

Igor Douven 2018 “Inference to the Best Explanation: What Is It? And Why Should We Care?”

IBE...

What is it?

IBE is best understood as “a slogan that not only permits different fleshings-out, but that may even rationally require different fleshings-out, depending on what our interests are.” (23)

And why should we care?

We (Bayesians) should care because “there exist versions of IBE that, at least in certain kinds of updating contexts, do better than Bayes’ rule in terms of speed and accuracy”. (18)

1. WHAT IS INFERENCE TO THE BEST EXPLANATION?

IBE: “hypotheses are to be assessed at least partly on the basis of their explanatory virtues.” (7)

Bas van Fraassen’s formulation of **IBE as a probabilistic update rule** was initially designed to illustrate vulnerability to Dutch book arguments, but it can be repurposed here for its formal virtues:

$$\Pr[E](H_i) = \frac{\Pr(H_i)\Pr(E|H_i) + f(H_i, E)}{\sum_{j=1}^n (\Pr(H_j)\Pr(E|H_j) + f(H_j, E))}$$

“the function f assigns a **bonus point** to the hypothesis that explains E best (given the background knowledge) and nothing to the other hypotheses.” (10), leaving Bayes’ rule as a limiting case.

2. IBE IN THE PSYCHOLOGY OF REASONING

Since “there is much experimental support for the claim that explanation influences the **attribution of prior probabilities**” (12),

it would be great to know in our context whether “explanatory goodness is also a factor that influences the **change of those probabilities** when new evidence is obtained is a different question”. (13)

“participants were first shown two urns—called ‘urn A’ and ‘urn B’—containing the same number of balls, some of which were black, the others white, but with different ratios of black and white balls. After the participants were told about the contents of the urns, one urn was chosen on the basis of a flip with a fair coin, and from that urn, ten balls were drawn, without replacement. The participants could not see which urn had been selected and were asked, after each drawing of a ball, how likely it was, in their opinion, that urn A had been selected, and they were also asked to rate the **explanatory goodness both of the hypothesis that urn A had been selected (H_A) and of the hypothesis that urn B had been selected (H_B)**.” (13)

Now, could such subjective judgments be replaced with **objective degrees of explanatory goodness** and still outperform objective probabilities alone?

$$\text{Popper: } \frac{\Pr(E|H) - \Pr(E)}{\Pr(E|H) + \Pr(E)}, \quad \text{Good: } \ln\left(\frac{\Pr(E|H)}{\Pr(E)}\right),$$

Indeed, applying these measures improved on Bayes' rule, so let's have that as an inspiration.

3. FLESHING OUT THE SLOGAN

We want our probabilistic update rule to be fast & accurate:

Fast: making us believe the truth to at least a high degree (more than 0.9)

Accurate: in the course of making us believe the truth to a high degree, it does not, along the way, make us believe falsehoods to a high degree (scoring rules: Brier / log)

Less asymmetric: let's spread out van Fraassen' bonus point, such that "explanatory bestness can be stated straightforwardly in terms of whichever measure of explanatory goodness the rule uses". (17)

EXPL:

$$\Pr[E](H_i) = \frac{\Pr(H_i)\Pr(E|H_i) + c\Pr(H_i)\Pr(E|H_i)\mathcal{M}(H_i, E)}{\sum_{j=1}^n (\Pr(H_j)\Pr(E|H_j) + c\Pr(H_j)\Pr(E|H_j)\mathcal{M}(H_j, E))}$$

constant / measure



"the rules to be considered will first update the hypothesis' probability via Bayes' rule,

then calculate the hypothesis' explanatory goodness according to a given objective measure,

next add or subtract a percentage of the hypothesis' probability in proportion to its explanatory goodness,

and finally normalize." (17)

Roulette Tables	Widget Machines	Mozart's Cause of Death
very slight bias → Good's rule	"bias" is going to be close to 1 → Bayes' or Popper's rule	accuracy may not be a concern → EXPL

❖ **Green Gem:** “I offer an analogy: psychologists studying problem solving have identified a number of general strategies that people use to solve problems. Among them is a strategy often called “divide and conquer,” which comes down to dividing up a problem into more manageable sub-problems, solving the sub-problems, and obtaining a solution to the initial problem by collating the solutions to the sub-problems.” (11)

❖ **Blue Gem:** “We see that Bayes’ rule appears as a **crystal clear**, simple mathematical formula, which is in stark contrast with the formulations of IBE given previously.” → What exactly is crystal clear? (9)

❖ **Grue Gem:** “because IBE might be descriptively correct—people might, as a matter of fact, respond to the receipt of new evidence as prescribed by (some version of) IBE—even if they are wrong to do so. That would be a valid reason to study IBE (**even if perhaps more for psychologists than for philosophers**).” (7) → Who is the psychologist here?!