Inductive Reasoning in the Deductive Science

Jonathan Henshaw

17 November 2008

The purpose of this essay is to explore some issues that arise out of the interaction between inductive and deductive logic in mathematics. I begin with a general account of how belief and uncertainty figure in the physical sciences, and move via analogy to mathematics. I then provide a brief normative account of inductive reasoning in mathematics.

Belief and uncertainty

How best to make sense of our naïve concepts of belief and uncertainty is a controversial issue in the philosophy of mind. I will present an informal account of the words as I intend to use them, neglecting many of the complexities involved. This will hopefully be enough for our purposes.

We say that an agent believes a proposition p if she thinks that p is true. Such a belief is *explicit* if the agent has previously thought about p and retains a representation of the truth of p in her accessible memory. The belief is *implicit* if, given her current explicit beliefs, the agent could quickly derive the truth of p as an explicit belief. For example, Jane might believe explicitly that there are 6 states of Australia. However, her belief that the number of states is divisible by 3 is most likely implicit, unless she has been in the unfortunate circumstance of dividing a cake between them. Similarly, she might believe explicitly that $1 \times x = x$ for all x. However, if she has never come across the number 13964 before then the instantiation $1 \times 13964 = 13964$ is only an implicit belief of hers.

By this definition, there is a sharp distinction between explicit and implicit beliefs, at least if we ignore problematic cases such as when memory access is difficult or when an agent repeatedly derives a particular proposition from a simpler one, such as a mnemonic.[1] However, there is no sharp distinction between what an agent believes implicitly and what she does not

[1] For instance, some mathematicians derive identities like $\sin 2\theta = 2 \sin\theta \cos\theta$ by using Euler's formula $e^{i\theta} = \cos\theta + i \sin\theta$ every time they need them.

believe, because quickness of derivability is a matter of degree (Schwitzgebel 2008). This has important consequences for our discussion of mathematical propositions. Suppose, for instance, that Jane is introduced to the axioms of Peano arithmetic along with the definition of a prime number. Does she now have an implicit belief that:

- 2 is prime?
- 53 is prime?
- there are infinitely many primes?
- the prime number theorem is true?

All these propositions are logical consequences of the Peano axioms, but some can be derived much more quickly than others. While the primality of 2 is obvious to anyone who understands the definition of a prime, an elementary proof of the prime number theorem was not published until 1949, sixty years after the Peano axioms were first defined (Cf. Peano 1889; Erdös 1949; Selberg 1949). With this in mind, I will understand 'quickly derivable' to mean derivable in a matter of seconds without the aid of technology. Thus, for most people in Jane's situation, only the first proposition above would be an implicit belief. As an agent gains in mathematical knowledge and skill, we would expect not only the number of her implicit mathematical beliefs to increase, but also the ratio of her implicit to explicit beliefs.

On to uncertainty. We say that an agent is uncertain about a proposition p if she is aware of p but believes neither p nor $\neg p$. It is sometimes useful to quantify an agent's uncertainty about a proposition. We do this in the standard way, by assigning to the proposition a real number $\mathcal{B}(p) \in [0, 1]$, where $\mathcal{B}(p) = 1$ represents the agent's belief in p and $\mathcal{B}(p) = 0$ her belief in $\neg p$. We call $\mathcal{B}(p)$ the agent's *degree of belief* in p.[2] As in the case of beliefs, we can talk about implicit and explicit degrees of belief. Before you read this sentence, you probably had no explicit degrees of belief regarding the 10^{99}th digit of π. Now you do.

We might measure an agent's degree of belief about p by ascertaining what betting odds she will accept over the truth of p. If $\mathcal{B}(p)$ lies in the interval $(a, b) \subset [0, 1]$ then she will accept any odds better than $\frac{1-a}{a} : 1$ that p is true and any odds better than $\frac{1-b}{b} : 1$ that $\neg p$ is true.

[2]Of course, if we are strict about requiring that beliefs occupy only endpoints of the interval $[0, 1]$ then many people may have no beliefs at all. In practice, we may be laxer and call a proposition a belief if it is 'near enough' to the endpoints.

Of course, this is a very fragile measure. What if our agent is especially risk-averse, or loves to gamble? What if, as Eriksson and Hájek (2007) suggest, she is a Zen Buddhist to whom money has no utility? These examples undermine the reliability of betting behaviour as a measure of degrees of belief, but I do not believe they undermine the coherence of the very idea. In the same way, my ability to lie about my beliefs undermines my word as a reliable measure of them, but not the coherence of the idea that I have beliefs. In both cases, we are considering a state of mind that is imperfectly indicated by external behaviour. We will return to this idea later on.

Uncertainty in the physical sciences

We may crudely suppose that a scientific theory consists of a set of data together with a model that is designed both to agree with existing data and to make predictions about future data. I propose that uncertainty enters into such a theory in four main ways:

1. *Indeterminism*: some models assume that particular physical processes are inherently indeterministic. The Copenhagen interpretation of quantum mechanics is an example of this. In such cases, no amount of information about the system will allow us to predict its future states perfectly. Uncertainty is built into the theory.

2. *Data uncertainty*: it is often impossible to determine accurately the current or past states of a physical system. This may be due to instrumental imprecision, to an inability to survey some parts of the system or to a complete lack of data. Data uncertainty can result in uncertain predictions, even when using an accurate model. For instance, astronomers often have difficulty predicting the future trajectory of asteroids due to uncertainty about their position and velocity.

3. *Model uncertainty*: even if accurate data are available for a system, there may be either no good model or several competing models to explain it. Uncertainty results if the models make different predictions about future data. The most general formulation of this type of uncertainty is the problem of induction.

4. *Computational complexity*: even if accurate data are available and a good model exists for predicting future states of the system, the model may require computations that are too complex to perform exactly. Some of the first accurate weather predictions were so computationally complex that the forecasters had difficulty keeping up with the weather! (Lynch 2008). In these cases, heuristic or statistical calculations may be very useful. However, they introduce uncertainty that is not inherent in the model.

Of course, these causes of uncertainty do not normally operate in isolation. Indeterminism and (in principle) data uncertainty interact in the Copenhagen interpretation of quantum mechanics. Weak underdetermination of cosmological theories often arises because a lack of data renders competing models inseparable. In chaotic systems, uncertainty about initial states combines with computational complexity to make long-term predictions impossible. It is even possible for all four types of uncertainty to interact in a single situation. How certain are scientists that the Large Hadron Collider at CERN will not destroy the earth? Any good answer to this question would have to involve all four concepts of uncertainty.

Uncertainty in mathematics

Every mathematical theory works from a set of assumptions. These are propositions that the relevant mathematical community considers *basic*, in the sense that they are not in need of formal justification. In most modern mathematics, such assumptions take the form of axioms. Before the nineteenth century, the assumptions were usually much broader than this, and this is still the case with fledgling contemporary theories.

As well as a set of assumptions, a successful mathematical theory provides a set of procedures by which new propositions can be derived from the assumptions and from previously derived propositions. We will refer to such procedures as *rules of deduction*, keeping in mind that they may operate at a higher level than the formal rules of an axiomatic system.

Time for an analogy. In the same way as scientists use a model to extrapolate from existing data to new predictions, we can think of mathematicians as extrapolating from assumptions to new propositions by means of the rules of deduction. The analogy works best if we consider scientific models that predict future system states based on initial conditions. Models of planetary motion, population dynamics and weather patterns all work in this way. We can think of the assumptions as corresponding to initial conditions, while the rules of deduction correspond to the algorithm that predicts future system states from past and current ones. We may visualise the extrapolation in the mathematical case as a (generally infinite) directed rooted tree, where the nodes represent sets of previously derived propositions and the branches represent the application of a rule of deduction.

Where does uncertainty enter this mathematical picture? Assumptions and rules of deduction are not subject to the same types of scrutiny as the data and model of a scientific theory are. They may be criticized for being unproductive, restrictive, inelegant, unintuitive or boring. However, isolated from any application they are not usually criticized as *inaccurate*. As a mathematical theory need not be representational, there may be nothing external against which to check its accuracy. It is possible, and I think helpful, to consider *any* assumptions or rules of deduction to be true by default. The truth of a mathematical proposition is then relative to the background theory. For instance, parallel lines never meet in Euclidean geometry, but they do in projective geometry. In the absence of a particular theory, the question of whether parallel lines meet has no definitive answer. Contrast this with scientific theories. Newtonian mechanics does not predict the perihelion precession of Mercury, while the theory of general relativity does. However, the truth of Mercury's precessing perihelion is *not* relative to which theory we are discussing, as the observable phenomena provide a check for the correctness of the theory.

These considerations make it difficult to attribute uncertainty in mathematics to assumptions or rules of deduction. What about indeterminism then? It is true that probability theory deals with indeterministic mathematical processes analogously to how some scientific theories deal with indeterministic physical processes. However, probability theory is only one branch of mathematics and has no monopoly on uncertainty.

A more fundamental cause of uncertainty in mathematics is computational complexity. As any mathematician has only finite resources with which to conduct deductive reasoning, it is not usually possible for her to determine deductively whether a given proposition is a logical consequence of her other beliefs. However, she may consider other types of evidence in forming a degree of belief, as we will see in the next section.

The sums are too hard

Let us begin with two examples. Firstly, the infinite series

$$\pi = 4\left(1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots\right)$$

allows us to calculate π to an arbitrary number of decimal places, given sufficient computing time.[3] There is no uncertainty about where to start or how to get there. However, if I ask a mathematician what the 10^{99}th digit of π is, she will not know the answer. She will have neither an explicit nor an implicit belief about the matter, even though the value of this digit is a logical consequence of other beliefs she holds. Her uncertainty arises out of the sheer complexity of calculation. However, this same complexity need

[3] For actual computations of π, much more sophisticated convergents are used.

not prevent her from forming a degree of belief about whether the digit is a 3. She may know, for instance, that amongst all the known digits of π, each of the integers from 0 to 9 appears roughly equally and that many mathematicians believe that π is normal. She may consider that even if π is not normal, she has no reason to believe that 3 is any likelier or less likely than any other digit.

Secondly, the Goldbach conjecture states that every even integer greater than 2 can be written as the sum of two primes. The conjecture has never been proved. However, it has been checked by brute force computation for all $n \leq 10^{16}$ (Oliveira e Silva 2005). There are also heuristic arguments that we should expect any exceptions to the conjecture to be small. Thus, most mathematicians have a high degree of belief in the Goldbach conjecture that nonetheless falls short of 1.

Thus, we can see that a formal proof is not the only factor that may change a mathematician's degree of belief about a proposition. Some other possible factors are: partial proofs or proof sketches; proofs of similar or analogous results; numerical verification of a large number of cases; agreement with the results of a physical experiment; and many others.

Dutch bookies shouldn't be too clever

I now discuss a frequent objection to the idea of degrees of belief in mathematical statements. To make matters simpler, we first develop some notation. Let A represent a set of assumptions and $\mathcal{R}$ a set of rules of deduction. Further, let p and q represent propositions. We write $p \vdash q$ if q is derivable from $A \cup p$ by (finite) repeated applications of the elements of $\mathcal{R}$.

Some philosophers argue that any agent whose degrees of belief are not deductively consistent is irrational. For instance, they argue that if p is a tautology, then any rational agent must have $\mathcal{B}(p) = 1$. This claim is often supported by a so-called Dutch book argument. Such arguments assume that if $\mathcal{B}(p)$ is defined then the agent should be willing to pay anything less than \$$\mathcal{B}(p)$ for a bet that is worth \$1 if p is true and \$0 if p is false. She should also be willing to sell such a bet for anything more than \$$\mathcal{B}(p)$. Now, suppose that p is a tautology but $\mathcal{B}(p) < 1$. Then a Dutch bookie can buy from the agent for only \$$\mathcal{B}(p)$ a bet that is worth \$1 if p is true. But p is guaranteed to be true, and so the agent will make a sure loss.

If accepted in its general form, this argument has the consequence that for any mathematical proposition p, it is irrational for $\mathcal{B}(A \vdash p)$ to lie in the interior of $[0, 1]$. This is because p either is or is not derivable from A via the rules of deduction and so $A \vdash p$ is either a contradiction or a tautology. Thus, it is irrational to have intermediate degrees of belief about mathematical statements.

I will argue that Dutch book arguments in their unqualified form are unreasonable. Let us first of all ignore the many problems with the presentation of Dutch book arguments and focus on what they are trying to show.[4] The idea is that a bookie *who knows no more than the agent* is able to make a book against her because her degrees of belief do not conform to the probability calculus. This symmetry of knowledge is important. Suppose, for instance, that the agent believes there is a one in three chance that her favourite horse will win the Melbourne cup. She is presumably tempted by her local bookie's odds of 6 : 1 on the horse. The bookie, however, has poisoned the horse himself and knows that it will die on the starting blocks. The bookie can make a sure win out of the agent if she bets on the horse. However, this is not because the agent is irrational. The bookie simply *knows more* than she does.

Consider now another situation that is similar to the one above. A bookie offers a mathematician a bet at 20 : 1 odds that the 10^{99}th digit of π is a 3. The mathematician believes that no-one has yet calculated this digit but knows that her distributed computing project will arrive at the answer in a matter of weeks. Now, suppose that the bookie has been secretly running his own distributed computation through a trojan virus and he already knows the digit's value. As before, the bookie can make a sure win out of the agent if she accepts the bet. But, as before, he knows something the agent does not: namely, that the definition of π implies by deduction that the 10^{99}th digit of π is whatever he knows it to be. The knowledge asymmetry in this scenario is not so different to that in the previous one. In both scenarios, the bookie is taking advantage of information that is inaccessible to the agent. In the first case, it is empirical knowledge; in the second, it is deductive knowledge.

I have argued that intermediate degrees of belief about mathematical statements are not irrational. However, I have not disputed that a mathematician should try to conform her beliefs to the probability calculus where she can. For instance, if the mathematician *knows* that p is a tautology, then she should adopt $\mathcal{B}(p) = 1$. The Dutch book argument makes sense here because there is no asymmetry of knowledge. Thus, we might think of a mathematician as having two strategies in beating the Dutch bookie of life. The first is to aim for as much consistency in her mathematical beliefs as her

[4] In keeping with my earlier comments, I believe Dutch book arguments are best viewed as attempts to display, rather than define, irrationality.

state of knowledge allows. The second is to maximise her access to relevant knowledge.

Inductive reasoning in the deductive science

In this final section I develop a very simple consistency model for beliefs about a mathematical theory. Defending and developing the requirements is beyond the scope of this essay. For convenience, I will use the notation $A \vdash p \vee q = (A \vdash p) \vee (A \vdash q)$ and $A \vdash p \wedge q = (A \vdash p) \wedge (A \vdash q)$.[5]

Let $\mathcal{P}$ be the set of mathematical statements about which an agent has an explicit degree of belief. I propose that in order to be considered rational, the agent's degrees of belief should obey the following consistency requirements (cf. Gaifman 2004):

1. If $\alpha \in A$ then $\mathcal{B}(A \vdash \alpha) = 1$.

2. If $A \vdash p$ is in $\mathcal{P}$ then so is $\neg(A \vdash p)$ and $\mathcal{B}(\neg(A \vdash p)) = 1 - \mathcal{B}(A \vdash p)$.

3. If $A \vdash p$, $A \vdash q$, $A \vdash p \vee q$ and $A \vdash p \wedge q$ are all in $\mathcal{P}$, then

$$\mathcal{B}(A \vdash p \vee q) = \mathcal{B}(A \vdash p) + \mathcal{B}(A \vdash q) - \mathcal{B}(A \vdash p \wedge q).$$

4. If $p \vdash q$, $A \vdash p$ and $A \vdash q$ are in $\mathcal{P}$ then

$$\mathcal{B}(A \vdash p)\mathcal{B}(p \vdash q) \leq \mathcal{B}(A \vdash q).$$

Among other things, these requirements imply that if p and q are known to be equivalent in a mathematical theory, then $\mathcal{B}(A \vdash p) = \mathcal{B}(A \vdash q)$. They can also be used to prove a version of Bayes' theorem once conditional probability has been defined.

Conclusion

Inductive reasoning does and should have a place in the thoughts and decisions of mathematicians. Unlike in the physical sciences, where uncertainty about data and models plays a key role, mathematical uncertainty is due primarily to computational complexity. This complexity cuts mathematicians off from most deductive consequences of their beliefs, making inductive reasoning both rational and useful. In order to be considered rational, the

[5] Note that I am not assuming that anything in the mathematical theory corresponds to an AND or OR connective. The connectives are external to the system.

degrees of belief that result from such reasoning should satisfy weak requirements of consistency.

Bibliography

Corfield, David (2004). *Towards a Philosophy of Real Mathematics.* Cambridge: Cambridge University Press.

Erdös, Paul (1949). On a new method in elementary number theory which leads to an elementary proof of the prime number theorem. *Proceedings of the National Academy of Sciences*, Vol. 35 (pp. 374–384).

Eriksson, Lina and Hájek, Alan (2007). What Are Degrees of Belief? *Studia Logica*, Vol. 86, No. 2 (pp. 183–213).

Gaifman, Haim (2004). Reasoning with Limited Resources and Assigning Probabilities to Arithmetical Statements. *Synthese*, Vol. 140 (pp. 97–119).

Hájek, Alan (2005). Scotching Dutch Books? *Philosophical Perspectives*, Vol. 19 (pp. 139–151).

Lynch, Peter (2008). The origins of computer weather prediction and climate modeling. *Journal of Computational Physics*, Vol. 227, Issue 7 (pp. 3431–3444).

Oliveira e Silva, T. (2005). *Goldbach Conjecture Verification.* Available at `http://listserv.nodak.edu/cgi-bin/wa.exe?A2=ind0512&L=nmbrthry&T=0&P=3233`.

Peano, Guiseppe (1889). *Arithmetices principia, nova methodo exposita.* Turin: Fratres Bocca.

Selberg, Atle (1949). An elementary proof of the prime-number theorem. *Annals of Mathematics*, Vol. 50, No. 2 (pp. 305–313).

Schwitzgebel, Eric (2008). Belief. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition).* Available at `http://plato.stanford.edu/archives/fall2008/entries/belief/`.

Talbott, William (2008). Bayesian Epistemology. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition).* Available at `http://plato.stanford.edu/archives/fall2008/entries/epistemology-bayesian/`.