The problem of grue

Jeff Speaks

April 3, 2008

1 The classical problem of induction and the new riddle of induction

Hume raised a famous problem about inductive reasoning, which can be thought of as an argument that we can provide no non-circular justification for inductive reasoning. Goodman's problem is different: he is not asking how we can justify induction, but rather is asking what sorts of inductive practices are legitimate.

There is no obvious answer to this question. As Goodman says,

> "As principles of *de*ductive inference, we have the familiar and highly developed laws of logic; but there are available no such precisely stated and well-recognized principles of inductive inference." (65)

The task of giving canons of inductive inference is the task of explaining when a certain set of premises can provide a good inductive argument for a given conclusion. It looks like a good starting point would be examples of *enumerate induction* of the following sort:

1. Emerald$_1$ is green.
2. Emerald$_2$ is green.

.

.

1000. Emerald$_{1000}$ is green.

C. All emeralds are green.

It seems clear that inductive arguments of this form are often good arguments. And what makes them good seems, intuitively, to be that the conclusions of arguments of this form are generalizations, and the premises are instances of that generalization. It seems like this is a good first step in putting together a logic of induction: a generalization is confirmed by its instances.

2 Grue and enumerative induction

Goodman's new riddle of induction shows that this is a false step: not all generalizations are confirmed by their instances. He shows this by inventing the predicate 'grue.' It is defined as follows:

> An object is grue if and only if the object is either (1) green, and has been observed before now, or (2), blue, and has not been observed before now.

This is a perfectly fine definition, in the sense that it gives us clear conditions on when the word 'grue' applies to an object. But it poses a problem when we use it in inductive arguments. Consider, for example, the following argument:

1. Emerald$_1$ is grue.
2. Emerald$_2$ is grue.

.

.

1000. Emerald$_{1000}$ is grue.

C. All emeralds are grue.

This argument seems, by the standard suggested above, to be a perfectly good inductive argument. But it cannot be, since it does not give us good reason to believe that all emeralds which have not been observed till now are blue.

Another way to see the problem is that the example of 'grue' seems to show that exactly the same evidence — observation of 1000 green/grue emeralds — provides equally good evidence for believing both that the next emerald to be observed will be green, and that it will be blue. But this is absurd.

3 What's wrong with 'grue'?

It is natural to respond to this puzzle by claiming that something must be wrong with the word 'grue.' If we could show that there was something wrong with it, then we could restrict the canons of induction to apply only to inductive arguments which do not contain terms which are defective in this way.

3.1 No made-up words

One might object that 'Grue' is a made-up word.

This is true; but what are new scientific terms? Aren't they just made-up words that are wholly or partially defined in terms of existing vocabulary?

3.2 Definability constraints

A second intuitive thought is that 'grue' is somehow unnatural, because it is defined in terms of two other predicates, 'green' and 'blue.' But, as Goodman points out, things are not so simple. Consider the new predicate, 'bleen', defined as follows:

> An object is bleen if and only if the object is either (1) blue, and has been observed before now, or (2), green, and has not been observed before now.

Again, this seems like a perfectly comprehensible, if unusual, definition. The problem is that we can now see that 'green' is also definable in terms of 'grue' and 'bleen': something is green if and only if it is either (1) grue and has been examined before now, or (2) bleen and has not been examined before now.

3.3 Reference to time and place

A further thought is that 'grue' is illegitimate because it makes reference to a specific time; it is defined in terms of what color something is if observed before *now*. This is part of what makes the predicate seem so artificial, so it is natural to think that it is also part of what makes its use in inductive arguments illegitimate. So maybe we should restrict the terms involved in inductive arguments to ones which do not involve any reference to a specific time and place.

This faces at least three problems:

- In what sense does 'grue' involve reference to a time and place? It can be defined partly in terms of time and place; but that's true of every term.
- We don't want to say that good inductive arguments can't include reference to time or place; sometimes we give inductive arguments about events in a certain time frame.
- We could come up with a 'grue'-like predicate which made no mention of time or place.

3.4 Unnatural properties

A third response to Goodman's problem is to appeal not to the way in which 'grue' is defined, but to differences between the properties of being grue and being green. The idea that inductive inferences are only reliable if they are restricted to properties which are in some sense *natural properties*.

What is the difference between natural and non-natural properties? How could we tell the difference between them?

4 Enumerative induction and background conditions

It is not easy to find something wrong with 'grue' which yields intuitively plausible results about which inductive arguments are good arguments, and which are not.

A different response to the paradox, which Sainsbury favors, is to reject the principle which was at work in generating the 'grue' paradox:

> Every instance of a generalization confirms that generalization.

This does solve the problem — and with it also the paradox of the ravens. However, this by itself is not very satisfactory. We want to say that the reasoning characteristic of science can sometimes give us good reasons for belief — so we should be able to say what sort of inductive reasoning can do this, if enumerate induction sometimes fails. This, after all, was the point of Goodman's challenge. So can we come up with some principle that tells us when an instance of a universal generalization *does* confirm that generalization?

Sainsbury considers a few options. One can be put as follows:

> A generalization that all A's are B's is confirmed by instances unless we have good reason to believe that there is some property, O, such that every A-instance is O, and if those A-instances had not been O, they would not have been B.

How does this help with the example of the grue emeralds?

This principle makes confirmation partly a matter of what background beliefs one brings to bear on the situation. Compare Sainsbury's two versions of the example of the lobsters (p. 87). Does it make sense to say that whether some piece of evidence counts in favor of a theory depends on what your other beliefs are?

Does this help at all with the paradox of the ravens?