Nelson Goodman — *Problems for a Theory of Confirmation*

• Recall the two questions we could ask about induction:
  
  Q1. Which inductions are good, and which are bad?
  
  Q2. Why are the good inductions good?

• Hume raised a worry about our ability to answer Q2. He argued that any answer we gave
to this question would be circular. Therefore, we cannot provide a good reason for thinking
that good inductions are good—ie, that they are likely to lead to true beliefs.

• Goodman raises problems for our ability to answer Q1. That is, he raises concerns about our
ability to provide a formal and systematic account of when a given body of evidence *supports*
or *confirms* a given hypothesis.

• At the time that Goodman was writing, philosophers of science were attempting to develop
a *logic* of induction, comparable to the logic of deduction. That is, they were attempting to
develop a logic that would tell them, in purely *formal* terms, whether a given piece of evidence
provided *inductive* support of a given hypothesis.

  – A first-pass thought: induction is in some sense the inverse of deduction. If a hypoth-
    ensis $h$ (along, perhaps, with some auxiliary assumptions) *deductively* entails that we will
    observe evidence $e$, then observing the evidence $e$ *inductively* supports the hypothesis $h$.

    \[
    \frac{h}{e} \quad \Rightarrow \quad \frac{e}{h}
    \]

  – Goodman: if we accept this, then if we also accept the principle that, if $e$ confirms $h$,
    then $e$ confirms whatever $h$ deductively entails, then we get the absurd conclusion that
    everything confirms everything else. That’s because, for any $A$ and any $B$, $A \land B$ deduc-
    tively entails $A$ and deductively entails $B$. So, if induction is the inverse of deduction,
    then $A$ should confirm $A \land B$, and if confirmation transmits through deductive entail-
    ment, then $A$ should confirm $B$, since $A \land B$ deductively entails $B$. So every proposition
    $A$ confirms every other proposition $B$.

    \[
    \frac{A \land B}{A} \quad \Rightarrow \quad \frac{A}{A \land B}
    \]

    \[
    \downarrow
    \]

    \[
    \frac{A \land B}{B} \quad \Rightarrow \quad \frac{A}{B}
    \]

    But that’s absurd. So the first-pass-thought can’t be right.
Goodman’s conclusion: not every entailment of a hypothesis confirms it. However, perhaps every instance of a general hypothesis confirms it. An instance of a general hypothesis \((\forall x)(Fx \rightarrow Gx)\) is an individual \(F\) that is also \(G\). So, if I say that all copper conducts electricity, then an individual piece of copper that conducts electricity will confirm this hypothesis.

Hempel’s Paradox of the Ravens

- Hempel presents the following problem for this analysis. If we accept these two principles:
  
  1. A universal hypothesis \((\forall x)(Fx \rightarrow Gx)\) is confirmed by an \(F\) which is \(G\).
  2. If \(h^*\) is logically equivalent to \(h\), then if \(e\) confirms \(h\), then it also confirms \(h^*\).

  then we are led to the absurd conclusion that observing a white shoe in our closet confirms the hypothesis that All ravens are black.

  - Here’s why: Consider the hypothesis \((\forall x)(\neg Bx \rightarrow \neg Rx)\) (where \(Bx\) is \(x\) is black and \(Rx\) is \(x\) is a raven)—i.e., ‘All non-black things are non-ravens’. By principle 1, this hypothesis is confirmed by a white shoe that I find in my closet. However, \((\forall x)(\neg Bx \rightarrow \neg Rx)\) is logically equivalent to \((\forall x)(Rx \rightarrow Bx)\) (‘All ravens are black’). So, by principle 2, a white shoe in my closet confirms the hypothesis that all ravens are black.

  - Goodman’s response: in the absence of any background information, just considering the white shoe in isolation, it does confirm the hypothesis that all ravens are black, but so too does it confirm the hypothesis all black things are ravens, and that everything is a non-black non-raven (i.e., there are no black things and there are no ravens). However, we know that these generalizations are false because they do not fit in with our total evidence.

Goodman’s ‘New Riddle of Induction’

- A problem for any purely formal theory of confirmation (one which does not take into account the meaning of \(F\) and \(G\) in a universal claim \((\forall x)(Fx \rightarrow Gx)\)): only universal claims with lawlike predicates \(F\) and \(G\) is capable of being confirmed by its instances.

  - For instance, the observation of a third son in a room does not confirm the hypothesis that everyone in this room is a third son.

  - The problem runs deeper: consider the property of being grue. A thing is grue just in case it is green and first observed before 2050 or else blue and first observed after 2050. Now, just as ever observed emerald has been green, so too has every observed emerald been grue.

  - So, if we just look at the formal structure of an inductive argument, we will think that both of the following arguments are equally good:
Emerald 1 is green
Emerald 2 is green

\[ \vdots \]
Emerald \( N \) is green

The first emerald observed after 2050 will be green

Emerald 1 is grue
Emerald 2 is grue

\[ \vdots \]
Emerald \( N \) is grue

The first emerald observed after 2050 will be grue (that is, blue)

- So our theory can’t get by on form alone. We need to pay attention to content as well.

- Additionally, we need some way of ruling these ‘gruesome’ properties out to settle upon the \textit{lawlike} ones; since only the \textit{lawlike} ones can be projected from observed instances onto unobserved instances.

- A first-pass thought: can’t we require that only \textit{qualitative} properties be used? ‘Grue’ doesn’t count as a qualitative property, since it involves reference to a specific time.

  - Goodman: yes, if we start with the terms ‘green’ and ‘blue’, then ‘grue’ involves reference to a certain time. However, if we start with ‘grue’ and ‘bleen’—where something is bleen iff it is first observed before 2050 and blue or first observed after 2050 and green—then we could define ‘green’ and ‘blue’ in such a way that their definitions make reference to a specific time.

    \[
    \text{Green} =_{df} \text{grue and first observed before 2050 or bleen and first observed after 2050.}
    \]

    \[
    \text{Blue} =_{df} \text{bleen and first observed before 2050 or grue and first observed after 2050.}
    \]

- The ‘New Riddle’ of Induction: what distinguishes those regularities which we \textit{can} project from observed cases to unobserved cases (like, \textit{e.g.}, ‘all emeralds are green’) from those which we \textit{cannot} project from observed cases to unobserved cases (like, \textit{e.g.}, ‘all emeralds are grue’)?