20 *Epistemic Warrant: Humans and Computers*[*]

I begin with some remarks about the epistemology of human cognition. Then I compare human cognition, taken by itself, with human cognition enhanced by relying on computers. Finally, I discuss some uses of computers in pure mathematics and in empirical science.

Knowing a proposition requires that the proposition be true, that one believe it, and that one's belief be epistemically warranted. These three conditions are not jointly sufficient for knowledge. But they are necessary. I will focus on the third condition—that one's belief be epistemically warranted.

Being epistemically warranted in having a belief is having the belief in a way that is good for having *true* beliefs, given limitations on one's information and cognitive capacities. One can be warranted but mistaken. But if one is warranted, one's belief is held through a natural competence that is epistemically good—conducive to the belief's being true and, usually, to the belief's constituting knowledge. To constitute warrant, this "good" route to truth must meet a certain minimum standard. It must be *reliable* in yielding true beliefs, in normal circumstances.

These remarks do not constitute a definition, analysis, or reduction of warrant. They are intended as orientation. They hide complexities. There are hard questions about limitations on information and capacities, about reliability, and about the normal circumstances in which reliability is required. I will not discuss the complexities.[1] I assume that we have an intuitive grip on the notion of epistemic warrant. I want to develop certain aspects of that notion.

Being warranted in having a belief is in the same ballpark as being justified. I count warrant the genus and justification a primary subspecies. I call the other main subspecies '*entitlement*'. What differentiates justification and entitlement? Being *justified* is having a reason that figures in an appropriate way in arriving at or in sustaining the relevant psychological state. (For *epistemic* justification, the primary relevant psychological state is belief.) *Having* a reason requires that the

* This essay was written for an interdisciplinary conference on applications of computers in empirical science, in Paris 2011.

[1] See my, 'Perceptual Entitlement', *Philosophy and Phenomenological Research* 67 (2003), 503–548.

reason is in one's psychology, or could easily be brought into one's psychology by simple, easy, obvious inferential transitions from what is already there.[2] So having a reason requires having the capacity to think the reason; and it requires being able to connect the reason with what it is a reason for. One can have a reason without the reason's being operative. A reason that one has for an attitude is *operative* if and only if the reason figures in a cognitively relevant causal way in forming or sustaining the attitude. Thus being epistemically warranted in having a belief, in the sense of being justified, requires having in one's psychology a reason that is operative.

I assume that reasons are propositional contents together with modes (like belief or intention). Reasons have the same structures as sentences, even though they need not be linguistic or even symbolic. Some think of perceptions or pictorial images, which I assume are not propositional, as reasons. I do not. Perceptions and images can figure in supporting an attitude, but they are not reasons for it. Reasons are explanatory as well as justificatory. Reasons are answers to potential 'why' questions. In effect, they provide a kind of explanation of the credibility of what they are reasons for.[3] Explanations and answers to questions are propositional. Perceptions and pictures, except as *glossed* by propositions, cannot explain—cannot answer a 'why' question or complete a 'because' clause. They are not themselves reasons. The point is a conceptual/grammatical one. Perceptions and pictures can certainly figure in entitlement and thus in support of belief.

Being *entitled* to a belief is being warranted in holding it, without depending for being warranted on having an operative reason for it. Entitlement is warrant without reason.

[2] This formulation allows modular justifications—justifications that occur in an individual's psychology, but that cannot, even in principle, be brought to consciousness by the individual. On this formulation, being justified hinges not on availability to consciousness, but on having a rational structure in one's psychology that functions to support and explain the relevant attitude. We might label being justified in this sense '$\textit{justified}_{\mathrm{p}}$', where the subscript indicates that an individual is justified in holding a propositional attitude if and only there is an operative (undefeated) reason for the attitude in the individual's *psychology*—whether or not the reason is in-principle accessible to the individual. A narrower conception of being justified would require in-principle accessibility to consciousness. We might label being justified in this sense '$\textit{justified}_{\mathrm{cp}}$', where the subscript indicates that an individiual is justified in holding a propositional attitude if and only if there is an operative (undefeated) reason for the attitude in the individual's *conscious psychology*, or at least that part of the psychology that could be brought to consciousness through introspection, prompting, or the like. One might use this distinction to produce correspondingly different notions of entitlement. I will return to this distinction later in the essay.

I take reasons to be the abstract representational contents, marked with a certain mode, of psychological states. Strictly, the abstract reasons do not cause anything. Only the psychological states do. Strictly, an operative reason is one that is the mode-content of a psychological state or occurrence that has that content and that figures causally in forming or sustaining the attitude whose mode-content the reason is a reason for.

[3] The explanation need not be in meta-representational terms. It is not essentially *about* belief or truth. It is fundamentally at the same level as the belief: *p* because *r*, where *p* is the content of the belief and *r* is the reason.

An individual can have both a justification and an entitlement for the same belief. Epistemic warrants derive from meeting standards for having epistemically good propositional attitudes, or for making epistemically good transitions among such attitudes. One can simultaneously meet different standards.

In the history of philosophy, epistemology featured justification, partly because of a focus on science—which aspires to justification. I focus first on entitlement.

A simple fact about human cognition is that we often have knowledge for which we lack a justification—a reason. Often an individual knows a proposition, and hence is warranted in believing it. Yet the individual may be unable to think a reason for the proposition. Or whatever reason that the individual does have, or could come to have, may not be operative in the individual's believing the proposition. The belief is not caused or sustained by a reason. Yet the individual is warranted in the belief, and the belief may even constitute knowledge.

A vivid example is a very young child's perceptual belief that red sphere is there. The child may lack the concepts necessary to having a reason for the belief. The simplest reason might be: I am having a perception as of a red sphere there. There are developmental reasons to believe that children have beliefs about colors, shapes, bodies, and locations before they can think about psychological states like *perceiving*. Even if they have psychological concepts (like perceiving) innately, their having meta-beliefs about perceivings is certainly not the primary warranting basis for perceptual beliefs about spheres. Children first form perceptual beliefs about spheres without any *thought* about perceivings. They form beliefs directly from their perceptions. They are warranted in doing so. Operative *reasons* come later.

An analogous point applies for perceptual beliefs in mature adults, including scientists. They may be able to cite their perceptions in rationalizing their perceptual beliefs. They may develop more articulated reasons. But they are first warranted in their perceptual beliefs because the beliefs were formed via a good, reliable, truth-conducive psychological competence, not because the believers can provide a reason that explains the beliefs' belief-worthiness. Their justifications are posterior to their entitlements. The justifications are not needed for the entitlements to hold.

A similar point applies to *transitions* in a deductive inference. To be warranted in believing the conclusion of an inference, one must be warranted in believing the premises and in relying on the inferential transitions. Children and some adults make warranted propositional inferences even when they are not in a position to think the inference rules that help explain their inferential transitions. One can make an inference that relies on a transition that is correctly explained as an instance of *modus ponens*. One can be entitled to rely on the transition in coming to a warranted conclusion, without *being able to* think the rule as a justification for a transition. Inference rules are meta-representational schematic generalizations about propositional contents. They hinge on isolating and representing the logical constants (here the conditional) on which the transition depends. Being competent to make a *modus ponens* inference does not require a capacity to think schematic generalizations or to think about propositional contents or logical constants—much less a capacity to think the rule and use it to justify a transition. Entitlement to rely on such transitions does not require *justifications* that cite the rules that codify and explain the transitions. They do not even require a *capacity* to cite the premises *as a reason* for the conclusion.

So at the most elementary level of empirical belief and of inferential transitions, entitlement (warrant without reason) precedes and is independent of justification (warrant with reason). Warrant without reason is warrant without full understanding. One can do well cognitively without being able to explain and rationalize what one is doing.

Many pieces of good reasoning mix entitlement and justification. Perceptual beliefs that one is entitled to can form premises for inferences. As premises, they can be reasons for an inference's conclusion. As just noted, one can be entitled to make inferential transitions without thinking the inference rules as justification for the transitions, even unconsciously. Thinking correct rules is not easy, even for mature thinkers. Even when one can think the rules, this ability is often not operative in a transition. Except for when an individual deliberately carries out an explicit proof, individuals who do think an inference rule as justification usually provide a further warrant, after the fact. Warrant for the conclusion of an inference is a combination of warrant for the premises and warrant for the transition-inferences. Even where the premises constitute reasons, justifications, for the conclusion, the reasons are often mediated by inferential transitions which the inferrer is entitled to, but not justified in. The transitions are warranted, but not justified by operative representation of the rule (consciously or unconsciously). So the full warrant for the conclusion is a mix of justification and entitlement.

In all inferences, a warrant for believing a conclusion depends not only on the reason-giving powers of the premises, but also on warrants for relying on the transitions. If the warrant for relying on a transition is an entitlement, the warrant for believing the conclusion will be a mix of justification (from antecedent steps for the propositional steps, including the conclusion) and entitlement (to the inferential transitions). Entitlement resides in an actual competence to make the relevant deductive transitions, not in an ability to understand and represent the rule governing the competence.[4]

When one is warranted in believing a conclusion because of an inference to it from premises, one commonly has some justification for the conclusion. It is natural to hold that the premises justify—provide reason for—the conclusion. I think that it is natural and correct to hold this even when one is just entitled to the transition steps, as long as the premises, together with the rules governing the

[4] Of course, although the premises in empirical inferences are often reasons for the conclusions, the premises themselves may be warranted by entitlement, not reason. As I have indicated, the basic warrants for perceptual beliefs are entitlements, not justifications.

transition steps, constitute a rationalizing explanation of the belief- worthiness of the conclusion.

I conjecture that, in *deductive* inference, when one is warranted in accepting a conclusion by virtue of the inference, the essential premises of the inference are always reasons for the conclusion. Often premises in good inductive inferences are reasons—justifications—for the conclusion. The premises constitute reasons, justifications, for the conclusion, by way of the inference, if but also only if those premises, when combined with the rule of inference (whether or not the inferrer understands and is justified in relying on the rule), yield some rationalizing explanation of the belief-worthiness of the conclusion. But as I will soon explain, this condition is probably not always met, even when one is warranted through the inference in believing the conclusion.

I have mostly concentrated on deductive inference. We do not understand induction very well. We do not have a theory of induction comparable to deductive logical theory. Many inductions—including what we call 'inferences to the best explanation' (abductions)—are complex, hard to articulate, and partly unconscious. There have been many attempts to codify induction into a logic. The best of these is probably Bayesian subjective probability theory. But what we call 'inductive inference' is probably a motley of significantly different kinds of transitions.

Some of what we call 'inductive inference' may not be genuine, reason-giving inference, even when it is warranted. Of course, any transition from one or more propositional attitudes to another one, according to some warranting transition pattern, is propositional inference. If inference is to provide a reason for a conclusion, the premises, together with the inferential transition, must constitute some sort of explanation of the acceptability of the conclusion. They must provide some answer to a 'why' question. They must be of the form: p because q, r, and s. The conclusion may be taken as only likely, relative to the inference from the premises. Or it may be taken as more reasonable than not, or as enhanced in credibility, relative to the inference. But if the inference is reason-giving for the inferrer, the premises and the transition rules must combine to go some way toward providing an (object-level) explanation for the individual inferrer of why the conclusion is credible. That is a minimum necessary condition on the premises' functioning as reasons for the inferrer.

As noted, even with respect to deductive inference, one need not be able to think the inferential rule that 'because' stands in for. One can be entitled to a transition without being able to explain the connection. One need not have a complete or fully satisfying explanation of the conclusion's being made credible by the premises, even at a non-meta-representational-, object-level of thinking. But if the premises are to provide reasons for the conclusion, they must provide some sort of rationalizing, explanatory support for the belief-worthiness, through the nature of the inferential transition, of the conclusion for the individual.

Spelling out this requirement is difficult. I do not know how to do so. Perhaps an example will help. Suppose that a human or higher animal has a perceptual belief. Suppose that some aspect of the belief triggers, by a natural psychological competence, an inference to a non-perceptual belief, such as a belief that it would be dangerous to move. Suppose that the connection between the truth of the perceptual belief and the truth of the triggered belief is a good, probabilistic one. Suppose, however, that nothing in the individual's experience could explain the connection. Perhaps the individual has no evidence that supports the inference. Perhaps the connection was innate, selected through evolution. It is well known that intuitively *very* unobvious probabilistic connections between properties can be significant and valuable psychological transitions.[5]

Of course, there is a biological/psychological explanation of why the perceptual belief is connected to belief about danger. There is some causal or statistical explanation of the relation in the environment between the perceived property and danger, that grounds the psychological connection. But the perceptual belief that is the premise may not provide the slightest explanation of the belief-worthiness of the conclusion for the individual. The individual and the individual's psychology cannot use the premise to rationalize, explain, or make sense of the belief-worthiness of the conclusion—even unconsciously. For the individual, it constitutes no *reason* for the conclusion.

In an abstract sense, one might claim that the premise *is* a reason for the conclusion. It is a small part of an explanation that could, in principle, be filled in by evolutionary theory or by an ideal epistemology. I think, however, that the premise would not function as a reason in the individual's psychology. The transition rule in the psychology may provide no insight into the connection, even for a theorist who knows the rule. The rule may be as simple as: if perceptual belief has concept $\underline{F}$ in it, belief about danger is to be formed. Only an explanation of why the rule is in place explains and rationalizes the belief-worthiness of the conclusion, given the premise. Such an explanation would have to show why the environmental property correlates with danger, and why the connection between the properties in the environment grounds the psychological connection.

Not only does the believer not understand a reason connecting premise and conclusion. The contents of the premise states that are in the (possibly unconscious) psychology combined with the rules of inductive transition do not add up to what even a super-psychologist would count as providing the slightest hint of a reason-explanation of the conclusion. As noted, a full account of why the rule came to be in place could show the premise as reason for the conclusion. From a God's eye point of view, one could reasonably say that the premise is a reason for the conclusion. But it is not a reason for the individual. It is not a rationalizing,

[5] W. S. Geisler, 'Visual Perception and the Statistical Properties of Natural Scenes', *Annual Review of Psychology* 59 (2008), 10.1–10.26; J. Burge, C. C. Fowlkes, M. S. Banks, 'Natural Scene Statistics Predict How the Figure-ground Cue of Convexity Affects Human Depth Perception', *Journal of Neuroscience*, 30 (2010), 7269–7280; W.S. Geisler, 'Contributions of Ideal Observer Theory to Vision Research' Vision Research 51 (2011), 771–781; Tyler Burge, *Origins of Objectivity* (Oxford: Oxford University Press, 2010), 359–366.

explanatory premise within the individual's psychology. I see too little in such an inductive inference to count it as reason-giving for the individual.

The premise and inferential transition still entitle the individual to believe the conclusion. For accepting the conclusion and relying on the inferential transitions are certainly as warranted as forming perceptual beliefs. In both cases, one uses a natural psychological competence that in fact constitutes a good, reliable route to getting things right.

What matters here is not whether or not one calls such inferences 'reason-giving'. What matters is to fix ideas on how dumb and non-rationalizing inductive transitions can be, while still providing entitlement to believe their conclusions.

It seems to me that more of our elementary induction may take this form than we philosophers are inclined to think. Patterns of inference are hammered into us by evolution. The principles that they instantiate and that explain them may yield little insight into why the conclusion is rationalized, made sense of, or explained by the premises. So, even knowledgeable reflection on the principle or rule according to which the inference is carried out may not show the premises to be reasons for the conclusion for the individual who carries out the inference. We do not know enough about how we actually carry out inductions to know whether the foregoing is so. But I think that we should be open to the possibility, indeed, I think, likelihood.

For some decades, computers have solved problems that are so complex that human beings lack the time and cognitive power to check the solutions. This situation emerged dramatically with the proof of the Four-Color Theorem by Appel and Haken in 1976.[6] Although any given step can be checked, the length of the proof prevents humans from checking the whole proof. Other computers checked the proof, however; and the theorem is considered proved.

This situation has grown more complex in the succeeding thirty-five years. Other important conjectures have apparently been proved with the help of computers. Checking them has sometimes been harder than checking the proof of the Four-Color Theorem. In 1998 Thomas Hales announced a proof of Kepler's conjecture.[7] The proof combined traditional geometrical analysis and

[6] K. Appel and W. Haken, 'Every Four Color Map is Colorable, Part I: Discharging', *Illinois Journal of Mathematics* 21 (1977), 429–490; 'Every Four Color Map is Colorable, Part II: Reducibility', *Illinois Journal of Mathematics* 21 (1977), 491–567. The theorem states that given any separation of a plane into contiguous regions, producing a figure called a map, no more than four colors are required to color the regions of the map so that no two regions that share a border segment have the same color.

[7] Thomas C. Hales, 'A Proof of the Kepler Conjecture', *Annals of Mathematics*, *2nd Series* 162 (2005), 1065–1185; 'Historical Overview of the Kepler Conjecture', *Discrete & Computational Geometry: An International Journal of Mathematics and Computer Science* 36 (2006), 5–20; Thomas C. Hales and Samuel P. Ferguson, 'A Formulation of the Kepler Conjecture', *Discrete & Computational Geometry: An International Journal of Mathematics and Computer Science* 36 (2006), 21–69. The Kepler Conjecture states that the highest density that can be achieved by filling a three-dimensional space with equal-sized spheres is *pi* divided by the square root of 18, or about 74%. This is the density of spheres stacked in a regular pyramid.

computer proof, with extensive descriptions of the computer portions. Several years of trying to check the proof by a committee of referees ended in failure. Again, the proof was too long to be checked by humans. But the computer program was not written with a view to being checked by computers. The referees regarded the proof as 99% *likely* to be correct. Subsequently, Hales launched an attempt to produce a fully formal proof that can be verified by computers. Hales regards it as a multi-year project. It is currently in progress.

Since computers are programmed to run *understood* mathematics the inferential transitions that they model in proofs, taken one by one, certainly give reasons, to humans who understand the mathematics and rely on the computers, for believing later inferential steps. Mathematicians understand steps as *rationalizing* later steps—providing some explanation of why they are belief-worthy. So the inferential transitions provide them with reason for the conclusions of the transitions. The problem, relative to classical conceptions of proof, lies simply in the uncheckably large number of transitions in the more complex proofs.

In earlier work, I gave an account of the epistemic status of relying on computers to complete and check a proof.[8] The account compared gaining knowledge from computers with gaining knowledge from communicating with other people. Computers are not people. However, their outputs have language-like character; and they produce propositional knowledge in a receiver much as human communicators do.

In exchanges with other people, I think that we have a default prima facie entitlement to accept what they assert unless there is reason not to. The entitlement is grounded in the prima facie rationality of a being that makes propositional assertions. Prima facie rationality implies prima facie competence and openness, at least on ordinary topics. As a constitutive point about rationality, rational individuals tend not to make incompetent assertions on such topics, although of course they sometimes do so. And rational beings tend not to lie without special reason. Lacking evidence that one's interlocutor is irrational, and that the subject matter is one on which competence and openness cannot be assumed, and lacking reason to believe that one's interlocutor has a special reason to lie, one can rely on one's interlocutor's prima facie rationality. I believe that recognizing such a default entitlement is the right way to account for childrens' gaining knowledge by being told things that they cannot evaluate, for adults' gaining knowledge by asking for unproblematic information from strangers, and for students' gaining knowledge when they are introduced to a new subject. This default entitlement structures most further warrants that can support or undermine an interlocutor's credibility on given occasions.[9]

[8] Tyler Burge, 'Computer Proof, Apriori Knowledge, and Other Minds', *Philosophical Perspectives* 12 (1998), 1–37; reprinted in this volume, section II. The discussion over the next few paragraphs summarizes some of the points in this article.

[9] See my 'Postscript: Content Preservation', this volume, section II.

In using computers to solve difficult problems, however, this default prima facie entitlement never suffices to yield knowledge, or even warranted belief. For the extreme difficulty of the subject matter grounds the need for supplementary warrant to believe that the computer is specially competent to solve the problems that mathematicians rely on it to solve.

Various types of warrant can support relying on a computer as credible. One division is between those that concern a machine's physical make-up and those that concern the content of its outputs. Knowing that a computer is made of stable materials, that its workings are predictable by laws of physics, that its power source does not die quickly, and that it is physically complex can help support believing that it could solve a difficult problem.

Such physical evidence is inevitably indirect. I think that it is not essential to warranting belief in a computer's outputs. Although the computer's success depends on the stability, reliability, and complexity of its physical operations, we do not *have to* know anything about such operations to have reason to believe that a computer's outputs are credible. Consider relying on a human mathematician or empirical scientist for information. Although the human must have a stable, reliable, and complex brain to solve a hard problem in mathematics or science, one need not know anything about biology or physiology to learn to trust his or her statements. One's warrant can derive entirely from considering the content of the statements.

The second, and primary, category of warrants that support relying on a computer depend on evaluating the content of the computer's outputs. For example, a mathematician can check shorter difficult proofs, or difficult parts of a long proof. One can study its program. Continuities in a proof and similarities of argument structure among different proofs enable one to ascribe highly competent processing to a single source, even apart from identifying the source as a single physical machine. One can check one computer by relying on another. The second computer can use different methods to solve the same problem. Thinking through such outputs and performing an induction on cases can provide warrant to rely on a computer to solve hard problems, even when one cannot check the full proof. Warrants that rely on understanding the content of a computer's output are the primary warrants for believing what a computer produces.

Warrant to rely on computers in solving hard problems bears comparison to warrant that a gifted mathematician has to rely on his or her own powers in arriving at mathematical beliefs without proof. Imagine a Ramanujan-type genius, who can do proofs, but who often proposes unproved mathematical results. Repeatedly, but not infallibly, the results are proved later. The individual has only a sketchy sense of how he or she arrives at the results. The individual firmly believes the results, however, even before they are proved.

I believe that a reliable individual, like Ramanujan, is often *warranted* in believing such results, before they are proved. The individual has exceptional mathematical powers and uses them well in forming true beliefs.[10] Lacking knowledge of Ramanujan's psychology, we cannot know exactly what type of warrant he had. I presume that no proof occurred in his psychology. If unconscious proofs were carried out in this psychology, he would have a justification at the unconscious-but-personal-level.[11] Perhaps he had inductive reasons to believe the theorems—unconscious but in principle available to consciousness. If his routes were non-propositional (say, pictorial), or if they were propositional but did not constitute an operative, explanation of the conclusion, he would have an entitlement. Then he would be like a person who reliably forms perceptual beliefs, but lacks a reason *anywhere* in his psychology that provides some explanation of the conclusion's belief-worthiness.[12]

A mathematician like Ramanujan is likely to have another warrant. The mathematician can develop an inductive *meta-representational* justification for relying on the powerful competence. He or she can reason: I have reliably come up with answers that have later been proved; so I have a reliable competence. The mathematician can strengthen the primary warrant or this meta-representational warrant by showing that a given result of the unconscious competence coheres well with other proved propositions, or by producing parts of a proof for it.

Thus the gifted mathematician can have both unconscious entitlement for belief in particular propositions and conscious, reasoned justification that falls short of proof. By contrast, a person who relies on a computer to solve difficult problems can, I think, have only inductive justification for believing the computer's results. Since the mathematician cannot produce solutions to the problems in a reliable way, the mathematician must have some inductive reason to accept the computer's offerings. The mathematician cannot rely on a general default entitlement to accept what one is told, other things equal. Other things are not equal. The known difficulty of the problems demands a reason to justify reliance on the computer's outputs. The justifications can be inductive. They have the defeasible, prima facie character of all inductive reasoning.

Even though reliance on computers to carry out proofs inevitably involves an inductive element, the way in which this element compromises the deductive heart of mathematics does not go very deep. Relying on computers to carry out proofs seems in some ways not substantially different from one mathematician's relying on another, as a source about what has been proved.

A deeper reason for rejecting the view that mathematics concerns only proof has been available for eighty years. Gödel's results show that for any given system of consistent axioms of sufficient power (including axioms as weak as

[10] I am assuming that the individual understands the relevant mathematics, can explain its significance, and can relate it insightfully to other mathematics. Thus the individual is not an idiot *savant* with marvelous calculating powers, but no broad mathematical competence. I will not take a position here on the epistemology of the beliefs of idiot *savants*.

[11] He would be justified$_{\mathrm{p}}$ but not justified$_{\mathrm{cp}}$. See note 2. In any case, he would still be warranted in his beliefs.

[12] I think it unlikely but possible that Ramanujan's warrants were mainly entitlements.

those for arithmetic), there are truths expressible in that system that are not provable from those axioms.[13]

Other mathematical practices also show that proof from compelling starting points is not always expected in mathematics. Many proofs in advanced descriptive set theory, for example, begin with unproved propositions that are postulated as plausible but that are far from compelling. The value of the postulations and proofs from them lies in unifying and helping to explain other results in set theory.[14] Of course, computers can be used in these proofs, as well as in proofs that begin with compelling assumptions.

Moreover, much work in descriptive set theory involves inductive support for conjectures that may frame further work—with no immediate expectation of proof for the conjectures.

Some uses of computers in mathematics go well beyond producing and checking proofs. A significant departure from traditional uses of computers in mathematics consists in computer-driven *probabilistic* tests for the truth of a hypothesis. There are, for example, probabilistic ways of determining whether a number is prime or for solving combinatoric problems. The computing itself is algorithmic and deductive. However, its starting point is a sample of possibilities. The result is argued to be true on probabilistic grounds, often assigning specific probabilities.[15]

Although expanded probabilistic uses of computers in pure mathematics is inevitable, there is some inertial resistance to such expansion in the mathematical community. Some of this resistance is simply a natural unease, born of the worry that any movement away from the gold standard of mathematical argument—rational compulsion through proof—may ultimately compromise mathematical standards. I think that no one doubts that proof from conceptually compelling

[13] It is sometimes said that the view that mathematical truth must be distinguished from mathematical theoremhood depends on specialized, doubtful philosophical views, such as Gödel's platonism. For an example of such confused writing, see Brian Davies, 'Wither Mathematics?', *Notices of the American Mathematical Society* 52 (2005), 1350–1356. The evidence of the truth of the relevant unprovable sentences does not depend at all on philosophical views. Contrary to Davies' assertions, the point depends neither on Gödel's platonism nor on reliance on an allegedly specialized notion of truth, such as Tarski's. I do regard ontological platonism about central mathematics as the natural and correct view. Of course, some mathematical problems are, inevitably, afflicted by vagueness in their key concepts. Then one cannot expect determinately true or false answers regarding determinate abstract structures. I think that we have no recipe—in particular, no general strictures (such as that vagueness is present when no proof is possible)—that indicate when and where there is vagueness in our mathematical concepts.

[14] Donald M. Martin, 'Mathematical Evidence', in H. G. Dales and G. Oliveri (eds.), *Truth in Mathematics* (Oxford: Clarendon Press, 1998), 215–232.

[15] David H. Bailey and Jonathan M. Borwein, 'Future Prospects for Computer-Assisted Mathematics', *Notes of the Canadian Mathematical Society* 37 (2005), 2–6; Leonard M. Adelman, 'Molecular Computation of Solutions to Combinatorial Problems', *Science* 266 (1994), 1021–1024; C. W. H. Lam, 'The Search for a Finite Projective Plane of Order 10', *American Mathematical Monthly* 98 (1991), 305–318; Carl Pomerance, 'The Search for Prime Numbers', *Scientific American* (December 1982), 136–147; Michael Rabin 'Probabilistic Algorithm for Testing Primality', *Journal of Number Theory* 12 (1980), 128–138.

starting points is to be preferred where it can be obtained. One simply has to maintain perspective on the relevant epistemic statuses of different methods. Sometimes one can use less traditional methods to put one in a better position to use traditional ones. But mathematics is very large, and mathematicians are comparatively small. Some mathematical problems simply do not submit to traditional methods. They can still be worth investigating.

I shall assume that non-traditional, probabilistic uses of computers in pure mathematics are intellectually worthwhile. I shall assume that they yield at least knowledge of the high probability of the truth of conclusions. I ask whether there are epistemically principled differences between probabilistic uses of computers and traditional methods of proof that include computer-assisted proof. I think that there are.[16]

The differences are not matters of rational certainty.[17] Although many proofs in elementary mathematics are rationally certain, use of inductive methods in set theory and other areas of pure mathematics is standard. Such methods cannot yield rational certainty. Computer proofs and long proofs that involve many mathematicians are also not rationally certain: One could, in principle, have reasonable grounds to doubt them, even when they are correct.

The differences are not matters of apriority. Arguments from self-evident axioms, broadly inductive-explanatory arguments, and probabilistic arguments can be all be apriori warranted. The force of the warrant for believing them can owe nothing to sense perception. Moreover, reliance on other mathematicians or on computers in carrying out deductive proofs[18] and reliance on computers to carry out probabilistic inferences are both warranted empirically.[19]

[16] In what follows, I disagree with a main conclusion of Don Fallis, 'The Epistemic Status of Probabilistic Proof', *The Journal of Philosophy* 94 (1997), 165–186. I agree that there are no good epistemic grounds for mathematicians to *reject* non-traditional probabilistic methods in pure mathematics. I do not agree that 'there is no important qualitative difference between [probabilistic uses of computers in pure mathematics] and the [more deductive, traditional methods] acceptable to mathematicians' (166). Although I disagree with the letter of the argument, I find its spirit largely congenial: probabilistic methods in mathematics are epistemically warranted and entirely legitimate. Fallis focuses exclusively on the epistemic value—*supporting the truth of mathematical claims*. As he says, probabilistic methods can make a mathematical truth just as credible as a very long human or computer-assisted deductive proof can. I think, however, that his article underplays epistemic values that reside in *how* a belief is warranted and what sort of understanding a given method provides. Probabilistic methods do not yield a direct understanding of mathematically necessary structures and relations.

[17] A belief is *rationally* certain if no possible rational consideration can justify rational doubt. *Psychological* certainty is just unshakeable total confidence. Such certainty can be irrational. Moreover a belief can be rationally certain without being psychologically certain, if the believer is overly cautious or timid, or does not sufficiently understand the power of his or her reasons. Although a lot of discussion of probabilistic methods focuses on their (high) degree of psychological certainty, psychological certainty is of no particular interest to epistemology.

[18] I note that I have given up my earlier view that reliance on computers or other human beings' reports can be strictly apriori warranted. See 'Postscript: "Content Preservation"', this volume, section III.

[19] A striking instance of empirical computation is Adelman's use of a computational interpretation of experiments involving DNA to solve a mathematical problem. See Adelman, 'Molecular

I believe that the epistemic difference lies in the sort of understanding associated with the different methods. Deductive inference yields understanding of why a mathematical truth must be true, at least relative to the premises. Purely probabilistic inference does not.

Understanding is not a precise notion. There are some ways of caricaturing the distinction that I am trying to get at. I want to resist such caricatures.

First, one should not think in terms of full understanding. Deductive mathematical practice allows many forms of incomplete understanding obtained from a deduction from self-evident starting points. A computer-assisted proof, like the proof of the Four-Color Theorem, does not give anyone full understanding of the proof. The mathematician knows much of the proof, understands the principles used in it, and has inductive reason to think that the computer has carried out a proof. Understanding is partial. It is partial understanding of how the proof goes, backed by inductive ground to believe that the proof has been completed. It is partial, idealized, but genuine understanding of the necessity of the conclusion relative to the premises.

The computer-assisted deductive proof is similar to a mathematician's sketching a proof and recognizing that the proof is completeable.[20] It is also similar to a mathematician's having in mind a proof of a closely related proposition and recognizing that a variant proof for the different proposition is viable—all, without actually going through the proof. Part of being a good mathematician is being able to recognize and understand, in a rough how-it-would-go way, the provability of a proposition. Such recognition provides understanding, even if not rigorous complete understanding, of the necessity of the conclusion relative to the premises.

It is common among writers on these subjects to point out cases in which individuals make mistakes in purporting to recognize that a proposition is provable. But centering on these cases tends, I think, to miss the most important point. The most important point is that in these cases, there is commonly a genuine mathematical competence that grounds epistemic warrant. The recognition competence is inevitably fallible. It varies in power and reliability with different mathematicians.

However, the competence is a reliable capacity in trained mathematicians that commonly warrants their belief in unproved but provable propositions. Assuming that a worked-out, operative proof is not present in the mathematician's unconscious psychology, there remain reasons—proof sketches or systematic analogies

Computation of Solutions to Combinatorial Problems'; and Keith Devlin, 'Test Tube Computing with DNA', *Math Horizons* 2 (1995), 14–21. The mathematical computation clearly is warranted partly through warrants to believe biological theory about DNA.

[20] This kind of situation is *very* common in mathematics. Gödel only *sketched* his second incompleteness theorem in 1931. The proof was not written out, or probably even thought through in full detail, by anyone until 1939. The mathematical community recognized the theorem as "proved" much sooner. See John W. Dawson, 'The Reception of Gödel's Incompleteness Theorems', *Philosophy of Science Association 1984* 2 (1985), 253–271.

to other proofs—that are available to the mathematician. These reasons can combine with an entitlement to rely on a capacity to transition from these reasons to belief. Thought, even unconscious thought, about the specific rule that governs these transitions may not (indeed, I am sure, often does not) figure operatively in the transition. So the warrant for relying on the transition is an entitlement, not a justification.

We do not know enough about the psychology of mathematicians to articulate these warrants. It seems likely, however, that the transition rules are often not represented and operative in the mathematicians' psychologies. And it seems even more likely that the whole process, including both the reason sketches and the transitions, often provides warranted understanding. Such understanding is *incomplete* understanding of why a mathematical proposition is a necessary consequence of the inference's premises.

In sum, there are different kinds of understanding of the necessities associated with deduction. The ideal is full deduction from first-principles. However, the starting points of deduction need not be compelling. They can be acceptable because they help explain other mathematical truths. And the understanding can be sketchy and incomplete, while still pointing toward how the explanatory reasons of a deduction would go.

The understanding need not be general. Brute force proofs in finite domains determine all relevant possibilities, though the truths are individual to the problem. (I am thinking, for example, of finite, brute force solutions to versions of the traveling salesman problem.) Understanding does not derive from general principles, and may provide little general insight. Still, such deductions yield singular understanding of necessary truths and necessary relations between truths.

Further, one should not caricature my point to mean that *probabilistic* proofs provide no understanding at all. Probabilistic proofs can yield insight into the structure of a problem, or inductive insight into mathematical relationships. They are part of an experimental approach to pure mathematics that is similar to the method of hypothesis and testing in empirical science.[21]

Deductive proof gives grounds not just for believing a proposition, but for understanding the necessity of its truth and its necessary relations to premises. It is not surprising that mathematical practice counts such understanding—even in idealized, partial form—epistemically distinctive. Mathematics aims to obtain such understanding, where possible.

I turn now from computers in mathematics to computers in empirical science. Empirical science cannot aspire to deductive proof from propositions that are, when understood, compelling. It must reach its conclusions through induction from perceptual belief.

Uses of computers in empirical science are more varied than those in pure mathematics. The variety of types of mathematical applications to empirical

[21] J. M. Borwein and D. H. Bailey, *Mathematics by Experiment: Plausible Reasoning in the 21st Century* (Wellesley, Mass.: A. K. Peters Ltd, 2004; 2nd edition 2008).

problems seems nearly endless. I discuss briefly just three such applications. I relate each to issues in human cognition.

A common computer application in biology, psychology, and engineering is computation of an optimal solution to a problem. In some such applications, the program is governed by a fixed, although possibly multiply re-applied, mathematical formula. In biology and psychology, such computations are often carried through to determine a baseline for testing actual performance—for example, by a colony of ants in foraging for a food supply, or by a visual system in encoding proximal stimulation for accurately representing some environmental condition.[22]

These computations consist in applying well-understood—often Bayesian—mathematical operations to huge data sets—perhaps repeatedly altered by Monte Carlo randomization—at very fast rates. These uses differ from human reasoning mainly with regard to length and complexity. Given a task with a clear goal, given a data set, given prior probabilities and limiting conditions, the computation provides an optimal solution to the problem, considering costs and benefits.

There are complications. Being certain that a solution is optimal often depends on a separate proof that there are no local maxima. Often one does without certainty and takes a computation of a local optimality result to probably apply globally. Even when one has a proof, the putative certainty lies in the machine's proof, not in the scientist's conclusion. The warrant for the scientist's conclusion consists in an inductive justification based on understanding the machine's program and reliability. The computer run is, however, often a proxy for reasoning that except for its complexity could be carried out by the scientist.

The Bayesian transitions in a computer run are commonly in a form that makes them propositional proxies for human reasoning *about* non-rational processes. Of course, the non-rational processes lack propositional structure. The subject matter of the programs—for example, ants' foraging behavior or processing in a visual system—are not pieces of reasoning. Even where the scientist has not fully thought through the computer's transitions, the scientist has a general

[22] Peter Nonacs and Joanne L. Soriano, 'Patch Sampling Behaviour and Future Foraging Expectations in Argentine Ants, *Linepithema humile*', *Animal Behavior* 55 (1998), 519–527; Sasha R. X. Dalla, Luc–Alain Giraldeaub, Ola Olssonc, John M. McNamarad, and David W. Stephense, 'Information and its Use by Animals in Evolutionary Ecology', *Trends in Ecology and Evolution* 20 (2005), 187–193; Thomas J. Valone, 'Are Animals Capable of Bayesian Updating? An Empirical Review', *Oikos* 112 (2006), 252–259; Yoram Buraka, Uri Roknia, Markus Meistera, and Haim Sompolinskya, 'Bayesian Model of Dynamic Image Stabilization in the Visual System', *Proceedings of the National Academy of Sciences of the United States of America* 107 (2010), 19525–19530; Wilson S. Geisler, Jiri Najemnik, and Almon D. Ing, 'Optimal Stimulus Encoders for Natural Tasks', *Journal of Vision* 9 (2009), 1–16.

Sometimes optimization algorithms are developed from observation of frequencies in actual behavior, for example, the swarm behavior of ants or bees. Such algorithms are often carried over to apply to different domains, including pure mathematics. See M. Dorigo and T. Stützle, *Ant Colony Optimization* (Cambridge, Mass.: MIT Press, 2004). These cases are rather like the DNA computing mentioned in note 16, in that they use natural empirical phenomena to help solve mathematical problems, as well as problems in natural science.

understanding of the explanatory relation between premises and conclusion of the computer run. The scientist can, commonly, understand any given step as a reason for a later step. Even where the scientist's understanding is partial, the computer run can be interpreted as a realization of reasoning. The scientist can enter into the reasoning at any point. The computer's transitions, supplemented by the scientist's inductive reasoning about the computer's performance, yield justification—warrant by reason—for the conclusion.

The transitions effected by the computer are not operative in the scientist's psychology. The scientist makes an induction about the reliability of the machine's outputs and is justified in relying on them. The scientist is *better* off than in many cases in which the scientist relies on his or her own inductive reasoning. Transitions in ordinary inductions are, to be sure, operative in the scientist's psychology, whereas transitions effected by the machine are not operative in the scientist's psychology. Principles governing transitions in human inductions are, however, often less well understood than the principles governing transitions effected by the machine.

A second type of computation in empirical science raises more interesting epistemic issues. I have in mind [what are broadly called] '*genetic algorithms*'. Genetic algorithms are search and optimization techniques based on principles governing mechanisms of evolution. The algorithms require individuals in a population to be represented as solutions to a problem, and they require a fitness function that maps solutions to a quality-of-solution evaluation. An initial population of solutions is generated randomly. A proportion of this population is selected through fitness evaluation to begin breeding a new population of individual solutions. Selection is biased toward better solutions. A new population is produced by genetic operators—the most common being mutation and crossover (or recombination). *Mutation* applies a probability that a random aspect of a solution is to be modified in some usually random way to produce an individual solution for the next population. *Crossover* is a genus of operations that combine aspects of the solutions of two or more "parent" individuals from a given population to produce a new individual solution for the next population. Different parents are chosen from the selected sub-population to produce new individuals. When a new population of a certain size is produced, the process of fitness evaluation, selection, and genetic operation is reapplied. The process is usually terminated when a given optimization level is reached, though it can be terminated after a set number of generations.[23]

Such algorithms are very good at producing optimal solutions in large, complex domains. An optimal solution can be used in science as a baseline for generating empirical hypotheses about actual domains and for empirically testing the nature of such domains.

[23] Melanie Mitchell, *An Introduction to Genetic Algorithms* (Cambridge, Mass.: MIT Press, 1996); J. R. Koza, *Genetic Programming: On the Programming of Computers by Means of Natural Selection* (Cambridge, Mass.: MIT Press, 1992).

What is the epistemic status of a scientist's belief that a certain population of solutions is optimal, when that belief rests on the result of a computer's running a long, complex genetic algorithm? As before, the belief can be epistemically justified through inductive reasoning about the program and the computer's reliability in implementing it. It seems to me, however, that there is a difference with the case in which a scientist bases a belief on the generation of an optimality result that derives from Bayesian updating.

As noted, like genetic algorithms, Bayesian updating processes often yield results about non-rational processes. But Bayesian updating processes are, at least on some formulations, proxies for *reasoning* in Bayesian probability theory. Each stage in the algorithm is naturally construed as a reason for the next stage. Genetic algorithms are not proxies for reasoning. The random mutations and relatively random recombinations are not analogs of reasoning. They lack the explanatory structure of reason transmission. The computer's transitions do not transmit reasons from previous steps to later steps. Earlier processing stages are not reasons that help constitute answers to why-questions about the credibility of later steps. The computer does help the scientist reason about a non-rational process. Its runs provides evidence for scientific inductions about the relevant subject matter. The computer's processing is *not*, however, to be construed as an instantiation or amplification for reasoning.

Given our understanding of evolution, we understand why the non-rational process yields an optimality result. We are justified—in principle *just as justified*—as in any other inductive case. But the computer's processing does not serve as a proxy and amplification for our own reasoning—as it does in carrying out mathematical proofs and in many construals of Bayesian algorithms. Of course, the evolutionary process itself is not an exercise of reasoning. Computer processing that serves as a proxy for that process is not a simulation of reasoning.

I emphasize that these reflections do not bear on the power of our warrants for believing the computer's results. Inductive warrants can be equally strong in the different cases. The reflections bear on the relations between natural construals of particular transitions in the computer's processing, on one hand, and steps in our reasoning, on the other.

I mentioned earlier that some warranted inductive propositional transitions in our own psychologies may not always constitute reasoning. Previous steps may enhance credibility of later steps without providing explanatory rationalization of the later steps—without being reasons for them. Some propositional inductive processing in our own psychologies may be as non-rational as evolution is, while still being a good route to truth. Given that our minds reflect nature in many other respects, it would surprising if they did not do so in this one. Any such warrants to believe conclusions from inferences would be purely entitlements, not justifications. Even the premises of the warranted inferences would not be justifications—reasons—for the conclusions.

There are Hume-inspired conceptions of inductive inference that treat *all* inferences as non-reason-giving associative transitions. Some psychologists who take connectionist programs to model human reasoning tend in this direction. I think that such models are not adequate to model deductive reasoning or most verbally articulated induction. Much of the reasoning in mathematics and natural science *is* reason-giving. What I am noting is that there may well be inductive inferential transitions, even in science, that support their conclusions by making them reliably more credible, but without providing even a partial explanation of the sort that reasons provide. Such inferences yield entitlements to their conclusions, all the way down, not justifications.

I turn briefly to a third example of uses of computers in empirical science. Recently, some striking results have been obtained in producing algorithms for discovery of scientific explanations. A computer was given data on certain types of motion, such as that of a double pendulum. It was given basic arithmetic, geometric, and trigonometric operations, and some basic parameters. It was programmed with a genetic, symbolic regression algorithm—one that looked simultaneously both for parameters to plug into equations and for equations that simulate conservation principles. It was also programmed to balance simplicity of equation against accuracy in fitting the data. The algorithm returned a small number of equations (on the order of ten). Although some were of no scientific interest, some were classical laws of physics—such as Newton's second law of motion and Lagrangian equations that apply to the double pendulum.[24]

Here the genetic algorithm works on formulae, weeding out the "less fit" ones in favor of equations that fit the data and are simple. Again, particular transitions in the computer processing do not yield reasons for subsequent stages. And again, the scientist has only general insight into the computer's operations. The algorithm does simulate the scientific method of posing a question, offering a hypothesis, testing the hypothesis, adjusting the hypothesis until one finds one that tests better, and so on. But the generation and adjustment of hypotheses have the random character of all genetic algorithms.[25] The process produces better hypotheses over time because only the better fitting equations are selected for each new cycle of testing.

In scientific discovery of new laws, we know very little about human inferences to the best explanation. It is clear that hypothesis revision in our reasoning is often more directed and more based on rational considerations than it is in any genetic algorithm. Such algorithms can afford to search for and produce new hypotheses by more nearly random methods.

Human discoveries are often described loosely as leaps of intuition. In such cases, the scientist has little conscious recognition of how he or she finds a

[24] M. Schmidt and H. Lipson, 'Distilling Free-Form Natural Laws from Experimental Data', *Science* 324 (2009), 81–85. There are important issues here, which I shall not go into, about the scientist's understanding non-basic parameters that the computer comes up with and about the scientist's role in recognizing which among the equations that the computer comes up with are scientifically interesting and genuinely explanatory.

[25] By contrast, as far as I can see, the testing phase of the algorithm *does* operate as proxy for rational testing in scientific reasoning.

hypothesis. But discovery often derives from weighing reasons at various levels of generality—from observations of evidence to attempts to fit in mathematically with high-level generalizations that are already known. Normally, a lot of reasoning goes into the inductive discovery of new explanatory principles.

Even so, we do not know the full psychology of such inductions. We cover our ignorance with the all-purpose, but uninformative epithets 'intuition' and 'insight'. Although human induction to scientific laws commonly involves reasoning, it may not consist purely in reasoning. There may be non-rational, evolutionarily drummed-in associations that are not backed by any explanatory reasons that are operative in our psychologies. Reasoning may sometimes figure mainly in testing hypotheses.

I think that this Popperean picture of scientific discovery is at most one factor in the right overall account of inference to best explanations. I emphasize, however, that although much human induction (and all deduction) surely differs from the genetic algorithms used in the case just described, we know very little about human induction. We are not in a position to say that our unconscious psychologies never use the brute, non-rational, "natural selection" methods that occur in genetic algorithms. Thus our sense of not understanding the computer's processing as *reasoning* should not be allowed to obscure the fact that we do not understand our own inductive processing very well—especially in the case of inferences to new scientific principles.

What we should insist upon is that science *looks for* rational explanations that yield reasons where it can find them. Fundamentally, science attempts to produce reasoned explanations. Both the non-rational processing in computers that simulates non-rational selectional processes in nature and whatever elements of non-rational processing there are in our own unconscious psychologies must be used to find reasoned explanations.

Empirical science, like mathematics, aims at reasons, explanation, and understanding. Science is a particular form of understanding. It strives not just for knowledge, but for *scientia*—systematic knowledge backed by reasoned, explanatory understanding of why its conclusions are to be believed. Relying on computers underscores that understanding is often partial, even in science. Reason is not itself inadequate. But our capacities to reason are small in comparison to the vast complexities of our mathematical, physical, and psychological subject matters. We must rely on computers, and perhaps on our own non-rational but warranted inductive processing, for all the help we can get, even though such help guarantees that our understanding of some aspects of science will remain partial.