Formal Methods and the Epistemology of Mathematics

Jeremy Avigad April 24, 2020

Carnegie Mellon University





















"Philosophy of mathematics is dead. Philosophy of mathematics remains dead. And we have killed it."

"Philosophy of mathematics is dead. Philosophy of mathematics remains dead. And we have killed it."

"Not only is philosophy of mathematics dead, but just try to find a plumber on weekends."

"Philosophy of mathematics is dead. Philosophy of mathematics remains dead. And we have killed it."

"Not only is philosophy of mathematics dead, but just try to find a plumber on weekends."

"Reports of its death are greatly exaggerated."

"Philosophy of mathematics is dead. Philosophy of mathematics remains dead. And we have killed it."

"Not only is philosophy of mathematics dead, but just try to find a plumber on weekends."

"Reports of its death are greatly exaggerated."

(Apologies to Friedrich Nietzsche, Woody Allen, and Mark Twain.)

Twentieth century philosophy of mathematics

The central questions, in the analytic tradition:

- 1. What is mathematical knowledge, and what justifies a claim to mathematical knowledge?
- 2. What sorts of things are mathematical objects, and how do we (or can we, or should we) come to have knowledge of them?

Nobody seems to care any more.

Twentieth century philosophy of mathematics

Why not?

- 1. For close to a century, there has been remarkable consensus about how to talk about mathematical objects, and about the standards of correctness for proofs.
- 2. Formal logic provides good accounts of these.
- 3. Of course, we learn about mathematical objects from our parents, teachers, and textbooks.
- Answering more fundamental questions (why these objects? why these rules?) requires hard work and sensitivity to the mathematics, not clever rhetoric.

Twentieth century philosophy of mathematics

There are normative assessments in mathematics that go beyond questions of correctness:

- Theorems can be interesting or illuminating or surprising.
- Concepts and methods can be powerful.
- Historical developments can be important.
- Proofs can be elegant.
- Questions can be natural.

We do care about these judgments.

Twenty-first century philosophy of mathematics

I'll argue for these claims:

- 1. There are philosophical questions about mathematics that are interesting and important beyond academic philosophy.
- 2. Contemporary developments in logic and computer science offer new analytic tools.
- 3. The new questions shed light on the old questions, and make it possible to address them in substantial and satisfying ways.

Outline

- motivating questions
- intuitions
- formal methods in mathematics
- from formal methods to epistemology

The problem of multiple proofs

On the standard account, the value of a mathematical proof is that it warrants the truth of the resulting theorem.

Why, then, do we often value a new proof of a previous established theorem?

For example, Gauss published six proofs of the law of quadratic reciprocity in his lifetime, and left us two unpublished versions as well.

Franz Lemmermeyer has documented 246 proofs through 2013. (The list, with references, is available online.)

The problem of multiple proofs

This question not new. For example:

It might be said: "—that every proof, even of a proposition which has already been proved, is a contribution to mathematics". But why is it a contribution if its only point was to prove the proposition? Well, one can say: "the new proof shews (or makes) a new connexion". — Wittgenstein, Remarks on the Foundations of Mathematics, III–60

Indeed, it is *not* a great mystery. There is a lot we can say about what we learn from different proofs.

But the philosophy of mathematics has had relatively little to say about this.

The problem of conceptual possibility

It is often said that some mathematical advance was "made possible" by a prior conceptual development.

For example, Riemann's introduction of the complex zeta function and the use of complex analysis made it possible for Hadamard and de la Vallée Poussin to prove the prime number theorem in 1896.

What is the sense of "possibility" here?

Intuition: a certain *understanding* guides us.

Proposition 16

In any triangle, if one of the sides be produced, the exterior angle is greater than either of the interior and opposite angles.

Let *ABC* be a triangle, and let one side of it *BC* be produced to *D*;

I say that the exterior angle *ACD* is greater than either of the interior and opposite angles *CBA*, *BAC*.

Let AC be bisected at E, [I. 10] and let BE be joined and produced in a straight line to F;

let EF be made equal to BE, [1.3]

let FC be joined, [Post. 1]

and let AC be drawn through to G. [Post. 2]

Then, since AE is equal to EC, and BE to EF,

the two sides AE, EB are equal to the two sides

CE, EF respectively;

and the angle AEB is equal to the angle FEC, for they are vertical angles. [I. 15]

Therefore the base AB is equal to the base FC, and the triangle ABE is equal to the triangle CFE,

and the remaining angles are equal to the remaining angles respectively, namely those which the equal sides subtend; [1. 4]

therefore the angle BAE is equal to the angle ECF.

















By side-angle-side, $\triangle AEB \equiv \triangle CEF$.



By side-angle-side, $\triangle AEB \equiv \triangle CEF$. So $\angle BAC = \angle ACF$.



By side-angle-side, $\triangle AEB \equiv \triangle CEF$. So $\angle BAC = \angle ACF$. Clearly $\angle ACD > \angle ACF$.



By side-angle-side, $\triangle AEB \equiv \triangle CEF$. So $\angle BAC = \angle ACF$. Clearly $\angle ACD > \angle ACF$. So $\angle ACD > \angle BAC$.



By side-angle-side, $\triangle AEB \equiv \triangle CEF$. So $\angle BAC = \angle ACF$. Clearly $\angle ACD > \angle ACF$. So $\angle ACD > \angle BAC$.

But why is it clear that $\angle ACD > \angle ACF$?

On a standard account, a proof is correct if each inference can be expanded to a formal derivation.

Such formal derivations can be extremely long. Even a single error renders one invalid.

How can ordinary mathematical proofs reliably warrant the existence of something so complex and fragile?

Why doesn't mathematics fall apart?

The role of abstraction

The value of algebraic reasoning is often attributed to its generality.

For example, the axiomatization of *groups* in the nineteenth century unified instances in Galois theory, number theory, and geometry.

But sometimes abstraction is valued even when there is only one instance.

In 1871, Dedekind introduced the notion of an *ideal* in a number ring. In 1882 he and Weber generalized it to rings of functions.

But Dedekind clearly thought the notion was useful, even before the 1882 generalization.

Why?

The use of computers in proofs

Kenneth Appel and Wolfgang Haken used extensive computation to prove the four-color theorem in 1976.

Thomas Hales announced a proof of the Kepler conjecture in 1998, again using extensive computation.

Propositional satisfiability provers are being used to solve combinatorial problems, in some cases, producing proofs that are terabytes long. (More on this later.)

Does this make for good mathematics?

Motivating questions

What the questions have in common:

- They have a general epistemological character.
- They have to do with mathematical *understanding*. (Proofs and concepts convey understanding, and require understanding.)
- They raise normative questions. (What do we value? What makes for good mathematics?)
- We have some intuitions.
- We care about the answers.

This doesn't guarantee that there is room for philosophy here.

But it should encourage us to take a look.

Intuitions

Responses will likely require talking about:

- methods
- concepts
- representations
- cognitive effort

Let's consider each, in turn.

Methods

Mathematical knowledge is often cast as *propositional* knowledge, like definitions and theorems.

But *understanding* seems to require something more dynamic, a kind of *procedural knowledge*.

Thinking means passing through epistemic states. Understanding guides thought.

One approach: talk about *methods*, i.e. heuristic, fallible, procedures for solving problems, searching for proofs, verifying inferences, etc.
Methods

Another approach: talk about *abilities*, or capacities for thought.

Understanding involves:

- Being able to recognize the nature of the objects and questions before us.
- Being able to marshall the relevant background knowledge and information.
- Being able to traverse space the of possibilities before us in a fruitful way.
- Being able to identify features of the context that help us cut down complexity.

In the psychological literature, concepts are sometimes thought of in terms of categorization (e.g. prototypes and exemplars).

From a logical perspective, a concept is given by a definition, in a suitable formal language.

These don't work so well for the philosophy of mathematics.

What does it mean to understand the concept of a *group*? Or the concept of a *function*? Or the concept of a *Riemannian manifold*?

Mathematical concepts have some interesting features:

- Membership is often sharply defined.
- Mathematical concepts evolve over time.
- Understanding a concept admits degrees.
- Various things can "improve our understanding" of a concept.
- One can speak of implicit uses of a concept.

One solution: think of a mathematical concept as a bundle of abilities.

For example, the group concept includes:

- Knowing the definition of a group.
- Knowing common examples of groups, and being able to recognize implicit group structures when it is fruitful to do so.
- Knowing how to construct groups from other groups or other structures, in fruitful ways.
- Recognizing that there are different kinds of groups (abelian, nilponent, solvable, finite vs. infinite, continuous vs. discrete) and being able/prone to make these distinctions.
- Knowing various theorems about groups, and when and how to apply them.

This renders "the group concept," for example, vague and open-ended.

But the notion *is* vague and open-ended:

- We can talk about student understanding.
- We can talk about the role of the concept in contemporary mathematics.
- We can talk about the historical development.

The proposal suggests that we can make our talk more precise by being more precise about the abilities (or methods, or capacities) we have in mind.

Representations

In philosophy of mind, sometimes a concept is taken to be some sort of mental *representation*, maybe in a language of thought.

Understanding seems to have something to do with having the right representations.

In contemporary philosophy of mathematics, there has been a lot of interest in the nature of representations, especially diagrammatic representations.

Ken Manders has advocated using the word *artifacts*. Roy Wagner likes *presentations*.

What is important is not what they represent, but what we can do with them.

Cognitive effort

As cognitive agents, we have limited time, energy, memory, processing capacity.

We value developments that make things easier.

But how can we measure difficulty?

- Computer science: algorithmic complexity
- Logic: descriptive complexity, length of proof
- Experimental psychology: timing tasks

We can consider the number of pages in a proof, the number of symbols in an expression, or the number of steps in a calculation.

But we need better ways of talking about cognitive difficulty.

Recap

Motivating questions:

- the problem multiple proofs
- the problem of conceptual possibility
- the nature of diagrammatic inference
- the problem of reliability
- the role of abstraction
- the use of computers in proofs

We need to understand:

- methods
- concepts
- representations
- cognitive effort

Goals

We are looking for a philosophical account that:

- is clear, precise, and internally coherent
- accords with our intuitions
- fits the data (what we see in mathematics)
- can inform (and can be informed by) other pursuits:
 - history of mathematics
 - interactive theorem proving and automated reasoning
 - psychology and cognitive science
 - mathematics education
 - mathematics itself

I will discuss formal methods, as one source of insight.

Outline

- motivating questions
- intuitions
- formal methods in mathematics
- from formal methods to epistemology

Disclaimers

There are some very good young people thinking about questions like the ones I have raised.

This is not a survey.

(See the references at the end of this talk.)

Computer science is certainly not the *only* promising source for insight.

(Go back two slides.)

Formal methods in computer science

Formal methods are used for

- specifying,
- developing, and
- verifying

complex hardware and software systems.

They rely on:

- formal languages to make assertions and express constraints,
- formal semantics to specify intended meaning, and
- formal rules of inference to verify claims and carry out search.

Formal methods in computer science

In short, they are used to

- say things,
- find things,
- and check things.

Examples:

- Model checkers search for counterexamples to specifications.
- Interactive theorem provers show that hardware and software designs meet their specifications

Formal methods hold promise for mathematics as well.

I will discuss three domains of application:

- verification
- discovery
- knowledge management

I will convey some of the things we have learned, and are learning.

Interactive Theorem Proving provides one method of verifying mathematical theorems.

Working with a proof assistant, users construct a formal axiomatic proof.

In many systems, this proof object can be extracted and verified independently.

Some systems with substantial mathematical libraries:

- Mizar (set theory)
- HOL (simple type theory)
- Isabelle (simple type theory)
- HOL Light (simple type theory)
- Coq (constructive dependent type theory)
- ACL2 (primitive recursive arithmetic)
- PVS (classical dependent type theory)
- Agda (constructive dependent type theory)
- Metamath (set theory)
- Lean (dependent type theory)

Accomplishments:

- There are good libraries for elementary number theory, real and complex analysis, point-set topology, measure-theoretic probability, abstract algebra, Galois theory, ...
- Lots of big name theorems have been verified.
- There have been some milestones: the verification of the Feit-Thompson theorem, the verification of the Hales' proof of the Kepler conjecture.
- Core mathematicians have begun to get involved.

But the technology is not where we want it to be.

Formalizing anything is a huge pain in the neck.

/- Author: Chris Hughes -/
def legendre_sym (a p : N) (hp : nat.prime p) : Z :=
if (a : zmodp p hp) = 0 then 0 else
if ∃ b : zmodp p hp, b ^ 2 = a then 1 else -1

```
theorem quadratic_reciprocity (hp1 : p % 2 = 1)
    (hq1 : q \% 2 = 1) (hpq : p \neq q) :
  legendre_sym p q hq * legendre_sym q p hp =
    (-1) \cap ((p / 2) * (q / 2)) :=
have hpq0 : (p : zmodp q hq) \neq 0,
  from zmodp.prime_ne_zero _ hp hpq.symm,
have hqp0 : (q : zmodp p hp) \neq 0,
  from zmodp.prime_ne_zero _ hq hpq,
by rw [eisenstein_lemma _ hq1 hp1 hpq0,
  eisenstein_lemma _ hp1 hq1 hqp0, <- _root_.pow_add,
  sum_mul_div_add_sum_mul_div_eq_mul _ hpq0, mul_comm]
```

One strategy for supporting interactive theorem proving is to use automation to justify small inferences.

Ideally, anything "obvious" would be handled automatically.

Some systems, like Isabelle, Coq, and Lean provide *metaprogramming* languages, which allow users to write their domain-specific, small-scale automation.

The moral: mathematical knowledge is not just the theorems and definitions, but also the procedures, methods, and heuristics.

Automated reasoners, constraint solvers, and theorem provers implement powerful search methods.

Alas, applications to mathematics to date are few and far between.

But there have been some.

In 1996, William McCune used an equational theorem prover to prove the Robbins conjecture, which states that a certain system of equations axiomatizes Boolean algebras.

This was posed by a logician, Tarski.

McCune showed that $(w((x^{-1}w)^{-1}z))((yz)^{-1}y) = x$ axiomatizes groups.

Kenneth Kunen showed that this is the shortest such axiom.

Theorem provers and model finders have added to the theory of some (fringe) algebraic structures, like quasigroups and loops.

Ronald Graham once posed the Pythagorean triples problem: Is it possible to color the positive integers red and blue such that there is no monochromatic pythagorean triple $(a^2 + b^2 = c^2)$?

In 2016, Marijn Heule, Oliver Kullmann, and Victor Marek showed:

- There is such a coloring of the integers from 1 to 7,824.
- There is no such coloring of the integers from 1 to 7,825.



They used a propositional satisfiability solver for this purpose.

The proof of the negative result is 200 terabytes long.

It is clearly not possible to tile the plane with unit squares in such a way that no two squares share a common edge.



It is clearly not possible to tile the plane with unit squares in such a way that no two squares share a common edge.



It is clearly not possible to tile the plane with unit squares in such a way that no two squares share a common edge.



It is clearly not possible to tile the plane with unit squares in such a way that no two squares share a common edge.



In 1930, Keller conjectured the corresponding claim holds in all dimensions.

It is clearly not possible to tile the plane with unit squares in such a way that no two squares share a common edge.



In 1930, Keller conjectured the corresponding claim holds in all dimensions.

In 1940, Perron showed that it is true for $n \leq 6$.

It is clearly not possible to tile the plane with unit squares in such a way that no two squares share a common edge.



In 1930, Keller conjectured the corresponding claim holds in all dimensions.

In 1940, Perron showed that it is true for $n \leq 6$.

In 1992, Lagarias and Shor showed that it is *false* for $n \ge 10$.

In 1986 Szabó reduced the conjecture to periodic tilings, and in 1990 Corrádi and Szabó reduced the problem to a question about the existence of a cliques of size 2^n in certain graphs $G_{n,s}$.

The Lagarias and Shor result exhibited a clique of size 2^{10} in $G_{10,2}$. In 2002, Mackey showed the conjecture is false for $n \ge 8$ by exhibiting a clique of size 2^8 in $G_{8,2}$.

In 2017, Kisielewicz showed it is enough to consider $G_{7,6}$ (which has 16^7 nodes).

Very recently, Brakensiek, Heule, Mackey, and Narvéz showed that there is no clique of size 2⁷ there, settling the last case.

They used a SAT solver, with clever reductions and methods to break the symmetry.

There are no 4-cliques in $G_{2,2}$:



A 256-clique in $G_{8,2}$:

*** *** *** *** *** *** *** ***	

Knowledge management

Latex, e-mail, the web, MathOverflow, MathSciNet, and so on have had a strong influence on mathematical research.

Contemporary digital technologies for

- storage,
- search, and
- communication

of mathematical information provide another market for formal methods in mathematics.

Knowledge management

Thomas Hales has launched the *Formal Abstracts* project to encourage mathematicians to write formal abstracts of their papers.

Think of *MathSciNet*, except:

- Every abstract is parsed to a formal language with a precise semantics.
- The library of abstracts be used by search engines.
- Formal proofs can be added later.
- Automated reasoners can use definitions and the results.
- The library is amenable to machine learning.

Knowledge management

The idea is to use controlled natural language.

The user enters something like this:

```
\begin{definition}
\deflabel{greatest element} We say that $y$ is a
\df{greatest\~element} in $R$ iff for all\ $x,\ x \le y$.
\end{definition}
```

```
Let x < y stand for x \le y and x \le y.
```

The Latex renders something like this:

Definition (greatest element)

We say that y is a **greatest element** in R iff for all $x, x \leq y$.

Formal methods hold promise for mathematics.

Three domains of application:

- verification
- discovery
- knowledge management
Lessons

Some of the things we have learned:

- Language and notation are important.
- Definitions are important.
- Representations are important.
- Abstraction is important.
- Structure is important.
- Pattern matching is important.
- Context is important.
- Relevance and salience are important.
- Expertise is important.
- Heuristics are important.

We need to understand all these better, from logical, computational, and philosophical perspectives.

Structures and domains

Some nontrivial aspects of mathematical reasoning are so fundamental as to be almost invisible.

A homomorphism between groups is a function $f : G \to H$ such that for every $g_1, g_2 \in G$, $f(g_1 \cdot g_2) = f(g_1) \cdot f(g_2)$.

What is the meaning of each \cdot ?

If X is a square matrix over \mathbb{C} , define

$$e^X = \sum_{i=0}^{\infty} \frac{1}{i!} X^i.$$

What sort of object is *i*? What does this expression mean?

Sometimes "is" means "is isomorphic to."

Both $\mathbb{Z}/7\mathbb{Z}$ and $\{\alpha\in\mathbb{C}\mid\alpha^7=1\}$ are the cyclic group on 7 elements.

As structures given by a construction, $\mathbb{Q}[\sqrt{2}][\sqrt{3}]$ and $\mathbb{Q}[\sqrt{3}][\sqrt{2}]$ are distinct but isomorphic.

As subsets of the complex numbers, they are identical. (This is a reason to embed structures in common completions.)

Structures and domains

We often identify not only isomorphic structures, but elements of different structures.

When talking about polynomials, we often identify the real number 3 with the polynomial 3, and the integer 3 with the real number 3.

Sometimes we care about a smaller domain in isolation, and then view it as part of a larger domain.

 \mathbb{R}^n denotes *n*-dimensional Euclidean space, whose elements are *n*-tuples of numbers. We commonly conflate \mathbb{R}^1 and \mathbb{R} .

Relevance and context

In interactive theorem proving, there is a notion of a *context* or *local context*.

1 goal p : \mathbb{N} , hp : prime p, a : zmodp p hp, ha : a \neq 0 \vdash a ^ (p - 1) = 1

This is distinguished from the *environment* or *global context*.

A lot of effort is involved in mediating between the two.

Relevance and context

Given an inference to verify, Isabelle's *sledgehammer* uses a relevance filter (which uses both heuristics and machine learning) to select \sim 200 facts from the library.

The problem is sent with these lemmas, and facts from the local context, to an external automated theorem prover.

If the external prover succeeds in verifying the inference, it reports on which fact from the library were used.

Isabelle uses this information (and often nothing else) to reconstruct a proof.

How do we make computers better at finding the relevant information?

- Use better heuristics.
- Let mathematicians provide annotations and advice.
- Use machine learning.

Modularity

Mathematics is modular:

- Mathematics has fields and subfields.
- Algebraic structures are defined hierarchically.
- Proofs are broken down into lemmas.

Well-defined interfaces support *encapsulation*.

- One can use a theorem without knowing the proof.
- One can combine algebraic structures as components.
- External developments are not sensitive to precise definitions.
- We can generalize theorems, definitions, notation, and proofs.
- We can interpret and reuse the mathematics of the past.

Modularity

Contemporary theorem provers provide lots of ways of supporting modularity:

- namespaces
- theories
- modules
- private or opaque annotations
- means for defining and extending algebraic structures
- generic notation and class inference

These support reuse of notation, facts, and patterns of inference.

Generality and specificity

Automation ranges from domain general to domain specific.

- propositional logic
- first-order logic
- equality
 - normalization and rewriting
 - symbolic calculation with operations that are associative, commutative, idempotent, ...
 - general equational reasoning
- linear arithmetic (integer / real)
- algebraic geometry, real algebraic geometry
- higher-order reasoning

Combination methods try to incorporate various domain-specific methods into a general search.

Outline

- motivating questions
- intuitions
- formal methods in mathematics
- from formal methods to epistemology

Problems and intuitions

Motivating questions:

- the problem multiple proofs
- the problem of conceptual possibility
- the nature of diagrammatic inference
- the problem of reliability
- the role of abstraction
- the use of computers in proofs

We need to understand:

- methods
- concepts
- representations
- cognitive effort

Insights from formal methods

We have more refined models of:

- mathematical language
- domains and structures
- relevance and context
- modularity
- methods of inference, both domain general and domain specific.

We have learned a lot about how to put formal methods to good use in mathematics, and what the challenges are.

This should help us develop a philosophical theory of mathematical understanding.

But we still have a long way to go, and it would help to have a better conceptual foundation.

I advocate a three-pronged approach.

First prong: articulate the big questions.

We're hacking through the wilderness.

We need to start mapping out the terrain, clarifying the issues.

Do this sparingly.

Second prong: focus on concrete, down-to-earth questions:

- How can we design formal languages to capture specific aspects of mathematical language?
- How can we automate various kinds of inference?
- What sorts of problems can we reduce to manageable search?
- How do diagrams support reasoning in specific domains?
- What types of activities promote student understanding?
- How can we model successful problem solving in specific domains?
- How did some particular historical development open up new possibilities for thought?

Third prong: be patient.

If we

- continue to make progress on specific questions and
- keep the general questions in mind,

a theory of mathematical understanding will eventually emerge.

What about the overarching questions: Why do we do mathematics the way we do?

View mathematics as a communal practice designed to meet fundamental constraints:

- scientific utility
- cognitive efficiency
- communicability
- reliability
- stability

The best justification for mathematics is that it serves its purposes well.

We need to better understand how and why.

The modes of analysis used in late twentieth-century analytic philosophy of mathematics have been thoroughly explored.

But there are new questions to ask, and new avenues to consider.

If we think of philosophy of mathematics in narrow terms, we *are* at the end.

But if we think of it more broadly, we have just started, and there is a lot to look forward to.

Concluding remarks

Philosophy of mathematics is dead.

Concluding remarks

Philosophy of mathematics is dead.

Long live philosophy of mathematics!

References

- Paolo Mancosu, ed., Philosophy of Mathematical Practice
- Yacin Hamami and Rebecca Morris have written a nice survey (ask them for a preprint).
- Association for the Philosophy of Mathematical Practice
 - The organization maintains a list of recent publications by members.
 - Follow links to find programs and abstracts from recent meetings.
- These conferences will be interesting, when they happen:
 - The 'End' of Philosophy of Mathematics
 - Understanding Mathematical Explanation
- Seminars I have taught have bibliographies:
 - Mathematics and Language (2015)
 - Mathematical Understanding and Cognition (2012)
- A recent paper of mine has a long bibliography.