Lecture 4
Induction

1. Introduction

The last source of knowledge that we will consider is **induction**. Induction is a non-deductive form of inference; that is to say, the premises of an inductive argument are consistent with the falsity of that argument's conclusion. For the purposes of this lecture, we will say that a belief formed via induction is an **inductive belief**, and if that belief constitutes knowledge then it is **inductive knowledge**. We will also say that beliefs justified by inductive are **inductively justified**.

Perhaps the most famous objection to inductive inference comes from Hume. We'll briefly rehearse Hume's arguments against induction before turning our attention to more contemporary discussions of this method of inference. These include Goodman's "new riddle of induction" and Lipton's "inference to the best explanation".

2. Hume & The Old Problem of Induction

But we begin, as we so often do, with Hume. To make sense of his arguments against induction, we must recall Hume's general epistemology. Earlier in the set of lectures we considered different propositions as either being **matters of fact** or **relations of ideas**. Hume identifies two kinds of reasoning that correspond roughly to this distinction:

(1) **Demonstrative reasoning** – "that concerning relations of ideas" (IV.2.18)
(2) **Reasoning from experience** – "that concerning matters of fact" (IV.2.18)

Now, consider an example of an inductive argument:

P1. Every cat I have observed has whiskers.
C1. All cats have whiskers.

First, note that **P1** is a matter of fact. It is something I know on the basis of observation (supposing that it is true that every cat that I've seen had whiskers). **C1**, on the other hand, is not something I could come to know by observing *every* instance of a cat; but it is something that I come to know via experience in some other way. After all, **C1** is not a relation of ideas; there is nothing in the idea of a cat that makes '*x* is a cat and *x* is whiskerless' entail a contradiction. So **C1**, if it constitutes knowledge, is a matter of fact.

By what argument are we entitled to conclude **C1** from **P1**, though? Hume notes that this argument is not a **demonstrative** on (i.e. it is not **deductive**). **P1** and **¬C1** are consistent—they can both be true together.

So the argument must involve some kind of reasoning from experience. But what could this method of reasoning be? And is it rightly called '*reasoning*'?

Hume notes that "all arguments from experience are founded on the similarity, which we discover among natural objects" (IV.2.20). But what reason have we for thinking that similarities we have observed in the past will persist in the future? The mere fact of having observed whiskered cats does not *entail* that all future cats will be whiskered. It is certainly true that I will **expect** them to be so. But, this expectation is not informed or legitimated by any valid argument.

Now, we could add a premise to make the argument valid:

P1. Every cat I have observed has whiskers.
P2. All cats are like the cats I have observed.
C1. All cats have whiskers.

But **P2** is unsubstantiated. So its addition is ad hoc. Of course, it is no use saying that experience provides evidence for **P2**, since doing this would be begging the question. This would effectively be to provide an inductive argument for induction!

Ultimately, Hume argues that it is **habit**, not reason or argument, that leads us to infer **C1** from **P1**.

3. Response to the Old Problem

There are several different kinds of response to Hume's argument. For now, let's just briefly consider one: **probable conclusions**. Part of Hume's challenge was that the conclusions we draw based on experience (like **C1**) do not follow deductively from their premises. This is certainly true when the conclusion is stated as **C1** is; but we could state the conclusion in less absolute terms:

C1*. It is highly likely that all cats have whiskers.

From here, the strategy involves providing an *a priori* argument for the claim that the premises of an inductive argument make their conclusion more probable. If this can be done, then inductive inference can be proven to be a legitimate form of inference by "demonstrative argument". But let's leave this for now.

4. Goodman & The New Problem of Induction

Let's suppose for the sake of argument that one of the responses to Hume' argument is successful. In other words, let's suppose that the method of inductive inference is a legitimate means of acquiring knowledge. Goodman's argument goes to show that, even if this were the case, induction would still face a challenge with respect knowledge-acquisition.

The general worry concerns what kinds of conclusions we can inductively draw from a given set of observations. Consider the **Paradox of the Raven** from Hempel (1945). We take it that seeing instances of black ravens provides evidence for the proposition (R) 'all ravens are black'. But (R) is equivalent to:

(R-cond) $\forall x(Rx \rightarrow Bx)$
(R-contra) $\forall x(\neg Bx \rightarrow \neg Rx)$

Now, on the plausible assumption that for any hypotheses H and H*, if H and H* are equivalent then E is evidence for H iff E is evidence for H*, it follows that any evidence for (R-contra) is evidence for (R). But that is to say that the whiteboard is evidence that all ravens are black. And intuitively it seems wrong to allow the following inductive inference:

P2. All non-black things I have observed are non-ravens.
C2. All non-black things are non-ravens.
C3. All ravens are black.

One way to understand what's wrong with the above argument is to say that non-blackness and non-ravenhood are not **projectable properties**. They are not the sort of properties about which we can form inductively justified beliefs. But how should we distinguish projectable properties from unprojectable properties? You might think that negative properties like 'non-black' should be disqualified, in which case the projectable properties are all and only the positive properties. But, as Goodman's "grue paradox" shows, this won't work either.

The Grue Paradox

Suppose you've seen lots of emeralds in the past. Each time, you observe an emerald, you observe that it's green. So you form a belief of the following sort:

P3. All of the emeralds I have observed are green.

And from this, using inductive inference, you conclude:

C4. All emeralds are green.

But now let's define a new predicate called **grue** in the following way:

Grue: For all x, x is grue iff x is observed before *t* and is green, or x not observed before *t* and is blue.
(Where *t* is some arbitrary time in the distant future)

For the sake of illustration, let's suppose that *t* is 1 January 3019. That means that all of the instances when I observed emeralds are times before *t*, and I observed them to be green. But being observed to be green before time *t* suffices for being grue. So now it is also the case that:

P4. All of the emeralds I have observed are grue.

From which inductive inference equally well licenses the conclusion:

C5. All emeralds are grue.

But now what should I say about emeralds? Are they green or are they grue? Induction itself does not tell us which of **C4** or **C5** to believe; both are equally well inductively justified.

There are a few things to notice. First, we could generate infinitely many grue-like predicates for which we have equally good inductive evidence. Simply change *t* by some non-zero increment each time and you'll have a new predicate. Second, *t* could just as well be in the distant past as in the distant future. Suppose *t** is 50 million years ago, we could define a predicate like the following:

Grue*: For all x, x is grue iff x is unobserved before *t** and is blue, or x is observed after *t** and is green.

You might think you could again respond by arguing that grue is not a projectable property. For instance, perhaps we should rule out **disjunctive** predicates on the grounds that they seem to be gerrymandered. The trouble with this proposal is that we can define green in disjunctive terms too. Let's define the predicate **bleen** as follows:

Bleen: For all x, x is bleen iff x is observed before *t* and is blue or is not observed before *t* and is green.

We can now define green in terms of grue and bleen:

Green: For all x, x is green iff x is observed before *t* and is grue or x is not observed before *t* and is bleen.

What is worse, we can perform this trick for any predicate you like. There was nothing special about green and blue. Suppose we defined predicates 'emermond' and 'diarald' the way we defined grue and bleen. We can define emerald thus:

Emerald: For all x, x is an emerald iff is observed before *t* and is an emermond or x is not observed before *t* and is a diarald.

This poses a serious threat to our inductive knowledge. Goodman's argument shows that I have equally good inductive justification for a whole slew of different beliefs about emeralds, and indeed about anything else I have observed. So how can I claim to know that emeralds are green? Any claim to know one of the propositions about emeralds and not the others would not be inductively justified. The choice would be ad hoc, and so an illegitimate claim to knowledge.

Another possible solution is to deny the principle that underlies the inductive inference that generates the problem. Call this the principle of **enumerative induction**:

Enumerative Induction: Every instance of a generalisation confirms that generalisation.

Instead you could say that instances confirm a given generalisation against a background of other beliefs. For instance, in this case you might say that the property 'grue' is not "well-entrenched" in our body of beliefs (this is Goodman's solution). Or you might say that we have a background belief that the intrinsic properties of emeralds do not depend on our observation, but this is inconsistent with thinking that the emeralds you observed are grue, since they are only grue *because you observed them* (Sainsbury 2009: 101).

These thoughts—that there is more to induction than mere enumerative induction—suggest a slightly different view of *ampliative* inferences (i.e. inferences that take us 'beyond' what is observed) like these may be called for.

5. Inference to the Best Explanation

Inference to the best explanation (IBE) is a form of ampliative inference, and one that may provide an account of induction that accounts for the the non-inductive factors that are relevant to a given inductive inference. One of the solutions to the grue paradox considered above was that we should include considerations that are not merely enumerative. IBE accommodates such considerations.

For a statement of IBE, I'll draw on Lipton (2004)—in particular, **Ch.4** of that book. The aim of IBE accounts is to give "an illuminating model of our inductive practices" (57). Roughly, we can state IBE as follows:

> **IBE:** Given a set of evidence, $\{E_1…E_n\}$, and distinct (potential) explanations of E, $H_1…H_n$, we have most justification for believing the member of $\{H_1…H_n\}$ that *best explains* $\{E_1…E_n\}$.

Notice that IBE is not enumerative (or at least, need not be). We are not looking for the H that has the most instances in the set $\{E_1…E_n\}$. We are looking for the H that best **explains** $\{E_1…E_n\}$, where the strength of an explanation is not reducible to the number of confirming instances.

So, in response to the grue paradox, you might say "that the emeralds are green best explains my observations"—i.e. better explains the observations than that the emeralds are grue. But, of course, this solution will only be satisfying if we can give a good account of how to adjudicate between candidate explanations. What makes one explanation better than the rest?

Lipton identifies two competing considerations: **likeliness** and **loveliness**.

Likeliness concerns the likelihood of truth, and is plausibly measured in enumerative terms (e.g. if there are more instances of green emeralds than of red emeralds, then 'all emeralds are green' is more likely than 'all emeralds are red').

Loveliness concerns how enlightening an explanation is, i.e. how much it helps us understand the evidence. This is much harder to measure.

According to Lipton, considerations of loveliness are "an important guide to judgements of likeliness". Here's his motivation for this view:

> *[…]we have an attempt to account for epistemic value in terms of explanatory virtue. This [view] claims that the explanation that would, if true, provide the deepest understanding is the explanation that is likeliest to be true. Such an account suggests a really lovely explanation of our inferential practice itself, one that links the search for truth and the search for understanding in a fundamental way.* (61)

Lipton argues that we've been misled into thinking that we're only interested in the likeliest explanations for evidence because, the loveliness and likeliness "tend to go together" (61).

IBE, then, can explain why it is that we discriminate on seemingly arbitrary grounds between predicates like 'green' and predicates like 'grue'. The explanation of our observations in terms of 'green' confers a better understanding of that evidence than the one in terms of 'grue'.

Notice, though, that this response doesn't so much try to remove the subjective considerations from inductive inference in order to legitimise such inferences as sources of knowledge. It rather leans into the subjectivity! Judging what counts as a good explanation is now part of induction on this view. Lipton tries to bridge the 'gap' to truth by saying that the subjective considerations are a good guide to truth. If you accept Lipton's argument, then you could say that we are justified in believing something like **C1** on the basis of IBE.

C1. All cats have whiskers.

However, if you don't, then you could still accept that we are justified in believing something like **C1'**:

C1'. 'All cats have whiskers' best explains my evidence.