Troubles with Bayesian Argumentation

Jonas Raab[0000−0003−4292−9566]

University of Manchester, Oxford Road, M13 9PL, Manchester, United Kingdom
JoRaab@gmail.com

Abstract. **Keywords:** Bayesian Argumentation · Argument Quality · Argument Strength.

In recent years, Ulrike Hahn (together with several collaborators) has proposed a *Bayesian* account of argumentation [3], [4], [5]. As the name suggests, the Bayesian approach to argumentation suggests that (i) arguments are about the change in one's degree of belief in propositions [3, p. 1837], and (ii) this change is normatively governed by Bayes' Theorem [3, p. 1838], i.e., the appropriate change in one's degrees of belief are determined via conditionalization. With this background, the main claim is that

> in evaluating the quality of an argument from e to h, we consider to what extent it rationally increases one's belief in conclusion h[.] [3, p. 1838]

Thus, the more an argument increases one's belief in the conclusion, the better the argument.

However, I argue that there are at least four distinct problems with this approach. I think three of these criticisms provide us with some insight regarding what an appropriate account needs to accomplish (Sections 1–3), whereas the fourth problem is on its own enough to reject the Bayesian approach because it does only presuppose that we use Bayes' Theorem (Section 4).

In a nutshell, the essential parts of the Bayesian approach are (i) the claim that argumentation is about belief change, (ii) that arguments are supposed to be convincing—which is a normative matter—and (iii) that the normative force stems from the application of Bayes' theorem. The problems concern different parts of these three claims.

1 The Two Standards Issue

The first problem I identify is that there is tension between (i) and (ii), viz., that arguments classified by the Bayesian approach as 'good' are not necessarily convincing—even if the change is maximal.

Let us consider how we would try to construe an argument that is supposed to convince someone else A. Well, we would first consider what A already believes. Let us call the set of sentences that A is convinced of A's *belief set* (Bel_A). In trying to convince A that c, we would use sentences from Bel_A as premises in our argument. It seems intuitively right that the best way to convince A that c is to show how it “follows” from A's already established belief set Bel_A. If we can construe a valid argument from Bel_A to the effect that c, then A really should be convinced that c.

Moreover, if A distinguishes between levels of conviction, i.e., if for $a, b \in \mathsf{Bel}_A$, A is more convinced of a than of b, then we would prefer a over b as a premise in our argument. Again, it seems normatively more compelling that A should believe that c if we can establish c from A's strongest beliefs—given a choice.

In a way, the Bayesian approach seems to agree with all this. In particular, it spells out the level of conviction in terms or degrees of belief, i.e., in probabilities. Thus, A is more convinced of a than of b iff $P_A(a) > P_A(b)$.

But this is also where we reach the point of departure. As Hahn & Hornikx put it, an “argument should be viewed as good when it can further increase our degree of belief in a claim even when we are already quite convinced that the claim is true” [3, p. 1845]. But there is no reason to suppose that arguments that are called ‘good’ by the Bayesian approach, are convincing arguments. Hahn & Hornikx note themselves that “better arguments are more likely to meet with greater persuasive success” [3, p. 1834]. Reading “persuasive” as “convincing”, it seems that this is, indeed, a different notion than “increases the posterior probability more”. But Hahn & Hornikx even use this as motivation for considering argumentation: “[. . .] persuasive success. Identifying what arguments should count as weak and which as strong is thus a fundamental question with both theoretical and practical implications” [3, p. 1834].

In the following, Hahn & Hornikx argue that we do not need to take absolute values into account but only odds: “in evaluating strength of evidence it is ratios that ultimately matter, not absolute values” [3, p. 1838]. But it is clear that absolute values do matter when it comes to persuasion and convincing. Suppose that someone (A) has as their personal rule only to be convinced of propositions a in which A has a degree of belief greater than $\alpha \in \mathbb{R}, 0 \leq \alpha \leq 1$ (cf. Foley [1]), e.g., $\alpha = 0.9$. Then, given an argument with conclusion c, even if $P_A(c|a) > P_A(c)$, $P_A(c|a)$ might still be less than α. Then, the argument would have been a good one in the sense that it raised the posterior degree of belief, but not a good one in not accomplishing to persuade A that c (cf. Kahneman [6, ch. 39, esp. pp. 314–319]).

Furthermore, the Bayesian approach seems to imply that we really should be using the “weakest” possible premises, viz., those premises a such that $P(a)$ is small. A good argument is one for which $P(c|a) > P(c)$, i.e., $P(c|a) - P(c) > 0$. Since $0 \leq P(a) \leq 1$ for all sentences a, the maximal difference can be $1 - 0 = 1$ (which is, of course, impossible; see also Sections 2–3). Thus, it seems easier to make a good argument with conclusion c from premise a if $P_A(c)$ is small: Let $\varepsilon \in \mathbb{R}^+$, $0 < \varepsilon < 1$, such that $P_A(c) = \varepsilon$ and ε is small. Take as premise a the sentence $c \wedge d$ for any sentence d. Then, obviously, $a \vdash c$. Therefore, $P_A(c|a) = 1$ and the “argument strength” is $1 - \varepsilon$. The Bayesian approach tells us now that this was an extremely good argument; even so good that we could not do much to make it much better. But given that A's degree of belief $P_A(c)$ is so small and, thus, A's degree of belief $P_A(a,c)$ is at best equally small, why should we be convinced that c? The Bayesian quality of the argument has absolutely failed to convince us—even its strongest argument cannot convince us in every case.

2 The First Completeness Issue: Logical Reasoning

The second problem is an application of Glymour's *problem of old evidence* [2]. The idea is that some good arguments don't lead to a change in degree of belief so that claim (i) cannot be correct. Yet, such arguments can be convincing. Indeed, we commonly take the best argument to be one which shows how a conclusion follows from true premises. According to the Bayesian approach, though, none of these can be classified as 'good'.

3 The Second Completeness Issue: No Degree of Belief

The third problem is that some arguments are missed altogether. For example, there are arguments which attempt to convince us that metaphysical propositions are not meaningless. I take it that some of those who claimed them to be meaningless have been (or, at least, might have been) convinced by some such argument. However, if we take a proposition to be meaningless, we refuse to assign it a degree of belief. Yet, the successful argument leads to a change in belief—but the Bayesian approach cannot account for such arguments, even though it takes the change in belief to be crucial.

4 The Arbitrariness Issue

Lastly, even if we waive the above three problems, there is a decisive fourth one. It is concerned with claim (iii).

As the above two criticisms make obvious, for the Bayesian approach to work we have to stipulate that all sentences in question are assigned probabilities strictly between 0 and 1, i.e., for all persons A for all a, $0 < P_A(a) < 1$ (I will call these *non-extreme*). In the limit cases there could not be a good argument. This also implies that A already had some degree of belief in the conclusion c since otherwise the prior probability would not affect the argument (cf. [3, p. 1838]). But what exactly is the argument then supposed to do? Apparently, it has to change the degree of belief in some c. But A already had a degree of belief that c, viz., $P_A(c)$.

Let there be an argument with premises $a_1, \ldots, a_n$ and conclusion c. Let $P_A(a_1), \ldots, P_A(a_n)$, $P_A(c)$ be non-extreme. For simplicity, let us assume that $n = 1$ so that the argument only has the one premise $a \coloneqq a_1$ and the conclusion c. Then A believes that a to the degree $P_A(a)$ and that c to the degree $P_A(c)$. But when exactly should A change their degree of belief that c to $P_A(c|a)$? Only when it is a good argument? Or in any case?

Since the Bayesian approach is "normative" [3, p. 1838], we are told that we really *should* change our degree of belief that c to $P(c|a)$. But it is absolutely arbitrary to just do so because someone brings forward any kind of argument which leads one to relate certain probabilities to one another. It is even more arbitrary if we just change our degree of belief that c given a good argument. So how should one proceed from here?

If we endorse this approach, we can (normatively!) convince A of any sentence c such that $P_A(c)$ is non-extreme—even and especially with bad arguments. Let c be arbitrary. Since for any sentence a, $0 < P_A(a) < 1$, let's argue from any premise d such that $d \wedge \neg c \vdash \bot$ to the effect that $\neg c$. Since $0 < P_A(\neg c) < 1$ and $P(\neg c|d) = 0$, the argument was not good. But, still, $P_A(\neg c|d) = 0$, i.e., A should now be convinced that c obtains since $P_A(c) = 1 - P_A(\neg c) \overset{\text{Bayesian approach}}{=} 1 - P_A(\neg c|d) = 1 - 0 = 1$.

This shows that the Bayesian approach is in desperate need of further standards that tell us when exactly we should update our degrees of belief.

Moreover, in the classical Bayesian understanding, degrees of beliefs are updated by conditionalization only if the evidence is certain (see Talbott [7, §2]). Even if it is plausible to change this requirement to less than certain evidence, i.e., premises, the question what exactly we are supposed to update remains. If we answer this question with "everything", we run into a generalized version of the above argument: we can convince anyone of anything—I can assume whatever I like and make you conditionalize on that. If we answer "nothing", we give up the Bayesian part of the approach. Thus, we are forced to say "something, but not everything". This leads to the question what exactly is supposed to change—a question which likewise has no good answer (but I cannot present the argument here).

References

1. Foley, R.: Beliefs, Degrees of Beliefs, and the Lockean Thesis. In: Huber, F., and C. Schmidt-Petri (eds.), *Degrees of Belief*, pp. 37–47. Berlin: Springer Science+Business Media B.V. (2009)
2. Glymour, C.: Why I am Not a Bayesian. In his *Theory and Evidence*, pp. 63–93. Princeton University Press (1980)
3. Hahn, U., Hornikx, J.: A Normative Framework for Argument Quality: Argumentation Schemes with a Bayesian Foundation. *Synthese* **193**(6), 1833–1873 (2016)
4. Hahn, U., Oaksford, M.: A Bayesian Approach to Informal Argument Fallacies. *Synthese* **152**(2), 207–236 (2006)
5. Hahn, U., Oaksford, M.: Rational Argument. In: Holyoak, K.J., Morrison, R.G. (eds.), *The Oxford Handbook of Thinking and Reasoning*, pp. 277–298. Oxford University Press (2012)
6. Kahneman, D.: *Thinking, Fast and Slow*. London: Penguin Books (2011)
7. Talbott, W.: Bayesian Epistemology. *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/epistemology-bayesian/> (2016)