Computers, mathematical proof, and the nature of the human mind

POMSIGMAA Keynote Address Philosophy of Mathematics

American Mathematical Society and Mathematical Association of America Joint Meetings

Jeff Buechner

Rutgers University-Newark and The Saul Kripke Center, CUNY, The Graduate Center

January 6th, 2017

1976 Appel and Haken prove the four-color theorem

- June 21, 1976 Wolfgang Haken and Kenneth Appel, with the aid of John Koch, completed their proof of the Four-Color Theorem (4CT). (Haken turned 48 years old on that day.)
- Their proof was published in 1977: "Every planar map is four colorable," Parts 1 and II, and Supplements I and II, Illinois Journal of Mathematics, XXI, 84, September 1977

1976 Appel and Haken prove the four-color theorem

- At one place in the proof of the 4CT, Appel and Haken need to find a finite list of reducible configurations with the property that every graph contains at least one configuration in the list. To do this, a lemma must be proved: that every configuration in an unavoidable set is reducible. A computer is needed to prove that all of the configurations are reducible. For instance, to show that one kind of configuration in the set is reducible requires 1,000,000 steps.

1976 Appel and Haken prove the four-color theorem

- Although computers had already been used to prove theorems in mathematics before the Appel-Haken proof, the importance of the 4CT brought the use of a computer in proving it to the forefront of attention of mathematicians, as well as a lay public. Moreover, some mathematicians did not believe that the theorem had been proved, since the computer proof of part of the 4CT is too long for a human being to survey.

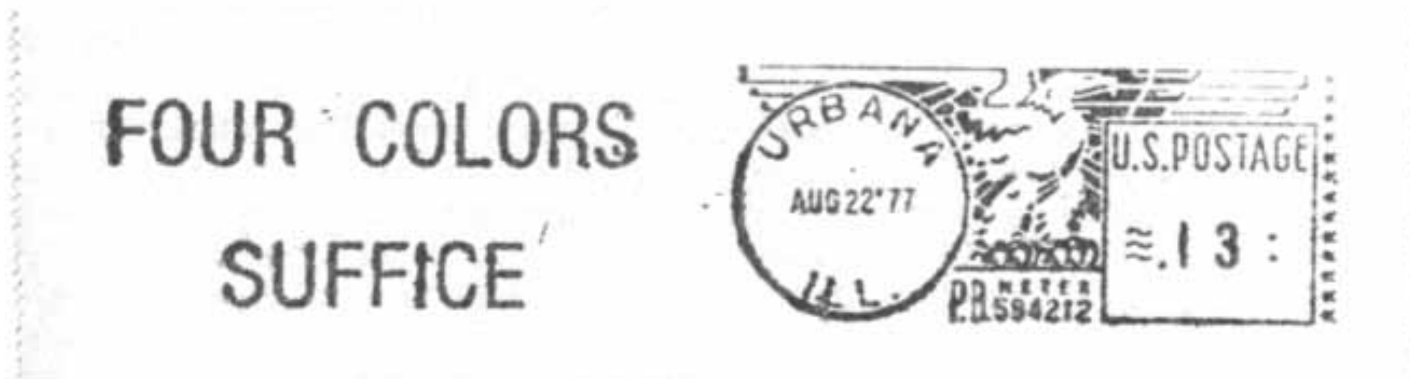
1976 Appel and Haken prove the four-color theorem: discord in the ranks

- “In the analysis of each case the computer only announced whether or not the procedure terminated successfully. The entire output from the machine was a sequence of yeses. This must be distinguished from a program which produces a quantity as output which can subsequently be verified by humans as being the correct answer… The real thrill of mathematics is to show that as a feat of pure reasoning it can be understood why four colors suffice. Admitting the computer shenanigans of Appel and Haken to the realm of mathematics would only leave us intellectually unfulfilled.” Daniel Cohen “The superfluous paradigm,” 1991

1976 Appel and Haken prove the four-color theorem: discord in the ranks

- “Nowhere in their long and often irrelevant account do they provide the evidence that would enable the reader to check what they say. It may or may not be ‘possible’ to prove the color theorem the way they claim. What is more certain is that they did not do so… not only is no proof to be found in what they published, but there is not anything that even begins to look like a proof. It is the most ridiculous case of ‘The King’s New Clothes’ that has ever disgraced the history of mathematics.” George Spencer-Brown, appendix to German edition of his Laws of Form

Their proof is implicitly recognized as valid by the United States Postal Authority



Their proof is implicitly recognized as valid by the United States Postal Authority

- The marking ‘FOUR COLORS SUFFICE’ was made by a University of Illinois-Urbana postal meter, not at a United States Post Office. But recognition is implicit, since anything which is illegal cannot be marked on a stamp by a university postal meter. So far, however, the United States Postal Authority does not take mistaken mathematical proofs to be illegal.

A shorter and improved proof

- Neil Robertson, Daniel Sanders, Paul Seymour, and Robin Thomas provide a new proof of the four-color theorem in 1994. An outline of their proof is published in Proceedings of the International Congress of mathematicians in 1995.

1979 Tymoczko on the Four-Color Theorem

- The first paper in the philosophy of mathematics on the philosophical importance of the four-color theorem appeared in February, 1979.
- Thomas Tymoczko “The Four-Color Problem and Its Philosophical Significance,” Journal of Philosophy Vol. 76, No. 21, pp. 57-83

1979 Tymoczko on the Four-Color Theorem

- “What reason is there for saying that the 4CT is not really a theorem or that mathematicians have not really produced a proof of it? Just this: no mathematician has seen a proof of the 4CT, nor has any seen a proof that it has a proof. Moreover, it is very unlikely that any mathematician will ever see a proof of the 4CT.” Tymoczko, op. cit. p. 58
- Elementary inference: Appel and Haken are mathematicians. So neither has ever seen a proof of the 4CT.

Tymoczko on what the four-color theorem shows

- “If we accept the 4CT as a theorem we are committed to changing the sense of the underlying concept of ‘proof.’”
- “The use of computers in mathematics, as in the 4CT, introduces empirical experiments into mathematics, and raises again for philosophy the problems of distinguishing mathematics from the natural sciences.” Tymoczko op. cit. p. 58

Tymoczko on the four-color theorem

- “The answer as to whether the 4CT has been proved turns on an account of the role of computers in mathematics.” op. cit. p. 58
- “The 4CT is substantial piece of pure mathematics which can be known by mathematicians only a posteriori. Our knowledge must be qualified by the uncertainty of our instruments, computer and program...The demonstration of the 4CY includes not only symbol manipulation, but the manipulation of sophisticated experimental equipment as well: the four-color problem is not a formal question.” Tymoczko, op. cit. pp. 77-78

Tymoczko on the four-color theorem

- “The idea that a proposition of pure mathematics can be established by appealing to empirical evidence is quite surprising. It entails that many commonly held beliefs about mathematics must be abandoned or modified. Consider:
- 1. All mathematical theorems are known a priori
- 2. Mathematics, as opposed to natural science, has no empirical content.
- 3. Mathematics, as opposed to natural science, relies only on proofs, whereas natural science makes use of experiments.
- 4. Mathematical theorems are certain to a degree that no theorem of natural science can match. Tymoczko, p. 63

Tymoczko on mathematical proof

- “Proofs are convincing. ..[In] Wittgenstein’s *Remarks on the Foundations of Mathematics*, this is all there is to proofs: they are convincing to mathematicians. This is to be taken as a brute fact, something for which no explanation can be given and none is necessary. Most philosophers are unhappy with this position and instead feel that there must be some deeper characterizations of mathematical proofs which explains, at least to some extent, why they are convincing.” Tymoczko, op. cit. p. 59

Tymoczko on mathematical proof

- Why are mathematical proofs convincing?
- "That proofs are surveyable and that they are formalizable are two such characterizations [J.B. of why mathematical proofs are convincing]." Tymoczko, op. cit. p. 59

Tymoczko on mathematical proof

- “We often say that a proof must be perspicuous, or capable of being checked by hand. It is an exhibition, a derivation of the conclusion, and it needs nothing outside of itself to be convincing.” Tymockzo, op. cit. p. 59

Tymoczko's circle

- Unfortunately, Tymoczko's definition of surveyability uses the idea that a mathematical proof needs nothing outside itself to be convincing. So one must already know what a mathematical proof is before one knows what surveyability consists in; but surveyability is one criterion of being a mathematical proof.
- "The mathematician *surveys* the proof in its entirety and thereby comes to know the conclusion." Tymoczko, op. cit., p. 59
- "The construction that we surveyed leaves no room for doubt." Tymoczko, op. cit. p. 60

Paul Teller on Tymoczko

- “Surveyability is needed, not because without it a proof is in any sense not a proof, but because without surveyability we seem not to be able to verify that a proof is correct. So surveyability is not part of what it is to be a proof in our accustomed sense.” Paul Teller “Computer Proof,” Journal of Philosophy, December 1980, pp. 797-803

Paul Teller on Tymoczko

- "…we may take advantage of new methods of surveying as long as these enable us to meet sensible demands on checking proofs, and a shift in the means of surveying actually used means only a shift in methods of checking proofs, not a shift in our conceptions of the things checked." Teller, op. cit., p. 798
- Not a shift in our conceptions of the things checked = not a shift in our concept of proof

The dispute between Teller and Tymoczko: the concept of mathematical proof

- Tymoczko: surveyability is an essential feature of the concept of a mathematical proof.
- Teller: surveyability is not an essential feature of the concept of a mathematical proof.
- Who is right? On what grounds are they right?

Quine’s problem for concepts

- W. V. Quine argued in his epochal paper “Two Dogmas of Empiricism,” that there is no hard and fast distinction between meaning-constituting beliefs and auxillary beliefs (beliefs that are not meaning-constituting).
- This means that it is impossible to draw a hard-and-fast line between essential (or necessary) features of a concept and non-essential (contingent) features of a concept.

Quine’s problem for concepts

- If A proposes that X is an essential feature of the concept of mathematical proof, and B proposes that it is an accidental feature of the concept of mathematical proof, there is no principled way of adjudicating between the two proposals.
- Adjudication should go by way of canons of rationality and canons of scientific inquiry, such as conservatism—upholding as many currently established beliefs as possible. How would that work for the concept of mathematical proof?

Quine’s problem for concepts

- There is no consensus view as to whether Quine is correct or not on this, but most philosophers take Quine to be correct.
- Teller’s claim that surveyability is not an essential feature of the concept of mathematical proof could be upheld if it satisfied more canons of rationality and of scientific inquiry than does Tymoczko’s claim.

Quine's problem for concepts

- However, the idea of surveyability surfaces in the context of using computers in mathematical proofs.
- There was not much data (i.e., features on which there is common agreement) concerning the use of computers in mathematical proof in the period 1976-1980.
- The disagreement between Teller and Tymoczko is a stalemate.

Detlefsen on Tymoczko

- “The need for the appeal to empirical evidence is brought about, in Tymoczko’s view, by the fact that the calculation performed by an IBM 370-160A in order to determine the reducibility of certain configurations is too long to be ‘checked’ or ‘surveyed’ by human mathematicians. Because of this, Tymoczko reasons, whatever evidence we have for the reliability of the IBM 370-160A in determining reducibility of configurations cannot take the form of a ‘surveyable’ proof of its reliability. And so, it is concluded, the evidence must be empirical in character.” Michael Detlefsen and Mark Luker“The Four Color Theorem and mathematical Proof,” Journal of Philosophy, 1980, pp. 803-820

Detlefsen on Tymoczko

- Detlefsen provides several examples of mathematical proofs which are surveyable and in which computations are made. He argues that such computations necessarily utilize empirical premises (such as: the computing agent correctly executes the program required to make the computation).
- If his argument is sound, Detlefsen has shown that unsurveyability is not necessary for the presence of an empirical element in mathematical proofs. This refutes a major claim in Tymoczko's paper.

Detlefsen on Tymoczko

- “This creates a dilemma for Tymoczko. For either one rejects his reasoning, in which case he is left without an argument for the empirical character of the proof of the 4CT or one accepts his reasoning, but is then forced to view the presence of calculation or computation in a proof as injecting an empirical element into that proof. The consequence of such a view is that empirical proofs are more widespread than Tymoczko himself indicates.” Michael Detlefsen, op. cit., p. 809

Detlefsen on Tymoczko

- “This creates a dilemma for Tymoczko. For either one rejects his reasoning, in which case he is left without an argument for the empirical character of the proof of the 4CT or one accepts his reasoning, but is then forced to view the presence of calculation or computation in a proof as injecting an empirical element into that proof. The consequence of such a view is that empirical proofs are more widespread than Tymoczko himself indicates.” Michael Detlefsen, op. cit., p. 809

Background to Burge: Descartes on mathematical proofs

- [In long deductions] “the last link is connected with the first, even though we do not take in by means of one and the same act of vision all the intermediate links on which that connection depends, but only remember that we have taken them successively under review…” Descartes, Rules for the Direction of Mind
- For Descartes, “if that knowledge is deduced from evident mathematical premises, it is certain and demonstrative.” Tyler Burge, Content Preservation, Philosophical Issues, 1995, p. 271

Background to Burge: Chisholm on mathematical proofs

- "What if S derives a proposition from a set of axioms, not by means of one or two simple steps, but as a result of a complex proof, involving a series in interrelated steps? If the proof is formally valid, then shouldn't we say that S knows the proposition *a priori*? I think that the answer is no." Roderick Chisholm, Theory of Knowledge, 2nd edition

Background to Burge: Chisholm on mathematical proofs

- “[I]f, in the course of a demonstration, we must rely upon memory at various stages, thus using as premises contingent propositions about what we happen to remember, then, although we might be said to have ‘demonstrative knowledge’ of our conclusion, in a somewhat broad sense of the expression ‘demonstrative knowledge,’ we cannot be said to have a *a priori* demonstration of the conclusion.” Roderick Chisholm, op.cit.,

Why is Chisholm’s point important?

- a posteriori knowledge: knowledge which is arrived at on the basis of sense experiences or perceptual beliefs.
- a priori knowledge: knowledge which is arrived at on the basis of intellectual processes which do not involve reference to or reliance upon sense experiences.
- a posteriori justification: justification which relies upon sense experiences.
- a priori justification: justification which employs intellectual processes which do not involve reference to or reliance upon sense experiences.

Why is Chisholm’s point important?

- If Chisholm is right that long mathematical proofs require a premise about what we happen to remember—and thus are not either known or justified a priori—then it easily follows that those proofs are known or justified empirically. They rely upon or refer to sense experiences.
- Tymoczko is right about the epistemological status of the 4CT if we accept Chisholm’s point. But he is wrong that only unsurveyable mathematical proofs require (in whole or in part) empirical justification.

Background to Burge: Fallis on dispensing with empirical evidence

- "… there is a sense in which the proof of the four-color theorem is an a priori justification. It need not appeal to any particular empirical data and in principle need not appeal to empirical data at all. For instance, the relevant computation could be performed by a device other than a digital computer and in principle could be performed in the mathematician's mind." Don Fallis, Mathematical Proof and the Reliability of DNA Evidence, American Mathematical Monthly, June-July, 1996, p. 496

Background to Burge: Fallis on dispensing with empirical evidence

- Fallis thinks that, unless it is necessary that a physical machine of some kind perform some computation, then the computation is a priori, because it is possible that a human mind could perform the computation.
- It is possible that a human mind could complete an infinite computational process (in, say, a Malament-Hogarth universe). Should we then say that such computations are a priori knowable?

Burge on the use of computers in mathematical proofs

- Tyler Burge, in 1998, in his important paper “Computer Proof, A Priori Knowledge, and Other Minds,” The Sixth Philosophical Perspectives Lecture (pp. 1-38), sidesteps the need to understand what a mathematical proof is in asking whether the use of computers in mathematical proofs adds an empirical element to such proofs. Burge will argue that no empirical element need be added when computers are used in mathematical proofs, such as the proof of the 4CT. The 4CT can be known a priori (to be true).

Burge on mathematical proofs

- Burge uses the terms ‘epistemic entitlement,’ ‘epistemic warrant,’ and ‘epistemic justification’ (sometimes without the qualifier ‘epistemic’).
- Unless you are a philosopher working in epistemology, it is best to treat all of them as meaning the same—namely, justification.

Burge on mathematical proofs

- “I conclude, given our assumptions, one can be in a position, from the third person point of view, to be a priori warranted in believing, in fact, knowing, on defeasible inductive grounds, that the [4CT] has been proved. One can know this even if one cannot replicate the proof.” Tyler Burge Computer Proof, A Priori Knowledge, and Other Minds, p. 23

Burge on mathematical proofs

- "The entitlement for relying on the source does specify the source [the computer]. But it need not specify the object's empirically discernible characteristics, or the empirical background conditions that enable the recipient to access and rely on the source. It can specify the source in the non-empirical way that the a priori justification does." Burge, op. cit., p. 29
- Work on proof assistants (e.g., by Harvey Friedman and by Jeremy Avigad) provides an entitlement for relying on the source (the IBM 370-160A used in the Appel-Haken proof of the 4CT).

Burge on mathematical proofs

- “Perceptual experience of the words or of the body of the source need play no role in justifying one’s understanding of, or intellectual uses of, the content of the words or the presentations of the source.” Notice that this point would, if correct, refute Chisholm.
- “Perception is only the mode of access, an enabling condition which makes no contribution to the epistemic force of the warrant.”
- Perception is merely a condition that enables one to make use of a resource for reason and understanding.” Burge, op. cit., p. 30

What has Burge shown?

- If his arguments are sound, Burge has shown that the unsurveyability of mathematical proofs is not sufficient for the existence of an empirical element in such proofs.

What have Detlefsen and Burge shown?

- Recall that Detlefsen has shown that the unsurveyability of mathematical proofs is not necessary for the existence of an empirical element in such proofs. Burge has shown that it is not sufficient.
- The results of Burge and Detlefsen, if both are correct, show there is no conceptual connection between the unsurveyability of mathematical proofs and the existence of an empirical element in such proofs. However, both cannot be correct.

What have Detlefsen and Burge shown?

- If there is no conceptual connection between the unsurveyability of mathematical proofs and the existence of an empirical element in such proofs, it easily follows that unsurveyability has nothing to do with the existence of an empirical element in mathematical proofs.
- Examining the arguments of both Detlefsen and Burge, this is not shocking, nor even surprising.

What have Detlefsen and Burge shown?

- Detlefsen argued that whether a mathematical proof is surveyable or unsurveyable, there might be empirical elements in it.
- Burge argued that whether a mathematical proof is surveyable or unsurveyable, there need not be empirical elements in it.
- Of course, neither rule out that there might be, nor that there might not be, empirical elements in a mathematical proof. But whether there are or are not is not a matter of what the concept of a mathematical proof consists in. It is, rather, an entirely contingent matter.

Four Problems for Burge

- There are three problems for Burge's account of how we can have a priori knowledge of the output of a computer.
- The first is that his account makes it too easy to have gettiered knowledge. Gettier counterexamples are cases in which a subject S has a true, justified belief that p, but in which S does not know that p.

Four Problems for Burge

- Here is how, following Burge’s account of epistemic justification in the context of mathematical truths, one can have a true, justified belief in the 4CT, but not know the 4CT.
- Suppose that the 4CT is true, but that the computer program for resolving the cases is fallacious. On Burge’s account, a subject S will be justified in believing the 4CT to be true. Since it is true (by assumption), S has a true, justified belief in the 4CT. But S does not know the 4CT.

Four Problems for Burge

- We should take Gettier counterexamples very seriously. (David Lewis has remarked that there are only two results that all philosophers take to be definitive: Gödel and Gettier.)
- If an account of epistemic justification makes it too easy for Gettier counterexamples (and not just possible for them to arise) to arise, that is a reason to reject such an account.

Four Problems for Burge

- The second problem for Burge concerns his claim that the mode of access to some epistemically warranted set of propositions is not necessary for being epistemically justified in believing those propositions: “[t]he entitlement for relying on the source … need not specify the empirical background conditions that enable the recipient to access and rely on the source.” Burge, op. cit., p. 29

Four Problems for Burge

- That is why we can discount the role of memory in determining whether we are epistemically justified in believing a theorem of mathematics on the basis of the proof of that theorem. Memory is a mere mode of access to the proof.
- For Burge, the same is true of computer proofs—we can discount the mode of access to the theorem and its proof (the computer program) which is the computer.
- If memory is faulty, that does not show that the proof is faulty. Indeed, a faulty memory has nothing to do with a proof—which is an abstract object. Can we say the same of a computer? Burge thinks we can.

Four Problems for Burge

- I say we cannot. Here is why. Unlike memory, a computer is not an organic intrinsic part of a human being. It is the unique mode of access to the computer program and the computations of that program—all of which are abstract objects.
- But it is more. It is the means by which the abstract objects are physically realized. Memory, on the other hand, need not be the means by which a proof is physically realized. Rather, a proof can be physically realized on a piece of paper using ink to make inscriptions.

Four Problems for Burge

- Imagine a culture in which there is no paper, no writing instruments, and no concepts of writing (on paper, using writing instruments). However, there is the concept of a proof. All proofs are in human memory.
- In such a case, we should say that a problem with human memory would create a problem in epistemic entitlement to the proof. Why? Because human memory is the only means by which the proof is physically realized, as well as the mode of access to the proof.
- Without memory, we do not have a proof, since we do not have any mode of access to the abstract object which is the proof. (Compare with a proof which is so difficult that no mathematical concepts available to the human mind are adequate for representing the proof. In such a case, even though the proof has an abstract existence, we should say that we cannot be epistemically entitled to it since we have no means by which to access it.)

Four Problems for Burge

- The third problem for Burge concerns four assumptions that Burge makes in his argument. One assumption is that "individual's knowledge of pure mathematics, resting on specifically mathematical understanding or reasoning, is ordinarily a priori." (Burge, op. cit., p. 4)
- This contradicts Detlefsen's claim—which depends on Tymoczko's definition of mathematical proof—that empirical premises are used in mathematical proofs that are surveyable (as well as those which are unsurveyable).
- We defer our exposition of the fourth problem for Burge.

The dialectics of how things stand

- Tymoczko and Teller: stalemate
- Tymoczko and Detlefsen: If Detlefsen is correct, empirical premises occur in mathematical proofs that are both surveyable and unsurveyable. This puts pressure on getting clear on what we mean by a mathematical proof.
- Tymoczko and Burge: If Burge is correct, then the use of computers in mathematical proofs does not introduce an empirical element into those proofs (nor does the use of computations in mathematical proofs). Tymoczko and Detlefsen are both refuted.

The dialectics of how things stand

- But we have presented reasons for thinking that Burge’s argument fails.
- It is clear that much philosophical work still needs to be done in explicating the concept of a mathematical proof.
- But no matter what that explication eventually consists in, it must be compatible with our views about the nature of computers and the nature of the human mind. That this (perhaps startling view) is so will be argued in the remainder of this talk.

A line of thought not taken

- We will now discuss a line of thought that is broached by Teller, Detlefsen, Davis, and Tymoczko, but not taken up by any of them.
- In his paper, Teller writes: “The alleged nonsurveyability also underlies Tymoczko’s second conclusion: the computer proof of the combinatorial lemma is subject to error—computers can make mistakes. We cannot guard against this possibility of mechanical failure or error in programming in the traditional way because we cannot survey the proof.” Teller, op. cit., p. 798

A line of thought not taken

- “What if the programming was erroneous? What if the initial data were fake? What if there was a machine malfunction?”
- “These considerations lead us to a position—which is rarely discussed in works on the philosophy of mathematics and which is very unpopular—that a mathematical proof has much in common with a physical experiment.”

P. Davis, “Formac Meets Pappus,” American mathematical Monthly, 1969, pp. 903-904.

John Horton Conway on computers

- A well-known mathematician, John H. Conway, has been quoted as saying: “I don’t like them [computers], because you sort of don’t feel you understand what’s going on.” New York Times, April 6, 2004 Kenneth Chang “In math, computers don’t lie. Or do they?” an article on the use of computers in mathematical proofs

Wittgenstein on machine computation

- “If we know the machine, everything else … seem[s] to be already completely determined. We talk as if these parts could only move in this way, as if they could not do anything else. Is this how it is? Do we forget the possibility of their bending, breaking off, melting, and so on? Yes, in many cases we don’t think of that at all. We use a machine, or a picture of a machine, as a symbol of a particular mode of operation. For instance, we give someone such a picture, and assume that he will derive the successive movements of the parts from it.” Ludwig Wittgenstein Philosophical Investigations, § 193

Kripke on Wittgenstein

- “Wittgenstein himself draws the distinction between the machine as an abstract program (‘der Machine als Symbol,’ PI 193) and the actual physical machine, which is subject to breakdown (‘Do we forget the possibility of their bending, breaking off, melting, and so on?’ PI 193)” Saul Kripke, Wittgenstein on Rules and Private Language, p. 35, fn. 24

Naive computer view of the mind

- "A machine can follow this rule; whence does a human being gain a freedom of choice in this matter which a machine does not possess?"

Sir Michael Dummett "Wittgenstein's Philosophy of Mathematics," Philosophical Review Vol. 68 (1959), pp. 324-348, at p. 351

The basic idea

- Since physical computing machines can break down in various ways, how do we really know what function F a given PCM computes?
- One might think that is not a serious problem. If F is the square function, and the PCM computes F(2) = 4, the PCM is operating normally. IF the PCM computes F(2) = 8, then it has suffered a breakdown.

The basic idea

- The basic idea is not that of the under-determination of theory by data. For instance, both the square function and the doubling function output ‘4’ when their input is ‘2.’ Indeed, there are many infinitely many functions whose initial segment consists of the integer ‘4.’ As more and more values of F are computed (say n), functions whose initial segment consist of the sequence of n-1 values will no longer share n values.
- But this is not a matter of underdetermination of theory by data. It is something quite different.

The basic idea

- That view is too naïve. There are many other functions (say, G) that PCM might be computing. Perhaps the output ‘4’ is when PCM suffers a breakdown in computing G. Perhaps the output ‘8’ is when PCM operates normally in computing G.
- Unless it is KNOWN that the PCM computes, say, F, it cannot be ruled out that, based on its behavior, it is computing, say, G.

The basic idea

- In short, we have to idealize the physical behavior of the PCM as computing, say, F, if we are to understand just what a PCM computes and what it does not compute.
- But to idealize the physical behavior of the computing machine as computing, say, F, we must already know that it does compute F.
- Where did we acquire this knowledge? Certainly, not from the physical behavior of the PCM (which physical behavior includes what PCM outputs), since we have idealized that physical behavior on the assumption that PCM computes F.

The basic idea

- We can’t identify the function a PCM computes by observing that it is operating normally, or is suffering a breakdown.
- We cannot do that because we cannot know whether PCM is operating normally or suffering a breakdown unless we already know what function PCM is computing.
- By idealizing the physical behavior of a PCM, we implicitly stipulate whether conditions are normal or breakdown.

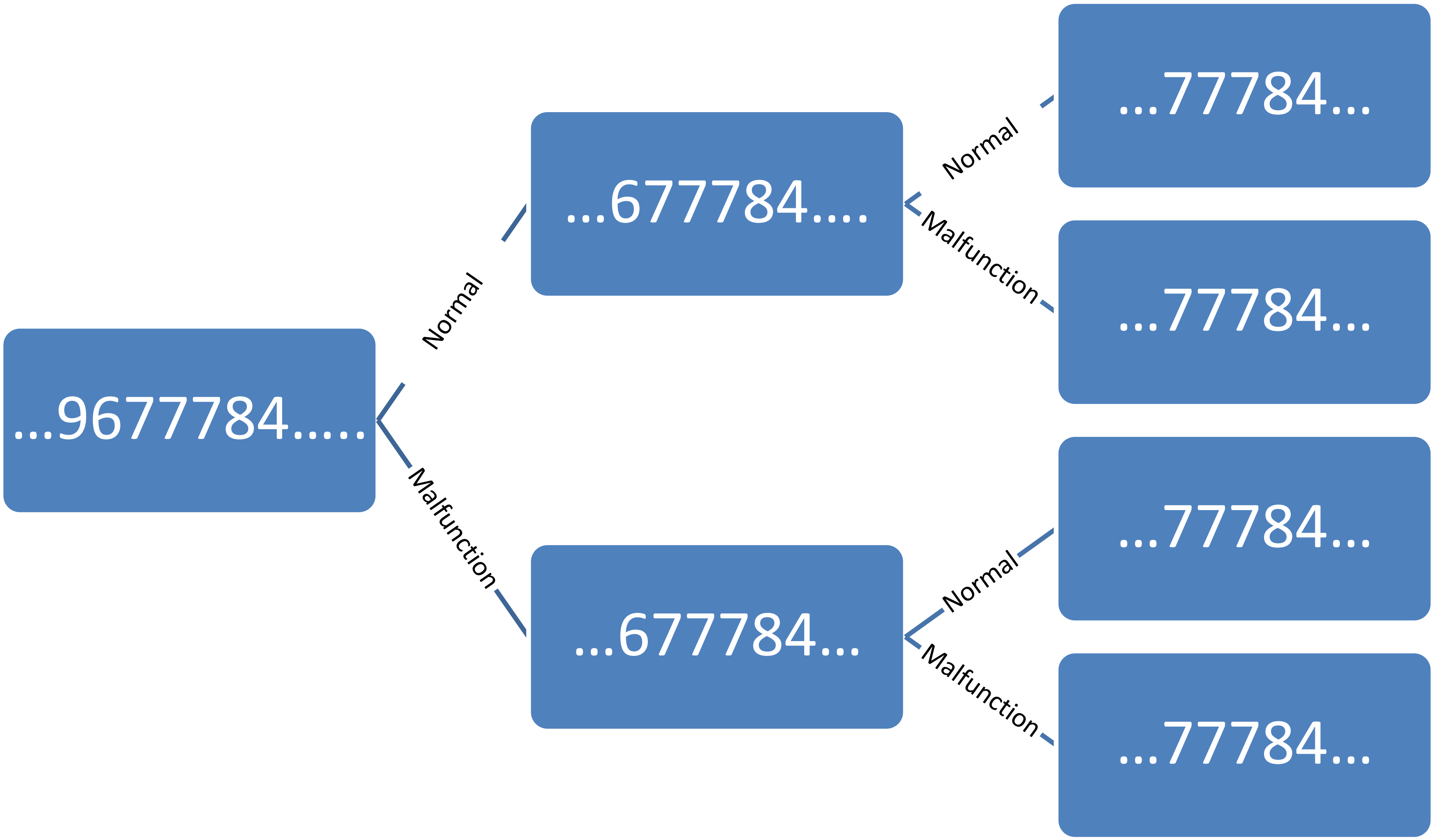
The basic idea

- We cannot appeal to the intentions of the designers of the PCM to determine what function the PCM computes unless we already know that they intend the PCM to compute, say, F.
- If we make such an appeal, then we can say whether the PCM is operating normally, or is in breakdown mode, only if we already know that the code for the PCM is the code for correctly computing, say, F.

How many distinct designers of the IBM 370-160A might there be?

- Construct a Boolean tree, where for any node, the top-most branch leaving it represents normal conditions and the bottom-most branches leaving it represent malfunction conditions.
- Feed the successive nodes in the tree successive digits in the sequence of output digits of some F.

How many distinct designers of the IBM 370-160A might there be?



A disturbing conclusion

- At any given stage in the computation of F, it is must be assumed that the computing machine is computing F, and not some other function, such as G. Even after the computation ends, and one can see (by observation) that the computing machine outputs the digits in the computation of F(n), it must still be assumed that F has been computed, and not some other function, such as G (because for each digit in F(n), it could have been computed by G, ….)

Kripke's argument is not an underdetermination argument

- An underdetermination argument: given evidence e, there are n hypotheses compatible with e. Where e is the output m = F(n), there are infinitely many recursive functions which agree with that output for that domain value. As other values of F are computed, the number of hypotheses compatible with e decreases.

Kripke's argument is not an underdetermination argument

- As more and more values of either F or of the digits of F(n) are computed, more and more possible functions arise that the computing machine might be computing. This is how the phenomenon of machine malfunction is importantly different from the phenomenon of the underdetermination of theory by evidence.

Kripke's argument is not an underdetermination argument

- In underdetermination arguments, we can meaningfully speak of how likely it is that some function F has been computed, since we have data concerning all of the functions which the computer might have computed.
- In Kripke's argument against functionalism, we cannot meaningfully speak of likelihoods.

Why likelihoods are ruled out

- Which function F a given computer computes might be any one of 2^n different functions.
- But unless one idealizes as to which function F a given computer computes, it won't be any of those 2^n functions.

Why likelihoods are ruled out

- It would appear to make sense to ask: “How likely is it that F is computed?”
- Given that 2^n functions could be computed, we answer: “It is $1/2^n$ likely that F is computed.”
- But this makes sense only if there is a fact-of-the-matter as to which F is computed.

Why likelihoods are ruled out

- However, in the absence of making an idealization as to which F a computer computes, there is no fact-of-the-matter as to which F it computes.
- And once the idealization is made, the fact-of-the-matter is that only one function F is computed. So it is certain that F is computed under the idealization.

What could we conclude from an underdetermination argument?

- Let’s briefly look at what we would say about a given physical computer physically computing some function F where we employ an underdetermination argument.
- This is useful to do, since one might confuse Kripke’s argument against functionalism with an underdetrmination argument.

What could we conclude from an underdetermination argument?

- No mathematician is ever justified that a computing machine is computing F because the probability that the machine is computing F is less than or equal to .5
- Indeed, for all computations of any function, the probability the machine is computing that function is less than or equal to .5
- We have no more reason to believe the computing machine is computing F than we have reason to believe that a fair flip of a fair coin will come up heads.

What could we conclude from an underdetermination argument?

- Since there are 2^n different functions computed that are compatible with a given output of a computing machine, the probability that the computing machine computes F is $1/2^n$.
- The more digits in F(n) that are computed, the more likely it is that F(n) has been computed by the computing machine.

What could we conclude from an underdetermination argument?

- Suppose that a computing machine outputs m, which happens to be the value of the recursive function F(n). It also happens to be the value of G(k), …
- Suppose that there are infinitely many recursive functions that output m for a given input value n. (There are infinitely many recursive functions that agree with F(n) for domain value n.)
- It would then follow that the probability that the computing machine computes F is $1/\infty = 0$ (as a limit, but of what function) or indeterminate

What could we conclude from an underdetermination argument?

- Underdetermination of theory by evidence is not what is the case where it is the computing machine which may or may not be exhibiting a breakdown. What function it is computing determines whether it is in breakdown mode or is operating normally. But one cannot know what function it is computing without knowing whether it is operating normally or is in breakdown mode.

Why computers are unreliable

- We have belabored the difference between Kripke’s argument against functionalism and underdetermination claims so that one can see fairly easily that underdetermination claims do not show computers are unreliable.
- However, Kripke’s argument against functionalism does show computers are unreliable, since in the absence of making an idealization as to which F a computer computes, there is <u>no</u> fact-of-the-matter as to which F it computes.

Why computers are unreliable

- That there is no fact-of-the-matter as to which F a given computer computes and that someone who uses the computer must stipulate which F it does compute shows they are unreliable.
- Reliability of a physical device is established by showing the extent to which the outputs of the device correspond to what we take the device to be registering, computing, measuring, etc. But if there is no fact-of-the-matter at to what the device registers, computes, measures, etc., then it cannot, by definition, be reliable.

Being right and saying a computer is right in what it computes

- In idealizing a computer as computing F, one is stipulating that the computer computes F. In the absence of such an idealization—or stipulation—there is no fact-of-the-matter as to what the computer computes—indeed, as to what it does.
- The distinction being the computer being right in what it computes and our saying it is right in what it computes vanishes.

Being right and saying a computer is right in what it computes

- If Wanda idealizes a given computing machine as computing F, then that is what it computes—viz., F.
- If Greg idealizes the very same computing machine as computing G, then that is what is computes—viz., G.
- There is no fact-of-the-matter as to which idealization is correct. So we cannot speak in these cases of 'correctness.'

Being right and saying a computer is right in what it computes

- Where we cannot speak of a fact-of-the-matter about which one is either correct or not correct, we have relativism.
- Truth-relativism is the doctrine that truth is relative to a speaker. It is an insidious doctrine that philosophers have done their best to refute.
- Computation-relativism is the doctrine that which computation a given computer makes is relative to the idealization a given person makes. It is a consequence of Kripke’s argument against functionalism.

Wittgenstein on being right and saying one is right

- “And now it seems quite indifferent whether I have recognized the sensation *right* or not. Let us suppose I regularly identify it as wrong, it does not matter in the least. And that also shows that the hypothesis that I make a mistake is mere show.” Ludwig Wittgenstein, Philosophical Investigations, paragraph 270.

Putnam on being right and saying one is right

- “the relativist cannot … make any sense of the distinction between *being right* and *thinking he is right*; and that means there is … no difference between *asserting* or *thinking*, on the one hand, and *making noises* … on the other. … To hold such a view is to commit a sort of mental suicide.” Hilary Putnam, Reason, Truth, and History, p. 122

Truth relativism and computation relativism

- Computation relativism appears to be such an absurd view (like truth relativism), that one naturally takes it to be a reductio of Kripke’s argument against functionalism.
- However, although there are compelling arguments which refute truth—relativism, there are no compelling arguments (thus far) which refute Kripke’s argument against functionalism.

Fourth Problem for Burge

- "It is a delicate and unresolved matter how to distinguish the cases in which warrant for continuing reliance on a source Q requires an empirical induction, or even an empirical entitlement, from the cases in which empirical recognition can be submerged into knowing how to access a rational resource." Burge, op. cit., p. 29

Fourth Problem for Burge

- “[Cases in which empirical recognition can be submerged] require that the perceivable properties of a computer or person that one uses as a rational resource be relatively simple. I think that they must be incorporated into a nearly automatic routine. It is important that the recipient need not engage in context-dependent empirical (or non-empirical) tracking exercises, or complex theorizing, to reidentify the resource … through its possibly changing physical characteristics.” Burge op. cit., p. 29

Fourth Problem for Burge

- In order to track the state of the system making the computations, the recipient will need to idealize the behavior of that system. Why? Because in the absence of an idealization, the recipient cannot say what the system is computing: whether it is computing the function the recipient takes it to be computing, or whether it is computing another function

Fourth Problem for Burge

- Without making the idealization, the recipient cannot know whether the machine is computing the function she takes it to be computing, under normal conditions of operation, or computing another function she does not take it to be computing, under abnormal conditions of operation.

Fourth Problem for Burge

- Without making such an idealization, the recipient cannot know whether the machine is operating under normal conditions, or operating under abnormal conditions.
- If the machine is idealized as operating under normal conditions, and it outputs what the recipient thinks it should output, then the machine is computing the function the recipient takes it to be computing.

Fourth Problem for Burge

- Making such an idealization is a necessary part of understanding what function the machine is computing. Notice that even if the recipient has established the mathematical powers of the machine by a priori reasoning, that does not establish her epistemic entitlement that she is warranted in believing the machine will correctly compute the functions she takes it to be computing.

Fourth Problem for Burge

- But making such an idealization is to engage in “complex theorizing to re-identify the resource through its possibly changing physical characteristics.”
- We need to refer to empirical constancy not just for access, but also refer to it in our warrant. (Burge: “We rely on empirical constancy for access, without having to refer to it in our warrant,” [J.B. unless the recipient engages in complex theorizing.]

Where do we go from here?

- There is much work to be done on developing a concept of mathematical proof and on proof assistants. But no matter what the development of these areas looks like in the future, unless we come to terms with the philosophical questions concerning the nature of the human mind, we will not be in a position to say whether a mathematical proof that uses computers (in the way the 4CT does) is a genuine mathematical proof.

Where do we go from here?

- Both the human brain and a physical computer are physical objects, subject to breakdown and malfunction.
- Modeling the human mind as a computational device works at the abstract level, but computational devices must be physically realized, and it is in their physical realization that problems arise.
- Could we re-think how a computer works by analogy with a non-computational model of the human mind? Would that get around the problems that arise with physical realizations?

Where do we go from here?

- Speculation: we will not have an adequate concept of machine computations until we have an adequate set of concepts on the nature of the human mind.
- Whether these concepts must respect the mathematical work on computation is an open question. It might be that, e.g., a new concept of computational complexity will be needed.
- This seems strange; indeed, it IS strange. But the arguments I have presented here today show that, although strange, it is (perhaps) necessary.

The End

- Thanks to Bonnie Gold for much helpful editorial advice and discussion.