THE RAVENS: WHY THE 'CANONICAL' BAYESIAN RESPONSE FAILS

JAKE CHANDLER*

Abstract

A common Bayesian response to the Paradox of the Ravens involves the claim that whilst (i) strictly-speaking, it *is true* that observing a non-black non-raven confirms the hypothesis that all ravens are black, (ii) we typically *judge it to be false*, due to the fact that the degree of support involved is far smaller than the degree of support conferred on the same hypothesis by the observation of a black raven. In this article, it is pointed out that this line of reply fails to address the issue at stake.

> 'C.G. Hempel has stated the following paradox. The sentence *a*: 'This is a man, and is mortal' confirms the general proposition *b*: 'Every man is mortal' Moreover, the sentence *a1*: 'This chair is not mortal and is not a man', confirms the general proposition *b1*: 'No non-mortal being is a man'. Now *b1*is equivalent to *b*. Thus *a1*, confirming *b1*, should at the same time confirm *b*. But this sounds paradoxical.' [Hosiasson-Lindenbaum, 1940, 136]

The above quote is the first formulation in print of Hempel's famous 'Ravens Paradox', as it is now known.[1] Swapping the predicates in the above quote for their more familiar contemporary counterparts, the paradox is generated by first of all noting that the two following assertions seem to be true:

Plausible Premise (PP): 'Observing that a is non-black and is a non-raven evidentially supports the hypothesis that all non-black things are non-ravens, relative to what we know.'

Equivalence Condition (EC): 'If E evidentially supports H_1, relative to background knowledge K, and H_1 is equivalent to H_2, then E evidentially supports H_2, relative to K, for all H_1, H_2, E, and K.'

Commitment to these assertions however, would appear to entail commitment to the following assertion, which is widely agreed to be counter-intuitive:

*Centre for Logic and Analytic Philosophy, HIW, KU Leuven, Kardinaal Mercierplein 2, 3000 Leuven, Belgium

Problematic Conclusion (PC): 'Observing that a is a non-raven and is non-black evidentially supports the hypothesis that all ravens are black, relative to what we know.'

The Ravens Paradox was originally raised as a problem for Hempel's instantiational account of confirmation, but the problem is also relevant to the traditional Bayesian account. The traditional Bayesian qualitative account of evidential support (TB) takes assertions of the form 'E evidentially supports H, given background knowledge K' to affirm the existence of a three-place probabilistic relation between E, H and K:

TB: Assertions of the form 'E evidentially supports H, given background knowledge K' are true iff $Pr(H \mid E\&K) > Pr(H \mid K)$.[2]

In other words, according to TB, assertions of the type 'E evidentially supports H, given background knowledge K' are true iff the probability of H conditional upon $E\&K$ is strictly superior to the probability of H conditional upon Kalone. Using the obvious abbreviations for the predicates, and assuming that the this gives us the following analyses of PP, EC and PC:

- PP is true iff: $Pr((x)(\neg Bx \supset \neg Rx) \mid \neg Ra\&\neg Ba\&K) > Pr((x)(\neg Bx \supset \neg Rx) \mid K)$.
- EC is true iff: if $Pr(H_1 \mid E\&K) > Pr(H_1 \mid K)$ and $H_1 \equiv H_2$, then $Pr(H_2 \mid E\&K) > Pr(H_2 \mid K)$.
- PC is true iff: $Pr((x)(Rx \supset Bx) \mid \neg Ra\&\neg Ba\&K) > Pr((x)(Rx \supset Bx) \mid K)$.

It has been well known, since [Good, 1967], that whether or not PP or PC are true on the traditional Bayesian account depends on the specifics of K.[3] Whether or not EC is true, however, doesn't: EC is true, according to TB, simply by virtue of the axioms of probability calculus. This means that, although Bayesians aren't committed to PP, *if* Bayesians accept that PP is true *then* they also have to accept that PC is true as well.

Now there have been many, many different Bayesian responses to this puzzle.[4] My concern here is with the most popular rejoinder, what Fitelson and Hawthorne [Fitelson and Hawthorne, 2006] call the 'canonical' Bayesian response. The idea here is to characterize our background knowledge K, in such a way that, at least:

(a) $c[(x)(Rx \supset Bx), Ra\&Ba, K] > c[(x)(Rx \supset Bx), \neg Ra\&\neg Ba, K] > 0$ (where $c[H, E, K] :=$ the degree to which E confirms H, relative to K). In other words: the degree to which the observation of a's being a raven and being black confirms all ravens being black is greater that the (strictly positive) degree to which the observation of a's being a non-raven and being non-black confirms all ravens' being black,

if not more strongly:

(b) $c[(x)(Rx \supset Bx), Ra\&Ba, K] \gg c[(x)(Rx \supset Bx), \neg Ra\&\neg Ba, K] > 0$ and $c[(x)(Rx \supset Bx), \neg Ra\&\neg Ba, K] \approx 0$. In other words: the degree to which the observation of a's being a raven and being black confirms all ravens being black is *far* greater that the degree to which the observation of a's being a non-raven and being non-black confirms all ravens being black, the latter being close to zero.

And what is this supposed to show? Well it is supposed to account for the fact that whilst, strictly speaking, the following propositions *are both true*:

PP_2: 'observing a black raven provides some degree of support for the claim that all ravens are black, relative to what we know', and

PC: 'observing a non-black object that is a non-raven provides some degree of support for the claim that all ravens are black, relative to what we know',

we nevertheless, as a matter of psychological fact, *have the intuition* that PP_2 is true, but that PC, the problematic conclusion generated in the Ravens Paradox, is false because the degree of support conferred on the hypothesis in PP_2 is vastly greater than the degree of support conferred on the hypothesis in PC. [Vranas, 2004] sums the approach up nicely:

> "The standard Bayesian solution to the paradox tries to vindicate PC. Bayesians argue that E [(i.e. observing a non-black object that is a non-raven)] does confirm H [(i.e. the claim that all ravens are black)]... Bayesians also claim that PC looks unacceptable (i.e., we have the impression that E does not confirm H at all) because we implicitly realize that the degree to which E confirms $H$is for all practical purposes negligible (and is much smaller than the degree to which the proposition that a is both black and a raven confirms H)."

Let us grant, for sake of argument, that the constraints on K are indeed plausible and that the quantitative results that follow do indeed provide an account of why we might judge PP_2 to be true and PC false, despite the fact that they are both, strictly speaking, true.

There remains a simple problem. The results obtained are irrelevant to the paradox at hand: PP_2 and PC are the *wrong pairs of propositions.* The paradox is, let us remember, generated by noting that it seems to be reasonable to claim that [1] PP: 'Observing a non-black object that is a non-raven provides some degree of support for the claim that all non-black objects are non-ravens, relative to what we know', that it also seems reasonable to claim that [2] EC: 'If E evidentially supports H_1, relative to background knowledge K, and H_1 is equivalent to H_2, then E evidentially supports H_2, relative to K, for all H_1, H_2, E, and K.', but that nevertheless it doesn't seem reasonable to claim that [3] PC: 'Observing that a is a non-raven and is non-black evidentially supports the hypothesis that all ravens are black, relative to what we know.' despite the fact that the third claim appears to follow from the first two.

If the Bayesian wanted to respond to the problem by appealing to quantitative/comparative considerations, she would have to argue that, whilst both PP and PC are true, PC seems false because the degree of support conferred in the latter case is significantly lower than the degree of support conferred in the former. The problem is, however, that this move is unavailable: a quantitative analogue of EC is true for all existing Bayesian accounts of degree of support.

Notes

[1][Hempel, 1943, 128] is Hempel's first published mention of the problem.

[2]On the standard view of conditional probability, one must also require that the evidence has a prior probability strictly greater than 0 ($0 < Pr(E)$), or else the left-hand side of the inequality would come out as undefined.

[3]His actual example is of a K such that Ra disconfirms $(x)(Rx \supset Bx)$ relative to K. The same move can however obviously be made with respect to $\neg Ra \& \neg Ba$. Good's case obviously demonstrates the falsity, within the Bayesian framework, of the stronger so-called 'Nicod Principle' (NC) (which holds that generalizations are invariably supported by their instances, relative to what we know) from which PP is sometimes derived in certain presentations of the paradox. The intuitive plausibility of PP, of course, needn't hinge on that of NC.

[4]See [Fitelson and Hawthorne, 2006] and [Vranas, 2004] for further references.

References

[Fitelson and Hawthorne, 2006] Fitelson, B. and Hawthorne, J. (2006). How Bayesian Confirmation Theory Handles the Paradox of the Ravens. In Eells, E. and Fetzer, J., editors, *Probability in Science*. Open Court.

[Good, 1967] Good, I. J. (1967). The White Shoe is a Red Herring. *Br J Philos Sci*, 17:322.

[Hempel, 1943] Hempel, C. (1943). A Purely Syntactical Definition of Confirmation. *Journal of Symbolic Logic*, 8:122–143.

[Hosiasson-Lindenbaum, 1940] Hosiasson-Lindenbaum, J. (1940). On Confirmation. *Journal of Symbolic Logic*, 5:133–148.

[Vranas, 2004] Vranas, P. (2004). Hempel's Raven Paradox: a lacuna in the standard bayesian solution. *Br J Philos Sci*, 55:545Ð–560.