Abstracta SPECIAL ISSUE III, pp. 18 – 26, 2009

INFERENCE TO THE BEST INDUCTIVE PRACTICES

Paul Thagard

Harman and Kulkarni (2007) provide a rigorous and informative discussion of reliable reasoning, drawing philosophical conclusions from the elegant formal results of statistical learning theory. They have presented a strong case that statistical learning theory is highly relevant to issues in philosophy and psychology concerning inductive inferences. Although I agree with their general thrust, I want to take issue with some of the philosophical and psychological conclusions they reach. I will first discuss the general problem of assessing norms and propose a metanormative decision procedure that can then be applied to the assessment of inductive methods. In particular, I will apply it to the assessment of the method of inference that Harman and Kulkarni call transduction.

METANORMATIVITY

Philosophy is inherently normative in that it is concerned not just with how people *do* think and act, but also with how they *ought* to think and act. Normative issues arise in general epistemological deliberations about how people should make inferences and in general ethical deliberations about the morality of kinds of actions. But normative issues also abound in much more specific practices, for example concerning how scientists should use theories and experiments to investigate the world and how doctors should decide what medical treatments are best. By a *norm* I mean a practice that ought to be followed in a particular domain. In this sense, Harman and Kulkarni can be viewed as working to help establish norms for inductive inference. By *metanormativity* I mean the meta-level consideration of how norms ought to be established (Hardy-Valée and Thagard, 2008).

In contemporary philosophy, the most popular approach to metanormativity is a method that John Rawls (1971) called *reflective equilibrium,* which involves an ongoing process modifying general principles and particular judgments to better accord with each other. Harman and Kulkarni have legitimate worries about this method, particularly that it is often fragile and unreliable. I have expressed similar worries about reflective equilibrium myself (Thagard 1988, 2000), and Harman and Kulkarni erroneously extrapolate from my enthusiasm for coherence-based approaches to many kinds of inference that I endorse an approach to metanormativity akin to reflective equilibrium.

To clarify my position on how to arrive at reasonable norms about how people ought to think and act, I will now offer a decision procedure for metanormativity generalized from less explicit previous accounts (Thagard 1988, Hardy-Valée and Thagard 2008). Abstractly, the procedure requires the following steps:

1. Identify a domain of practices.
2. Identify candidate norms for these practices.
3. Identify the appropriate goals of the practices in the given domain.
4. Evaluate the extent to which different practices accomplish the relevant goals.
5. Adopt as domain norms those practices that best accomplish the relevant goals.

Of course, this kind of inference is revisable as new information comes in about goals and the efficacy of practices. I will illustrate how this procedure works by describing its application to a narrow medical domain.

In specific medical domains, physicians are concerned with establishing norms of diagnosis and treatment that prescribe the best practices for recognizing and improving patients' conditions. For example, in accord with step 1, we could identify as a narrow domain of practice the need of cardiologists to determine how to deal with patients who have arteries sufficiently blocked that chest pain results. Step 2 requires identifying candidate best practices, such as bypass surgery, stent insertion, treatment with drugs for controlling blood pressure and cholesterol levels, or watchful waiting. A major way of accomplishing step 2 is to survey the broad range of current practices, but a more creative approach requires considering new kinds of treatments that need to be invented or refined. Step 3 requires figuring out what are the goals of treating people with blocked arteries, which might include improving the patients' life expectancy and quality of life, but could also include matters of cost to the public health system in countries sufficiently civilized to have one. Step 4 requires the onerous process of determining as much as possible the extent to which each of the different processes accomplishes the various goals of treating arterial disease. Once this is done using the best available medical evidence, we should have a good idea of the best practices for treatment, which can then be adopted as norms in step 5.

This decision procedure can also be applied to the higher-order problem in step 4 of how to assess the efficacy of various treatments. A positive recent movement is *evidence-based medicine,* which advocates replacing unreflective judgments based on clinical experience with systematic evaluation of the best kinds of experimental evidence, particularly data gathered from randomized, controlled, clinical trials (e.g. Guyatt et al 1992). I won't attempt it here, but I would argue that the norms of evidence-based medicine should be adopted as best practices for establishing lower-level norms for particular kinds of treatment. Another interesting exercise would be to apply my decision procedure for metanormativity to itself, arguing that it is a better practice for adopting norms than such alternatives as reflective equilibrium and a priori reasoning. Instead, I want to apply my metanormativity decision procedure to the central problem of Harman and Kulkarni's book, establishing norms for inductive inference.

EVALUATING INDUCTIVE METHODS

The domain of practices now to be evaluated concerns inductive inferences, ones that introduce uncertainty (step 1). Despite the risk of making errors by inferring false conclusions, people have many ways of making inductive inferences, including: generalizing from samples to populations, using probability theory, statistical inference, the hypothetical-deductive method, inference to the best explanation, analogy, and wishful thinking. Step 2 of my decision procedure requires identifying these and many other possible ways of reasoning inductively in order to be able to assess which of them should be adopted as norms of inductive inference. Harman and Kulkarni are concerned with only a small range of kinds of inductive inference that can be analyzed using the tools of statistical learning theory.

Step 3 is much more problematic. Harman and Kulkarni interpret the philosophical problem of induction as the problem of the *reliability* of inductive inference. They do not define reliability, but I presume they mean something like the ratio of the number of true conclusions to the number of all conclusions reached (Goldman 1986). However, they seem to presuppose that reliability is the *only* goal of inductive inference, but I find this implausible.

First, in both science and everyday life, the goals of inductive inference include understanding as well as reliability. We want not only to achieve truths and to avoid error, but also to grasp why things happen. In everyday life, emphasis on reliability alone would restrict us to a kind of behaviorism, noting regularities in how people respond to their environments. But people cannot resist attributing mental states to each other, going beyond behavior to infer that people have various beliefs, desires, and emotions that cause their behavior. This ancient kind of reasoning was felicitously dubbed “inference to the best explanation” by Harman (1965). Inference to the best explanation can take us beyond observed regularity to non-observable states that tell us why regularities occur. It often leads to false conclusions, for we often err in our judgments about the mental states of other people, and even sometimes err about our own mental states. Still, inference to the best explanation about mental states should not be eschewed, because there is no better way of figuring out why people behave as they do.

More systematically, inference to the best explanation is an established part of the practice of inductive inference in science (Thagard 1988, 1992). In physics, chemistry, biology, and medicine, science has progressed by inferring the existence of entities that are not directly observable, such as electrons, molecules, genes, and viruses. This kind of inference is frequently *un*reliable, as we see from the pantheon of scientific mistakes that includes inferences to the existence of such discredited entities as phlogiston, caloric, and the luminiferous aether. Ultracautious positivist and empiricist philosophers of science want to stick closely to observation, but without the goal of understanding why things happen we would not have the best current theories. Hence we should include understanding beside reliability as a main goal of inductive inference.

Even more controversially, I would like to propose another goal to be used in evaluating competing methods of inductive inference: potential for practical importance. At any given moment, there is a huge range of inductive inferences that a person might make. I might devote the rest of the afternoon to collecting evidence that would support the generalization that all the items in my study weigh less than 100 kilograms. My inference would be reliable, but pointless. Inductive methods should be capable of producing conclusions that are useful, not just true. Once again, inference to the best explanation to non-observed entities wins out over more restrictive inductive methods that might be more reliable. Without theories about non-observable entities such as electromagnetic radiation and viruses, we would not have such marvels of modern technology as computers, television, and antiviral drugs. Hence I think that Harman and Kulkarni are unduly restrictive in considering only reliability as the concern of inductive reasoning.

Step 4 of my metanormativity decision procedure recommends assessing different methods with respect to all relevant goals. For inductive inference, the goals certainly do include reliability, and the assessment of inductive methods can legitimately be formal as well as empirical. Statistical learning theory strikes me as highly useful for assessing some inductive practices with respect to that particular goal, and I was impressed by the insights provided by Harman and Kulkarni concerning the VC dimension. My point is just that the goals and range of practices of inductive inferences are far broader than can be studied using these formal methods.

TRANSDUCTION

A particularly interesting part of the discussion by Harman and Kulkarni is their treatment of Vapnik's theory of transduction, which they mark as being novel in two main aspects. First, transduction does not involve inferring an inductive generalization that is then used for classification, but instead proceeds directly from information about previous cases to classification of new ones. Second, transduction uses information about what new cases have come up in its classification of them. This second aspect is indeed novel, but the first has a long history in philosophical and psychological discussions.

On the standard view, inductive inferences go from cases to rules, which can then be applied to new cases. The alternative view that inference can go directly from cases to cases was proposed by John Stuart Mill in the nineteenth century (Mill 1970, p. 364 – Ch. XX of Book III of the eighth edition):

> But we conclude (and that is all which the argument of analogy amounts to) that a fact *m,* known to be true of A, is more likely to be true of B if B agrees with A in some of its properties, (even though no connection is known to exist between *m* and those properties) than if no resemblances at all could be traced between B and any other thing known to possess the attribute *m.*

Similarly, Bertrand Russell (1967, p. 44) wrote early in the twentieth century:

> But the newness of the knowledge is much less certain if we take the stock instance of deduction that is always given in books on logic, namely, 'All men are mortal; Socrates is a man, therefore Socrates is mortal.' In this case, what we really know beyond reasonable doubt is that certain men, A, B, C, were mortal, since, in fact, they have died. If Socrates is one of these men, it is foolish to go the roundabout way through 'all men are mortal' to arrive at the conclusion that *probably* Socrates is mortal. If Socrates is not one of the men on whom our induction is based, we shall still do better to argue straight from our A, B, C, to Socrates, than to go round by the general proposition, 'all men are mortal'. For the probability that Socrates is mortal is greater, on our data, than the probability that all men are mortal. (This is obvious, because if all men are mortal, so is Socrates; but if Socrates is mortal, it does not follow that all men are mortal.) Hence we shall reach the conclusion that Socrates is mortal with a greater approach to certainty if we make our argument purely inductive than if we go by way of 'all men are mortal' and then use deduction.

Thus something like transduction has long been recognized by philosophers.

Similarly, contrary to what Harman and Kulkarni suggest, the importance of inference from cases to cases has also been noted by many psychologists, from at least three different perspectives. First, there is a large body of psychological research on analogical inference, which is obviously inference from cases to cases without intervening generalizations, although its cognitive structure is more complex than Mill recognized (e.g. Holyoak and Thagard, 1995). Second, one prominent theory of concepts is called the *exemplar* view, which proposes that concepts are not stored as general representations but merely as collections of particular cases that are then used to classify new cases (e.g.

Murphy 2002, ch. 4). Third, the standard interpretation of neural networks that learn from examples using backpropagation is not that the connection weights encode rules, but that they encode statistical patterns rather than generalizations (e.g. Rumelhart and McClelland 1986). Thus the aspect of transductive inference that it goes from cases to cases without intervening rules has a strong place in psychology as well as philosophy.

Should people use transduction? It would seem to support the inductive goal of reliability, because going from cases to cases avoids the danger of overgeneralization that inferring rules can easily introduce, as the quote from Russell suggests. However, inductive generalization may have some benefits with respect to explanation, if it leads to the adoption of causal rules that can explain why things happen. Another currently prominent theory of concepts emphasizes their role not just in classification but in explanation (Murphy 2002, ch. 6). I may, for example, have seen many cases of drunks that I can use to classify a new staggering person as a drunk, but an inductive generalization may enable me to explain why he is drunk. For example, the generalization that people who drink a lot of alcohol lose motor control provides a causal explanation of why someone is staggering.

My third goal of inductive inference was practical importance, and it is here that the second aspect of transduction seems most relevant. Because transduction takes into account information about new information to be classified, it can potentially come up with more useful new classifications. I cannot think of any current psychological theory of concept learning that has this property, nor of any experimental evidence that people have the capability of using such anticipations in their concept learning. However, a notable aspect of the multiconstraint theory of analogical inference proposes that the retrieval, mapping, and application of cases to be used as analogies all involve the constraint of purpose, which concerns the practical use of the analogy (Holyoak and Thagard 1995). So perhaps analogical inference counts as transductive both in the sense of going from cases to cases and in the sense of taking into account the characteristics of cases to be inferred about. Purpose may not increase reliability in the abstract sense of improving the truth to error ratio, but should make case-to-case inferences more useful for the various problem-solving goals of analogical inferences.

In conclusion, Harman and Kulkarni's *Reliable Reasoning* is a highly informative and stimulating exploration of important topics in inductive inference, but many important descriptive and normative questions about induction inference require further investigation.

Paul Thagard
University of Waterloo

pthagard@uwaterloo.ca

References

Goldman, A. (1986). *Epistemology and cognition*. Cambridge, MA: Harvard University Press.

Guyatt, G. , et al. (1992). 'Evidence-based medicine: A new approach to teaching the practice of medicine'. *Journal of the American Medical Association, 268*, 2420-2425.

Hardy-Vallée, B., & Thagard, P. (2008). 'How to play the ultimatum game: An engineering approach to metanormativity'. *Philosophical Psychology, 21*, 173-192.

Harman, G. (1965). 'The inference to the best explanation'. *Philosophical Review, 74*, 88-95.

Harman, G., & Kulkarni, S. (2007). *Reliable reasoning: Induction and statistical learning theory*. Cambridge, MA: MIT Press.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press/Bradford Books.

Mill, J. S. (1970). *A system of logic* (8 ed.). London: Longman.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Rawls, J. (1971). *A theory of justice*. Cambridge, MA: Harvard University Press.

Rumelhart, D. E., & McClelland, J. L. (Eds.). (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge MA: MIT Press/Bradford Books.

Russell, B. (1967). *The problems of philosophy*. Oxford: Oxford University Press.

Thagard, P. (1988). *Computational philosophy of science*. Cambridge, MA: MIT Press/Bradford Books.

______. (1992). *Conceptual revolutions*. Princeton: Princeton University Press.

______. (2000). *Coherence in thought and action*. Cambridge, MA: MIT Press.