COMPUTERS IN MATHEMATICAL INQUIRY

JEREMY AVIGAD

1. Introduction

Computers are playing an increasingly central role in mathematical practice. What are we to make of the new methods of inquiry?

In Section 2, I survey some of the ways that computers are used in mathematics. These raise questions that seem to have a generally epistemological character, although they do not fall squarely under a traditional philosophical purview. The goal of this article is to try to articulate some of these questions more clearly, and assess the philosophical methods that may be brought to bear. In Section 3, I note that most of the issues can be classified under two headings: some deal with the ability of computers to deliver appropriate "evidence" for mathematical assertions, a notion that is explored in Section 4, while others deal with the ability of computers to deliver appropriate mathematical "understanding," a notion that is considered in Section 5. Final thoughts are provided in Section 6.

2. Uses of computers in mathematics

Computers have had a dramatic influence on almost every arena of scientific and technological development, and large tracts of mathematics have been developed to support such applications. But this essay is not about the numerical, symbolic, and statistical methods that make it possible to use the computer effectively in scientific domains. We will be concerned, rather, with applications of computers *to* mathematics, that is, the sense in which computers can help us acquire mathematical knowledge and understanding.

Two recent books, *Mathematics by experiment: plausible reasoning in the 21st century* (Borwein and Bailey, 2004) and *Experimentation in mathematics: computational paths to discovery* (Borwein *et al.*, 2004) provide a fine overview of the ways that computers have been used in this regard (see also the associated "Experimental Mathematics Website," which provides additional links and resources). Mounting awareness of the importance of such methods led to the launch of a new journal, *Experimental Mathematics*, in 1992. The introduction to the first book nicely characterizes the new mode of inquiry:

> The new approach to mathematics—the utilization of advanced computing technology in mathematical research—is often called *experimental mathematics*. The computer provides the mathematician with a "laboratory" in which he or she can perform experiments: analyzing examples, testing out new ideas, or searching for patterns... To be precise, by experimental mathematics, we mean

---

I am grateful to Ben Jantzen and Teddy Seidenfeld for discussions of the notion of plausibility in mathematics; to Ed Dean, Steve Kieffer, and Paolo Mancosu, for comments and corrections; and to Alasdair Urquhart for pointing me to Kyburg's comments on Pólya's essay.

the methodology of doing mathematics that includes the use of computations for:

(1) Gaining insight and intuition.
(2) Discovering new patterns and relationships.
(3) Using graphical displays to suggest underlying mathematical principles.
(4) Testing and especially falsifying conjectures.
(5) Exploring a possible result to see if it is worth a formal proof.
(6) Suggesting approaches for formal proof.
(7) Replacing lengthy hand derivations with computer-based derivations.
(8) Confirming analytically derived results.

In philosophical discourse it is common to distinguish between discovery and justification; that is, to distinguish the process of formulating definitions and conjectures from the process of justifying mathematical claims as true. Both types of activities are involved in the list above.

On the discovery side, brute calculation can be used to suggest or test general claims. Around the turn of the nineteenth century, Gauss conjectured the prime number theorem after calculating the density of primes among the first tens of thousands of natural numbers; such number-theoretic and combinatorial calculations can now be performed quickly and easily. One can evaluate a real-valued formula to a given precision, and then use an "Inverse Symbolic Calculator" to check the result against extensive databases to find a simplified expression. Similarly, one can use Neil Sloan's "On-Line Encyclopedia of Integer Sequences" to identify a sequence of integers arising from a particular calculation. These, and more refined methods along these lines, are described in (Bailey and Borwein, 2005). Numerical methods can also be used to simulate dynamical systems and determine their global properties, or to calculate approximate solutions to systems of differential equations where no closed-form solution is available. Graphical representations of data are often useful in helping us understand such systems.

Computers are also used to justify mathematical claims. Computational methods had been used to establish Fermat's last theorem for the first four million exponents by the time its general truth was settled in 1995. The Riemann hypothesis has been established for all complex numbers with imaginary part less than 2.4 trillion, though the general claim remains unproved. Appel and Haken's 1977 proof of the four-color theorem is a well-known example of a case in which brute force combinatorial enumeration played an essential role in settling a longstanding open problem. Thomas Hales's 1998 proof of the Kepler conjecture, which asserts that the optimal density of sphere packing is achieved by the familiar hexagonal lattice packing, has a similar character: the proof used computational methods to obtain an exhaustive database of several thousand "tame" graphs, and then to bound nonlinear and linear optimization problems associated with these graphs. (This pattern of reducing a problem to one that can be solved by combinatorial enumeration and numerical methods is now common in discrete geometry.) Computer algebra systems like *Mathematica* or *Maple* are used, in more mundane ways, to simplify complex expressions that occur in ordinary mathematical proofs. Computers are sometimes even used to find justifications that can be checked by hand; for example, William McCune used a theorem prover named EQP to show that

a certain set of equations serve to axiomatize Boolean algebras (McCune, 1997), settling a problem first posed by Tarski.

The increasing reliance on extensive computation has been one impetus in the development of methods of formal verification. It has long been understood that much of mathematics can be formalized in systems like Zermelo-Fraenkel set theory, at least in principle; in recent decades, computerized "proof assistants" have been developed to make it possible to construct formal mathematical proofs in practice. At present, the efforts required to verify even elementary mathematical theorems are prohibitive. But the systems are showing steady improvement, and some notable successes to date suggest that, in the long run, the enterprise will become commonplace. Theorems that have been verified, to date, include Gödel's first incompleteness theorem, the prime number theorem, the four color theorem, and the Jordan curve theorem (see Wiedijk, 2006). Hales has launched a project to formally verify his proof of the Kepler conjecture, and Georges Gonthier has launched a project to verify the Feit-Thompson theorem. These are currently among the most ambitious mathematical verification efforts underway.

Thus far, I have distinguished the use of computers to suggest plausible mathematical claims from the use of computers to verify such claims. But in many cases, this distinction is blurred. For example, the Santa Fe Institute is devoted to the study of complex systems that arise in diverse contexts ranging from physics and biology to economics and the social sciences. Computational modeling and numeric simulation are central to the institute's methodology, and results of such "experiments" are often held to be important to understanding the relevant systems, even when they do not yield precise mathematical hypotheses, let alone rigorous proofs.

Computers can also be used to provide inductive "evidence" for precise mathematical claims, like the claim that a number is prime. For example, a probabilistic primality test due to Robert Solovay and Volker Strassen works as follows.[1] For each natural number n, there is an easily-calculable predicate, $P_n(a)$, such that if n is prime then $P_n(a)$ is always true, and if n is not prime then at least half the values of a less than n make $P_n(a)$ false. Thus, one can test the primality of n by choosing test values $a_0, a_1, a_2, \ldots$ less than n at random; if $P_n(a_i)$ is true for a large number of tests, it is "virtually certain" that n is prime.[2]

In sum, the new experimental methodology relies on explicit or implicit claims as to the utility of computational methods towards obtaining, verifying, and confirming knowledge; suggesting theorems and making conjectures plausible; and providing insight and understanding. These claims have a patent epistemological tinge, and so merit philosophical scrutiny.[3] For example, one can ask:

- In what sense do calculations and simulations provide "evidence" for mathematical hypotheses? Is it rational to act on such evidence?
- How can computers be used to promote mathematical understanding?

---

[1]A probabilistic test later developed by Michael Rabin, based on a deterministic version by Gary Miller, has similar properties and is now more commonly used.

[2]This can be made mathematically precise. For example, suppose a 100-digit number is chosen at random from a uniform distribution. Number-theoretic results show that there is a non-negligible prior probability that n is prime. If one then chooses $a_0, \ldots, a_l$ at random, one can show that the probability that n is prime given that $P_n(a_i)$ holds for every i approaches 1, quickly, as l increases.

[3]The list is not exhaustive. For example, uses of computers in storing, organizing, and communicating mathematical knowledge also raise issues that merit philosophical attention.

- Does a proof obtained using extensive computation provide mathematical certainty? Is it really a proof?
- Is knowledge gained from the use of a probabilistic primality test any less certain or valuable than knowledge gained from a proof? What about knowledge gained from simulation of a dynamical system?
- Does formal verification yield absolute, or near absolute, certainty? Is it worth the effort?

As presented, these questions are too vague to support substantive discussion. The first philosophical challenge, then, is to formulate them in such a way that it is clear what types of analytic methods can have a bearing on the answers.

3. The epistemology of mathematics

A fundamental goal of the epistemology of mathematics is to determine the appropriate means of justifying a claim to mathematical knowledge. The problem has a straightforward and generally accepted solution: the proper warrant for the truth of a mathematical theorem is a mathematical proof, that is, a deductive argument, using valid inferences, from axioms that are immediately seen to be true. Much of the effort in the philosophy of mathematics has gone towards determining the appropriate inferences and axioms, or explaining why knowledge obtained in this way is worth having. These issues will not be addressed here.

There are at least two ways in which one may wish to broaden one's epistemological scope, neither of which denies the correctness or importance of the foregoing characterization. For one thing, one may want to have a philosophical account of warrants for mathematical knowledge that takes into consideration the fact that these warrants have to be recognized by physically and computationally bounded agents. A formal proof is an abstract object, albeit one that we may take to be reasonably well instantiated by symbolic tokens on a physical page. But proofs in textbooks and mathematical journals are somewhat further removed from this idealization: they are written in a regimented but nonetheless imprecise and open-ended fragment of natural language; the rules of inference are not spelled out explicitly; inferential steps are generally much larger than the usual formal idealizations; background knowledge is presupposed; and so on. Few can claim to have verified any complex theorem from first principles; when reading a proof, we accept appeals to theorems we have learned from textbooks, journal articles, and colleagues. The logician's claim is that the informal proof serves to indicate the existence of the formal idealization, but the nature of this "indication" is never spelled out precisely. Moreover, we recognize that proofs can be mistaken, and often express degrees of faith depending on the nature of the theorem, the complexity of proof, the methods that have been used to prove it, and the reliability of the author or the authorities that are cited. Just as mathematical logic and traditional philosophy of mathematics provides us with an idealized model of a perfect, gapless deduction, we may hope to model the notion of an "ordinary" proof and ask: when is it rational to accept an ordinary proof as indicating the existence of an idealized one?[4]

To explore this issue, one need not conflate the attempt to provide an idealized account of the proper warrants for mathematical knowledge with the attempt to provide an account of the activities we may rationally pursue in service of this

---

[4]For an overview of issues related to the "surveyability" of proofs, see (Bassler, 2006).

ideal, given our physical and computational limitations. It is such a conflation that has led Tymozcko (1979) to characterize mathematics as a quasi-empirical science, and Fallis (1997, 2002) to wonder why mathematicians refuse to admit inductive evidence in mathematical proofs. The easy answer to Fallis's bemusement is simply that inductive evidence is not the right sort of thing to provide mathematical knowledge, as it is commonly understood. But when their remarks are taken in an appropriate context, Tymoczko and Fallis do raise the reasonable question of how (and whether) we can make sense of mathematics, more broadly, as an activity carried out by agents with bounded resources. This question should not be dismissed out of hand.

A second respect in which one may wish to broaden one's epistemological ambitions is to extend the analysis to value judgments that go beyond questions of correctness. On the traditional view, the role of a proof is to warrant the truth of the resulting theorem, in which case, all that matters is that the proof is correct. But when it comes to proofs based on extensive computation, a far more pressing concern is that they do not provide the desired mathematical insight. Indeed, the fact that proofs provide more than warrants for truth becomes clear when one considers that new proofs of a theorem are frequently judged to be important, even when prior proofs have been accepted as correct. We tend to feel that raw computation is incapable of delivering the type of insight we are after:

> ...it is common for people first starting to grapple with computers to make large-scale computations of things they might have done on a smaller scale by hand. They might print out a table of the first 10,000 primes, only to find that their printout isn't something they really wanted after all. They discover by this kind of experience that what they really want is usually not some collection of "answers"— what they want is *understanding*. (Thurston, 1994, page 162)

We often have good intuitions as to the ways that mathematical developments constitute conceptual advances or further understanding. It is therefore reasonable to ask for a philosophical theory that can serve to ground such assessments, and account for the more general epistemological criteria by which such developments are commonly judged.

In sum, questions about the use of computers in mathematics that seem reasonable from a pre-theoretic perspective push us to extend the traditional philosophy of mathematics in two ways: first, to develop theories of mathematical evidence, and second, to develop theories of mathematical understanding. In the next two sections, I will consider each of these proposals, in turn.

4. Theories of mathematical evidence

We have seen that some issues regarding the use of computers in mathematics hinge on assessments of the "likelihood" that a mathematical assertion is true:

- a probabilistic primality test renders it highly likely that a number is prime;
- numeric simulations can render it plausible that a hypothesis is true;
- formal verification can render it nearly certain that a theorem has a correct proof.

Since judgments like these serve to guide our actions, it is reasonable to ask for a foundational framework in which they can be evaluated. Such a framework may

also have bearing on the development of computational support for mathematics; for example, systems for automated reasoning and formal verification often attempt to narrow the search space by choosing the most "plausible" or "promising" paths.

Probabilistic notions of likelihood, evidence, and support have long played a role in characterizing inductive reasoning in the empirical sciences, and it is tempting to carry these notions over to the mathematical setting. However, serious problems arise when one tries to do so. Roughly speaking, this is because any mathematical assertion is either true, in which case it holds with probability 1, or false, in which case it holds with probability 0, leaving no room for values in between.

Put more precisely, classical approaches to probability model "events" as measurable subsets of a space whose elements are viewed as possible outcomes of an experiment, or possible states of affairs. The laws of probability dictate that if an event A entails an event B, in the sense that $A \subseteq B$, then the probability of A is less than or equal to the probability of B. In particular, if a property holds of all possible outcomes, the set of all possible states of affairs that satisfy that property has probability 1. So, to assign a probability other than 1 to an assertion like "5 is prime," one needs to characterize the primality of 5 as a property that may or may not hold of particular elements of a space. But 5 is prime, no matter what, and so it is difficult to imagine what type of space could reasonably model the counterfactual case. I may declare X to be the set $\{0, 1\}$, label 0 the state of affairs in which 5 is not prime, label 1 the state of affairs in which 5 is prime, and then assign $\{0\}$ and $\{1\}$ each a probability $1/2$. But then I have simply modeled a coin flip; the hard part is to design a space that can convincingly be argued to serve as an appropriate guide to behavior in the face of uncertainty.

It is tempting to resort to a Bayesian interpretation, and view probabilities as subjective degrees of belief. I can certainly claim to have a subjective degree of belief of $1/2$ that 5 is not prime; but such claims cannot play a role in a theory of rationality until they are somehow linked to behavior. For example, it is common to take the outward signs of a subjectively held probability to be the willingness to bet on the outcome of an experiment (or the result of determining the true state of affairs) with corresponding odds. In that case, F. P. Ramsey and Bruno de Finetti have noted that the dictates of rationality demand that, at the bare minimum, subjective assignments should conform to the laws of probability, on pain of having a clever opponent "make book" by placing a system of bets that guarantees him or her a profit no matter what transpires. But such coherence criteria still depend, implicitly, on having a model of a space of possible outcomes, against which the possibility of book can be judged. So one has simply shifted the problem to that of locating a notion of coherence on which it is reasonable to have a less-than-perfect certainty in the fact that 5 is prime; or, at least, to develop a notion of coherence for which there is anything interesting to say about such beliefs.

The challenge of developing theories of rationality that do not assume logical omniscience is not limited to modeling mathematical beliefs; it is just that the difficulties involved in doing so are most salient in mathematical settings. But the intuitions behind ascriptions of mathematical likelihood are often so strong that some have been encouraged to overcome these difficulties. For example, Pólya (1941) discusses a claim, by Euler, that it is nearly certain that the coefficients of two analytic expressions agree, because the claim can easily be verified in a number

of specific cases. Pólya then suggested that it might be possible to develop a "qualitative" theory of mathematical plausibility to account for such claims. (See also the other articles in Pólya 1984, and Kyburg's remarks at the end of that volume.) Ian Hacking (1967), I. J. Good (1977), and, more recently, Haim Gaifman (2004) have proposed ways of making sense of probability judgments in mathematical settings. David Corfield (2003) surveys such attempts, and urges us to take them seriously.

Gaifman's proposal is essentially a variant of the trivial "5 is prime" example I described above. Like Hacking, Gaifman takes sentences (rather than events or propositions) to bear assignments of probability. He then describes ways of imposing constraints on an agent's deductive powers, and asks only that ascriptions of probability be consistent with the entailments the agent can "see" with his or her limited means. If all I am willing to bet on is the event that 5 is prime and I am unable or unwilling to invest the effort to determine whether this is the case, then, on Gaifman's account, any assignment of probability is "locally" consistent with my beliefs. But Gaifman's means of incorporating closure under some deductive entailments allows for limited forms of reasoning in such circumstances. For example, if I judge it unlikely that all the random values drawn to conduct a probabilistic primality test are among a relatively small number of misleading witnesses, and I use these values to perform a calculation that certifies a particular number as prime, then I am justified in concluding that it is likely that the number is prime. I may be wrong about the chosen values and hence the conclusion, but at least, according to Gaifman, there is a sense in which my beliefs are locally coherent.

But does this proposal really address the problems raised above? Without a space of possibilities or a global notion of coherent behavior, it is hard to say what the analysis does for us. Isaac Levi (1991, 2004) clarifies the issue by distinguishing between theories of *commitment* and theories of *performance*. Deductive logic provides theories of the beliefs a rational agent is ideally committed to, perhaps on the basis of other beliefs that he or she is committed to, independent of his or her ability to recognize those commitments. On that view, it seems unreasonable to say that an agent committed to believing "A" and "A implies B" is not committed to believing "B," or that an agent committed to accepting the validity of basic arithmetic calculations is not committed to the consequence of those calculations.

At issue, then, are questions of performance. Given that physically and computationally bounded agents are not always capable of recognizing their doxastic commitments, we may seek general *procedures* we can follow to approximate the ideal. For example, given bounds on the resources we are able to devote to making a certain kind of decision, we may seek procedures that provide correct judgments most of the time, and minimize errors. Can one develop such a theory of "useful" procedures? Of course! This is exactly what theoretical computer science does. Taken at face value, the analysis of a probabilistic primality test shows that if one draws a number at random from a certain distribution, and a probabilistic primality test certifies the number as prime, then with high probability the conclusion is correct. Gaifman's theory tries to go one step further and explain why it is rational to accept the result of a test in a *specific* case where it provides a false answer. But it is not clear that this adds anything to our understanding of rationality, or provides a justification for using the test that is better than the fact that the procedure is efficient and usually reliable.

When it comes to empirical events, we have no problem taking spaces of possibilities to be implicit in informal judgments. Suppose I draw a marble blindly from an urn containing 500 black marbles and 500 white marbles, clasp it in my fist, and ask you to calculate the probability that the marble I hold is black. The question presupposes that I intend for you to view the event as the result of a draw of a ball from the urn. Without a salient background context, the question as to the probability that a marble I clasp in my fist is black is close to meaningless.

In a similar fashion, the best way to understand an ascription of likelihood to a mathematical assertion may be to interpret it as a judgment as to the likelihood that a certain manner of proceeding will, in general, yield a correct result. Returning to Pólya's example, Euler seems to be making a claim as to the probability that two types of calculations, arising in a certain way, will agree in each instance, given that they agree on sufficiently many randomly or deterministically chosen test cases. If we assign a probability distribution to a space of such calculations, there is no conceptual difficulty involved in making sense of the claim. Refined analyses may try to model the types of calculations one is "likely" to come across in a given domain, and the outcome of such an analysis may well support our intuitive judgments. The fact that the space in question may be vague or intractable makes the problem little different from those that arise in ordinary empirical settings.[5]

Along the same lines, the question as to the probability of the correctness of a proof that has been obtained or verified with computational means is best understood as a question as to the reliability of the computational methods or the nature of the verification. Here, too, the modeling issues are not unlike those that arise in empirical contexts. Vendors often claim "five-nines" performance for fault-tolerant computing systems, meaning that the systems can be expected to be up and running 99.999% of the time. Such judgments are generally based on past performance, rather than on any complex statistical modeling. That is not to say that there are not good reasons to expect that past performance is a good predictor, or that understanding the system's design can't bolster our confidence. In a similar manner, formal modeling may, pragmatically, have little bearing on our confidence in computational methods of verification.

In sum, there are two questions that arise with respect to theories of mathematical evidence: first, whether *any* philosophical theory of mathematical plausibility can be put to significant use in any of the domains in which the notions arise; and

---

[5]Another nice example is given by (Wasserman, 2004, Example 11.10), where statistical methods are used to estimate the value of an integral that is too hard to compute. As the discussion after that example suggests, the strategy of suppressing intractable information is more congenial to a classical statistician than to a Bayesian one, who would insist, rather, that all the relevant information should be reflected in one's priors. This methodological difference was often emphasized by I. J. Good, though Wasserman and Good draw opposite conclusions. Wasserman takes the classical statistician's ability to selectively ignore information to provide an advantage in certain contexts: "To construct procedures with guaranteed long run performance, ... use frequentist methods." In contrast, Good takes the classical statistician's need to ignore information to indicate the fragility of those methods; see the references to the "statistician's stooge" in (Good, 1983). I am grateful to Teddy Seidenfeld for bringing these references to my attention.

I have already noted, above, that Good (1977) favors a Bayesian approach to assigning probabilities to outcomes that are determined by calculation. But, once again, Levi's distinction between commitment and performance is helpful: what Good seems to propose is a theory that is capable of modeling "reasonable" behavior in computationally complex circumstances, without providing a normative account of what such behavior is supposed to achieve.

second, if so, whether a fundamentally different concept of rationality is needed. It is possible that proposals like Pólya's, Hacking's, and Gaifman's will prove useful in providing descriptive accounts of human behavior in mathematical contexts, or in designing computational systems that serve mathematical inquiry. But this is a case that needs to be made. Doing so will require, first, a clearer demarcation of the informal data that the philosophical theories are supposed to explain, and second, a better sense of what it is that we want the explanations to do.

5. Theories of mathematical understanding

In addition to notions of mathematical evidence, we have seen that uses of computers in mathematics also prompt evaluations that invoke notions of mathematical understanding. For example:

- results of numeric simulation can help us understand the behavior of a dynamical system;
- symbolic computation can help shed light on an algebraic structure;
- graphical representations can help us visualize complex objects and thereby grasp their properties (see Mancosu, 2005).

Such notions can also underwrite negative judgments: we may feel that a proof based on extensive computation does not provide the insight we are after, or that formal verification does little to promote our understanding of a theorem. The task is to make sense of these assessments.

But the word "understanding" is used in many ways: we may speak of understanding a theory, a problem, a solution, a conjecture, an example, a theorem, or a proof. Theories of mathematical understanding may be taken to encompass theories of explanation, analogy, visualization, heuristics, concepts, and representations. Such notions are deployed across a wide range of fields of inquiry, including mathematics, education, history of mathematics, cognitive science, psychology, and computer science. In short, the subject is a sprawling wilderness, and most, if not all, of the essays in this collection can be seen as attempts to tame it. (See also the collection Mancosu *et al.*, 2005.)

Similar topics have received considerably more attention in the philosophy of science, but the distinct character of mathematics suggests that different approaches are called for. Some have expressed skepticism that anything philosophically interesting can be said about mathematical understanding, and there is a tradition of addressing the notion only obliquely, with hushed tones and poetic metaphor. This is unfortunate: I believe it is possible to develop fairly down-to-earth accounts of key features of mathematical practice, and that such work can serve as a model for progress where attempts in the philosophy of science have stalled. In the next essay, I will argue that philosophical theories of mathematical understanding should be cast in terms of analyses of the types of mathematical abilities that are implicit in common scientific discourse where notions of understanding are employed. Here, I will restrict myself to some brief remarks as to the ways in which recent uses of computers in mathematics can be used to develop such theories.

The influences between philosophy and computer science should run in both directions. Specific conceptual problems that arise in computer science provide effective targets for philosophical analysis, and goals like that of verifying common mathematical inferences or designing informative graphical representations provide concrete standards of success, against which the utility of an analytic framework can

be evaluated. There is a large community of researchers working to design systems that can carry out mathematical reasoning effectively; and there is a smaller, but significant, community trying to automate mathematical discovery and concept formation (see e.g. Colton *et al.*, 2000). If there is any domain of scientific inquiry for which one might expect the philosophy of mathematics to play a supporting role, this is it. The fact that the philosophy of mathematics provides virtually no practical guidance in the appropriate use of common epistemic terms may lead some to wonder what, exactly, philosophers are doing to earn their keep.

In the other direction, computational methods that are developed towards attaining specific goals can provide clues as to how one can develop a broader philosophical theory. The data structures and procedures that are effective in getting computers to exhibit the desired behavior can serve to direct our attention to features of mathematics that are important to a philosophical account.

In (Avigad, 2006), I addressed one small aspect of mathematical understanding, namely, the process by which we understand the text of an ordinary mathematical proof. I discussed ways in which efforts in formal verification can inform and be informed by a philosophical study of this type of understanding. In the next essay, I will expand on this proposal, by clarifying the conception of mathematical understanding that is implicit in the approach, and discussing aspects of proofs in algebra, analysis, and geometry in light of computational developments. In focusing on formal verification, I will be dealing with only one of the many ways in which computers are used in mathematics. So the effort, if successful, provides just one example of the ways that a better interaction between philosophical and computational perspectives can be beneficial to both.

6. Final thoughts

I have surveyed two ways in which the philosophy of mathematics may be extended to address issues that arise with respect to the use of computers in mathematical inquiry. I may, perhaps, be accused of expressing too much skepticism with respect to attempts to develop theories of mathematical evidence, and excessive optimism with respect to attempts to develop theories of mathematical understanding. Be that as it may, I would like to close here with some thoughts that are relevant to both enterprises.

First, it is a mistake to view recent uses of computers in mathematics as a source of philosophical puzzles that can be studied in isolation, or resolved by appeal to basic intuition. The types of questions raised here are only meaningful in specific mathematical and scientific contexts, and a philosophical analysis is only useful insofar as it can further such inquiry. Ask not what the use of computers in mathematics can do for philosophy; ask what philosophy can do for the use of computers in mathematics.

Second, issues regarding the use of computers in mathematics are best understood in a broader epistemological context. Although some of the topics explored here have become salient with recent computational developments, none of the core issues are specific to the use of the computer *per se*. Questions having to do with the pragmatic certainty of mathematical results, the role of computation in mathematics, and the nature of mathematical understanding have a much longer provenance, and are fundamental to making sense of mathematical inquiry. What we need now

is *not* a philosophy of computers in mathematics; what we need is simply a better philosophy of mathematics.

References

AVIGAD, Jeremy, 'Mathematical method and proof', *Synthese*, 153:105-159, 2006.

BAILEY, David and BORWEIN, Jonathan (2005), 'Experimental mathematics: examples, methods and implications', *Notices of the American Mathematical Society*, 52, 502–514.

BASSLER, O. Bradley (2006), 'The surveyability of mathematical proof: A historical perspective', *Synthese* 148:99–133.

BORWEIN, Jonathan and BAILEY, David (2004), *Mathematics by experiment: plausible reasoning in the 21st century* (Natick, MA: A K Peters Ltd.).

BORWEIN, Jonathan, BAILEY, David, and GIRGENSOHN, Roland (2004), *Experimentation in mathematics: computational paths to discovery* (Natick, MA: A K Peters Ltd.).

COLTON, Simon, BUNDY, Alan, and WALSH, Toby (2000), 'On the notion of interestingness in automated mathematical discovery', *International Journal of Human-Computer Studies*, 53, 351–365.

CORFIELD, David (2003), *Towards a philosophy of real mathematics* (Cambridge: Cambridge University Press).

FALLIS, Don (1997), 'The epistemic status of probabilistic proof', *Journal of Philosophy*, 94, 165–186.

——— (2002), 'What do mathematicians want?: probabilistic proofs and the epistemic goals of mathematicians', *Logique et Analyse*, 45, 373–388.

GAIFMAN, Haim (2004), 'Reasoning with limited resources and assigning probabilities to arithmetical statements', *Synthese*, 140, 97–119.

GOOD, I. J. (1977), 'Dynamic probability, computer chess, and the measurement of knowledge', in E. W. Elcock and Donald Michie, eds., *Machine Intelligence 8* (New York: John Wiley & Sons), 139–150. Reprinted in Good (1983), 106–116.

GOOD, I. J. (1983), *Good thinking: The foundations of probability and its applications* (Minneapolis: University of Minnesota Press).

HACKING, Ian (1967), 'A slightly more realistic personal probability', *Philosophy of Science*, 34, 311–325.

LEVI, Isaac (1991), *The fixation of belief and its undoing* (Cambridge: Cambridge University Press).

——— (2004), 'Gaifman', *Synthese*, 140, 121–134.

MANCOSU, Paolo (2005), 'Visualization in logic and mathematics', in Mancosu *et al.* (2005).

MANCOSU, Paolo, JØRGENSEN, Klaus Frovin, and PEDERSEN, Stig Andur (2005), *Visualization, explanation and reasoning styles in mathematics* (Dordrecht: Springer-Verlag).

McCUNE, William (1997), 'Solution of the Robbins problem', *Journal of Automated Reasoning*, 19, 263–276.

PÓLYA, George (1941), 'Heuristic reasoning and the theory of probability', *American Mathematical Monthly*, 48, 450–465.

PÓLYA, George (1984), *Collected papers. Vol. IV: Probability; combinatorics; teaching and learning in mathematics*, (Cambridge, MA: MIT Press). Edited by Gian-Carlo Rota, M. C. Reynolds and R. M. Shortt.

THURSTON, William P. (1994), 'On proof and progress in mathematics', *Bulletin of the American Mathematical Society*, 30, 161–177.

TYMOZCKO, Thomas (1979), 'The four-color problem and its philosophical significance', *Journal of Philosophy*, 76, 57–83. Reprinted in (Tymoczko, 1998), pages 243–266.

TYMOCZKO, Thomas (ed.) (1998), *New directions in the philosophy of mathematics*, expanded edition, (Princeton, NJ: Princeton University Press).

WASSERMAN, Larry (2004), *All of statistics* (New York: Springer-Verlag).

WIEDIJK, Freek (2006), *The seventeen provers of the world* (Berlin: Springer-Verlag).