, G., “Endlichkeitssätze für abelsche Varietäten über Zahlkörpern,” *Inventiones Mathematicae*, vol. 73 (1983), pp. 349–366.
- [3] Feferman, S., “Intensionality in mathematics,” paper presented to the Symposium on Intensions and Set Theory, meeting of the Pacific Division of the American Philosophical Association, March 22–24, 1984, Long Beach, California.
- [4] Feferman, S., “Between constructive and classical mathematics,” preprint.
- [5] Frege, G., *Posthumous Writings*, ed. H. Hermes, F. Kambartel, and F. Kaulbach, trans. P. Long and R. White, University of Chicago Press, Chicago, 1979.
- [6] Goodman, N. D., “The knowing mathematician,” *Synthèse*, vol. 60 (1984), pp. 21–38.
- [7] Henson, C. W., “Review of Set Theory: An Introduction to Independence Proofs, by K. Kunen”, *Bulletin of the American Mathematical Society (New Series)*, vol. 10 (1984), pp. 129–131.
- [8] Linsky, L., *Names and Descriptions*, University of Chicago Press, Chicago, 1977.
- [9] Linsky, L., *Oblique Contexts*, University of Chicago Press, Chicago, 1983.
- [10] Quine, W. V., *Word and Object*, Wiley Publishing Co., New York, 1960.
- [11] Quine, W. V., *Set Theory and its Logic*, Harvard University Press, Cambridge, 1963.
- [12] Quine, W. V., *Ontological Relativity and Other Essays*, Columbia University Press, New York, 1969.
- [13] Searle, J. R., “Minds, brains, and programs,” *The Behavioral and Brain Sciences*, vol. 3 (1980), pp. 417–457. Reprinted as pp. 353–373 in Hofstadter, D. R. and Dennett, D. C., eds., *The Mind's I*, Basic Books, New York, 1981.
- [14] Shapiro, S., “On the notion of effectiveness,” *History and Philosophy of Logic*, vol. 1 (1980), pp. 209–230.
- [15] Shapiro, S., ed., *Intensional Mathematics*, North-Holland Publishing Co., Amsterdam, 1985.

*Department of Mathematics
State University of New York at Buffalo
Buffalo, NY 14214*