

UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

How to Revise Beliefs from Conditionals: A New Proposal

Permalink

<https://escholarship.org/uc/item/0882v36b>

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN

1069-7977

Authors

Hartmann, Stephan
Hahn, Ulrike

Publication Date

2021

Peer reviewed

How to Revise Beliefs from Conditionals: A New Proposal

Stephan Hartmann (S.Hartmann@lmu.de)

Munich Center for Mathematical Philosophy, LMU Munich, 80539 Munich (Germany)

Ulrike Hahn (U.Hahn@bbk.ac.uk)

Centre for Cognition, Computation, and Modelling

Department of Psychological Sciences, Birkbeck, University of London, WC1E 7HX (UK)

Abstract

A large body of work has demonstrated the utility of the Bayesian framework for capturing inference in both specialist and everyday contexts. However, the central tool of the framework, conditionalization via Bayes' rule, does not apply directly to a common type of learning: the acquisition of conditional information. How should an agent change her beliefs on learning that "If A, then C"? This issue, which is central to both reasoning and argumentation, has recently prompted considerable research interest. In this paper, we critique a prominent proposal and provide a new, alternative, answer.

Keywords: Indicative conditional reasoning; belief change; probability

Introduction

An agent entertains the propositions A and C with a prior probability distribution P defined over the corresponding propositional variables. An important question, both theoretically and practically, is how the agent should change P once she learns the natural language indicative conditional "If A, C" (Collins, Krzyżanowska, Hartmann, Wheeler, & Hahn, 2020). Following a proposal by Eva, Hartmann, and Rafiee Rad (2020), learning "If A, C" imposes the probabilistic constraint $Q(C|A) = 1$ on the new probability distribution Q . See also Eva and Hartmann (2018). The full new probability distribution is then found by minimizing an appropriate distance measure (such as the Kullback-Leibler divergence) between Q and P , taking the constraint into account. This procedure yields intuitively plausible results for test cases such as Douven's ski trip example discussed below.

It is natural to generalize this proposed procedure to the learning of non-strict conditionals. Such situations occur if the agent factors in the partial reliability of the information source or if

she takes disabling conditions into account. In such situations the constraint is $Q(C|A) = p' < 1$. Eva et al. (2020) argue that this procedure yields plausible results for key test cases in the literature.

Here, we scrutinize the generalized procedure, showing that it has normatively unacceptable consequences. As a remedy, we propose a rather minimal modification of the procedure: increasing the probability of the conjunction of A and C instead of the conditional probability. We show that minimizing the Kullback-Leibler divergence between Q and P , taking that new constraint into account, leads to intuitive, normatively plausible, consequences.

The Standard Approach

We introduce binary propositional variables A and C (in italic script) which have the values A and $\neg A$, and C and $\neg C$ (in roman script), respectively. These beliefs are represented by a probability distribution P which can be parameterized by the prior probability

$$P(A) = a \quad (1)$$

with $a \in (0, 1)$ and the conditional probabilities

$$P(C|A) = p \quad , \quad P(C|\neg A) = q \quad (2)$$

with $p, q \in (0, 1)$. The following proposition summarizes our findings about the new probability distribution Q after learning a natural language indicative conditional.¹

Proposition 1 *An agent considers the propositions A and C with a prior probability distribution P defined in eqs. (1) and (2). Learning the*

¹Here and below we use the convenient shorthand notation $P(A, C)$ for $P(A \wedge C)$.

conditional “If A, C ” then imposes the constraint $Q(C|A) = p'$ (with $p < p' < 1$) on the new probability distribution Q . Minimizing the Kullback-Leibler divergence between Q and P then implies the following claims: (i) $Q(A) < P(A)$, (ii) $Q(C) > P(C)$, (iii) $Q(A, C) > P(A, C)$, and (iv) $Q(A|C) > P(A|C)$.

These qualitative results seem plausible. Item (i) is perhaps most controversial, but it can be justified by observing that learning a conditional makes its antecedent more informative. But more informative propositions have a lower probability, and so the probability of the antecedent should decrease.

We now consider the behavior of $Q(A)$, $Q(C)$ and $Q(A, C)$ as p' increases from p to 1. To do so, we parameterize² $p' = \lambda + p \cdot \bar{\lambda}$ with $\lambda \in (0, 1)$ and plot the three quantities as a function of λ . Fig. 1 shows that $Q(A)$ decreases strictly monotonically as p' increases. This is plausible as there is no reason why $Q(A)$ should have an extreme value.

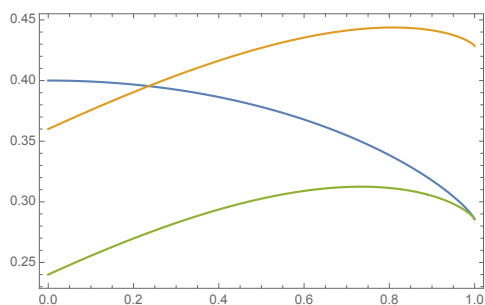


Figure 1: $Q(A)$ (blue), $Q(C)$ (orange) and $Q(A, C)$ (green) as a function of λ for $a = .4, p = .6, q = .2$.

Interestingly, this is exactly what happens for $Q(C)$ and $Q(A, C)$ —a behavior which is not normatively plausible. Once the strength of the conditional (as measured by the parameter p') increases, one would expect the impact of learning the conditional on the two respective probabilities to be a strictly monotonic function of p' . But it is not if one uses the present procedure. Here is a more complex test case from the literature:

The Ski Trip Example. Harry sees his friend Sue buying a ski outfit. This surprises him a

²Here and below we use the convenient shorthand notation $\bar{x} := 1 - x$.

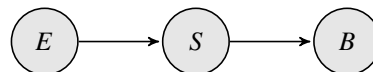


Figure 2: The Bayesian network for the ski trip example

bit, because he did not know that she had any plans to go on a ski trip. He knows that she recently had an important exam and thinks it unlikely that she passed it. Then he meets Tom, his best friend and also a friend of Sue’s, who is just on his way to Sue to hear whether she passed the exam, and who tells him: “If Sue passed the exam, her father will take her on a ski vacation.” Recalling his earlier observation, Harry now comes to find it more likely that Sue passed the exam. So in this example upon learning the conditional information Harry should intuitively increase the probability of the antecedent of the conditional. (Douven & Dietz, 2011)

Eva et al. (2020) analyze this example as follows. First, they introduce the variables E (“Sue passed the exam”), S (“Sue’s father invites her for a ski trip”) and B (“Sue buys a new ski outfit”). Second, they assume that B and E are probabilistically independent given S (see Fig. 2). Third, they fix the prior probability of E , i.e.

$$P(E) = e \quad (3)$$

with $e \in (0, 1)$ and the conditional probabilities

$$\begin{aligned} P(S|E) = p_1 & \quad , & P(S|\neg E) = q_1 \\ P(B|S) = p_2 & \quad , & P(B|\neg S) = q_2 \end{aligned} \quad (4)$$

with $p_1, p_2, q_1, q_2 \in (0, 1)$. Fourth, they note that the agent does not only learn a conditional, but also that Sue bought a new ski outfit. If one conditionalizes on these two pieces of information (representing the conditional $E \rightarrow S$ as the material conditional $\neg E \vee S$), then one obtains $P^*(E) = e / (e + \bar{e} \cdot l_0)$ with the likelihood ratio $l_0 := (q_1 p_2 + \bar{q}_1 q_2) / (p_1 p_2)$. Hence, $P^*(E) > P(E)$ iff $l_0 < 1$. Eva et al. (2020) then show that this condition is fulfilled given the information provided in the description of the example. The authors also show that one obtains exactly the same result if one minimizes the Kullback-Leibler divergence with the

constraints $Q(B) = 1$ and $Q(S|E) = 1$. The following proposition shows what happens if one relaxes the second constraint.

Proposition 2 *An agent considers the propositions B, E and S with a prior probability distribution P defined in eqs. (3) and (4). Fig. 2 represents the assumed conditional independencies. Learning the information mentioned in the example then imposes the constraints $C_1 : Q(B) = 1$ and $C_2 : Q(S|E) = \lambda + P(S|E) \cdot \bar{\lambda}$ (with $\lambda \in (0, 1)$) on the new probability distribution Q. Minimizing the Kullback-Leibler divergence between Q and P then implies that $Q(E)$ has a maximum in the open interval $(0, 1)$ as a function of $P(S|E)$.*

To illustrate Proposition 2 we plot $Q(E)$ as a function of λ which measures the increase of the conditional probability from $P(S|E)$ (for $\lambda = 0$) to 1 (for $\lambda = 1$).

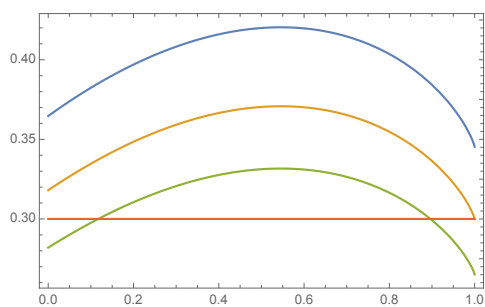


Figure 3: $Q(E)$ as a function of λ for $e = .3$, $p_1 = .4$, $p_2 = .8$, $q_2 = .2$ and $q_1 = .1$ (blue), $q_1 = .2$ (orange) and $q_1 = .3$ (green).

Fig. 3 clearly show that $Q(E)$ has a maximum. That is, after a certain value of λ , $Q(E)$ decreases. It is even possible that $Q(E)$ becomes smaller than $P(E)$ as λ (and therefore p'_1) increases. This challenges our normative intuitions, suggesting a rejection of the present approach. As we would like to keep the spirit of the distance-minimization approach, we propose to replace the conditional probability constraint by a normatively more appealing alternative.

Before presenting our new proposal, we would like to note another interesting feature exhibited in Fig. 3: For small values of λ it is possible that $Q(E) < P(E)$. In this case, the learned information

disconfirms E. This, however, is plausible as for small values of λ the contribution of learning the other item of information (i.e. the proposition B) dominates and we know already (see Proposition 1) that learning only this information *decreases* the probability of E. But once the conditional becomes sufficiently strong (i.e. once λ is above a certain threshold), learning the conditional dominates the learning of B and $Q(E) > P(E)$.

A New Proposal

We have seen that the standard approach, which assumes that learning a conditional prompts the agent to increase the corresponding conditional probability, leads to normatively implausible results. Alternatively, we propose that the agent increases the probability of the conjunction of the antecedent and the consequent instead. The following proposition states what this entails.

Proposition 3 *An agent considers the propositions A and C with a prior probability distribution P defined in eqs. (1) and (2). Learning the conditional “If A, C” then imposes the constraint $Q(A, C) = \alpha + P(A, C) \cdot \bar{\alpha}$ (with $\alpha \in (0, 1)$) on the new probability distribution Q. Minimizing the Kullback-Leibler divergence between Q and P then implies the following claims: (i) $Q(A) > P(A)$, (ii) $Q(C|A) > P(C|A)$, (iii) $Q(C) > P(C)$, and (iv) $Q(A)$, $Q(C|A)$ and $Q(C)$ are increasing functions of α .*

This is an interesting result. We first note that the increase of the probability of the conjunction of A and C (i.e. $a' p' > a p$) implies an increase of the probability of the antecedent (i.e. $a' > a$) and an increase of the conditional probability of the consequent, given the antecedent (i.e. $p' > p$). This could, of course, be different. It would have been possible that one of the two factors increases strongly, while the other one decreases, but not so much, making sure that the product increases. Next, it is interesting to see that the probability of the antecedent increases in the same proportion as the probability of the conjunction of A and C increases. That is, as a result of learning the conditional “If A, C”, the probability of the antecedent does not decrease as for the constraint $Q(C|A) = 1$, but it increases.

This is intuitively plausible as the conditional informs us about the relation between A and C which has something to do with the joint occurrence of A and C. Hence, after the conditional is stated, we expect A to be more likely to occur than if the conditional would not have been stated. Note also that the new probability of the antecedent only depends on the prior probability of the antecedent (and of course on α), but not on the prior conditional probabilities p and q . Finally we note that p' increases with α (for fixed a and p) and decreases with a (for fixed α and p). p' does not depend on q . The following proposition presents the results for the ski trip example on our new proposal.

Proposition 4 *An agent considers the propositions B, E and S with a prior probability distribution P defined in eqs. (3) and (4). Fig. 2 represents the assumed conditional independencies. Learning the information mentioned in the example then imposes the constraints $C_1 : Q(B) = 1$ and $C_2 : Q(E, S) = \alpha + P(E, S) \cdot \bar{\alpha}$ (with $\alpha \in (0, 1)$) on the new probability distribution Q. Minimizing the Kullback-Leibler divergence between Q and P then implies the following claims: (i) $Q(E) > P(E)$ for $\alpha > \alpha_c$ (specified in the proof). If $P(E) \approx 0$ and $Q(B|\neg S) \approx 0$, then $\alpha_c \approx P(E)P(\neg S|E) \approx 0$, (ii) $Q(S|E) > P(S|E)$ if $P(B|S) > P(B|\neg S)$, (iii) $Q(E)$ and $Q(S|E)$ are increasing functions of α , and (iv) $Q(S|\neg E) > P(S|\neg E)$ iff $P(B|S) > P(B|\neg S)$.*

It is interesting to compare these results to the results of the standard approach presented in Proposition 2. The standard approach assumes that learning a conditional can be modeled by increasing the conditional probability of the consequent given the antecedent. Then the new probability of E, i.e. $Q(E)$ first increases and then, as the conditional probability (measured by the parameter λ in Fig. 3) increases, decreases. This is counter-intuitive if one takes the conditional probability as a measure of the “strength” of the corresponding conditional. However, if one follows the new proposal, one finds that $Q(E)$ strictly monotonically increases as a function of the “strength” of the conditional (i.e. the respective probability of the conjunction of the antecedence and the consequent measured by the parameter α). One also finds that $Q(E) > P(E)$ once α passes a certain

threshold α_c . This is plausible as for small values of α the contribution of learning the other item of information (i.e. the proposition B) dominates and we can argue as above. Once the conditional becomes sufficiently strong (i.e. once $\alpha > \alpha_c$), learning the conditional dominates the learning of B and $Q(E) > P(E)$. To get a sense of the numerical value of the threshold in the ski trip example, we make two observations: (i) $P(E) \approx 0$ because Harry “thinks it is unlikely that [Sue] passed the exam”. (ii) $Q(B|\neg S) \approx 0$ because the story does not mention any reason for Harry to expect Sue to buy a skiing outfit if she is not invited on a skiing trip. Hence, Proposition 4 informs us that $\alpha_c \approx 0$. For larger values of $P(E)$ and $Q(B|\neg S)$, α_c is correspondingly larger.

Discussion

We have provided a new normative proposal for belief change in response to a conditional. It remains for the discussion to relate this proposal to concerns in the wider literature. One such concern is that the natural-language conditional is directional. Our constraint, however, is symmetrical. This issue is part of the wider concern about “centering” as commonly understood in the psychology of reasoning: the claim that the inference from conjunction to conditional is valid (Cruz, Baratgin, Oaksford, & Over, 2015).

First, while this claim is clearly true for the material conditional, it is not at all clear that it is true for the indicative conditional of natural language. Not only is it clear that the material conditional is not an adequate formalization of the indicative conditional, there is also a growing body of experimental evidence suggesting that the indicative conditional involves some “connection” between antecedent and consequent (e.g., Krzyżanowska, Wenmackers, and Douven (2013); Mirabile and Douven (2020)), though it remains a topic of debate whether this connection is properly construed as semantic or pragmatic). On such an account of the indicative conditional, inference from conjunction to conditional is no more valid than the erroneous inference from correlation to causation against which students of even the most introductory statistics class are warned: from the fact that cigarette packaging and lung cancer co-occur it does not follow that cigarette packaging causes

lung cancer. Indeed, even where conjunctions have probability 1, so-called “missing link” conditionals such as *if roses are plants, roses have thorns* are infelicitous (Krzyżanowska, Collins, & Hahn, 2017).

Crucially, the account presented in this paper does *not* fall prey to warranting such inference. The relationship detailed is between the conjunction and the conditional probability. This relationship is indeed inferentially symmetric in that increasing one will increase the other (*ceteris paribus*). However, this is distinct from what constitutes the semantics/pragmatics of the natural language conditional. Some accounts equate the indicative conditional with the conditional probability (Evans, Handley, & Over, 2003; Stalnaker, 1970), but those accounts need not be, and, in our view, likely are not empirically adequate for the reasons just outlined. On any account where there is “more” to the semantics or pragmatics of the indicative conditional than the conditional probability, it will simply not follow that raising the probability of the conjunction raises that of the conditional. Indeed, there may be no such quantity as the probability of the conditional, because the conditional does not constitute a proposition in the first place (Adams, 1975).

As a result, the constraint posited in this paper is *not* symmetrical, and thus does not fall prey to familiar problems of asymmetry known from the literature on explanation (Salmon, 1992). Our account is largely neutral on the details of the semantics of the conditional itself and assumes only that the indicative conditional imposes a probabilistic constraints. This must not be confused with the idea that this is what the conditional *means*.

Proofs

We begin with three preliminary remarks. First, the Kullback-Leibler divergence between Q and P is given by

$$D_{KL}(Q||P) := \sum_{i=1}^n Q(S_i) \log \frac{Q(S_i)}{P(S_i)}.$$

Second, following Eva et al. (2020), we define

$$\Phi_x := x' \log \frac{x'}{x} + \bar{x}' \log \frac{\bar{x}'}{\bar{x}}.$$

Note that $\Phi_x > 0$ for $x' \neq x$ and that $\partial\Phi_x/\partial x' = 0$ implies that $x' = x$. Third, readers unfamiliar with the use of Bayesian networks in rationality research are referred to Hartmann (2020) for a concise introduction.

Proof of Proposition 1

Proposition 4 in Eva et al. (2020) implies that $q' = q$ and $a' = a/(a + \bar{a} \cdot l)$ with $l = (p'/p)^{p'} (\bar{p}'/\bar{p})^{\bar{p}'}$. Consider $Q(A)$ first: $l > 1$ is equivalent to $\Phi_p > 0$ which holds for $p' > p$. Hence, $Q(A) < P(A)$. Second, we show that $Q(A, C) = a' p' > a p = P(A, C)$. To do so, we first note that the previous inequality is equivalent to $p'/p - 1 > \bar{a}(l - 1)$. It therefore suffices to show that $p'/p > l$ which follows from simple algebraic manipulations. Third, $Q(C) = a' p' + \bar{a}' q > a p + \bar{a} q = P(C)$ follows from $a' p' > a p$ and the fact that $a' \leq a$ implies that $\bar{a}' \geq \bar{a}$. Fourth, we consider $Q(A|C) = a' p' / (a' p' + \bar{a}' q)$. With $a' = a/(a + \bar{a} \cdot l)$ we find that $Q(A|C) = a p' / (a p' + \bar{a} l q)$. Hence, defining $\Delta := Q(A|C) - P(A|C)$, we obtain

$$\begin{aligned} \Delta &= \frac{a p'}{a p' + \bar{a} l q} - \frac{a p}{a p + \bar{a} q} \\ &= \frac{a \bar{a} q \cdot (p' - p l)}{(a p' + \bar{a} l q)(a p + \bar{a} q)}, \end{aligned}$$

from which we conclude that $Q(A|C) > P(A|C)$ iff $p'/p > l$, which holds as shown above. ■

Proof of Proposition 2

We first note that the constraint C_1 implies that $p'_2 = q'_2 = 1$. The Kullback-Leibler divergence is then given by

$$\begin{aligned} KL &= \Phi_e + e' \Phi_{p_1} + \bar{e}' \Phi_{q_1} - \\ &= (e' p'_1 + \bar{e}' q'_1) \cdot \log p_2 - (e' \bar{p}'_1 + \bar{e}' \bar{q}'_1) \log q_2. \end{aligned}$$

Next, we differentiate KL with respect to q'_1 and set the resulting expression equal to zero. This yields

$$q'_1 = \frac{q_1 p_2}{q_1 p_2 + \bar{q}_1 q_2}.$$

Analogously, we obtain $e' = e/(e + \bar{e} l)$ with

$$l = \left(\frac{p'_1}{u}\right)^{p'_1} \left(\frac{\bar{p}'_1}{v}\right)^{\bar{p}'_1} \cdot w$$

and $u := p_1 p_2$, $v := \bar{p}_1 q_2$ and $w := q_1 p_2 + \bar{q}_1 q_2$. We can now show that l always has a minimum in $(0, 1)$ as a function of p'_1 . (Hence, e' always has a maximum in $(0, 1)$ as a function of p'_1 .) To do so, we note that $\log l$ is a strictly monotonically increasing function of l . Hence, l and $\log l$ have the same extrema provided that l does not vanish at the extremal points. (As $l > 0$, this does not happen.) We therefore differentiate $\log l$ by p'_1 ,

$$\frac{\partial \log l}{\partial p'_1} = \log \left(\frac{p'_1 v}{p'_1 u} \right) + u - v,$$

and set the resulting expression equal to zero. This yields

$$p'_1 = \frac{u \exp(v)}{u \exp(v) + v \exp(u)}.$$

Note that $p'_1 \in (0, 1)$ as $u, v \in (0, 1)$. Hence, p'_1 is an extremum point. The extremum is a minimum as $\partial^2 \log l / \partial p_1'^2 = 1 / (p'_1 \bar{p}'_1) > 0$. ■

Proof of Proposition 3

We have to minimize $L = KL + \mu(a' p')$ with

$$KL = \Phi_a + a' \Phi_p + \bar{a}' \Phi_q$$

and a Lagrange multiplier μ . To do so, we first differentiate L with respect to q' and set the resulting expression equal to zero. This yields $q' = q$. Analogously for a' and p' , yielding

$$a' = \frac{a(p + \bar{p}x)}{a(p + \bar{p}x) + \bar{a}x}$$

and $p' = p / (p + \bar{p}x)$ with $x := \exp(\mu)$. To determine x (and therewith μ), we insert the expressions for p' and a' into the constraint $a' p' = \alpha + a p \bar{\alpha}$. This yields

$$x = \frac{a p \bar{\alpha}}{\alpha + a p \bar{\alpha}}.$$

Inserting this expression into the expressions for a' and p' and yields

$$\begin{aligned} a' &= \frac{\alpha + a \bar{\alpha}}{\alpha + a p \bar{\alpha}} \\ p' &= \frac{\alpha + a p \bar{\alpha}}{\alpha + a \bar{\alpha}}. \end{aligned}$$

With this, we can show the claims made in the proposition. (i) $a' = \alpha + a \bar{\alpha} = a + \bar{a} \alpha > a$. (ii) We

note that

$$\begin{aligned} p' &= \frac{\alpha + a p \bar{\alpha}}{\alpha + a \bar{\alpha}} \\ &= \frac{p \alpha + \bar{p} \alpha + a p \bar{\alpha}}{\alpha + a \bar{\alpha}} \\ &= p + \frac{\bar{p} \alpha}{\alpha + a \bar{\alpha}} > p. \end{aligned}$$

(iii) We first set $c := P(C) = a p + \bar{a} q$ and $c' := Q(C) = a' p' + \bar{a}' q$. Next, inserting the expressions for a' and p' , we obtain $c' = \alpha + c \bar{\alpha}$. Hence, $c' > c$. (iv) We compute $\partial a' / \partial \alpha = \bar{a} > 0$, $\partial p' / \partial \alpha = a \bar{p} / (\alpha + a \bar{\alpha})^2 > 0$, and $\partial c' / \partial \alpha = \bar{c} > 0$. ■

Proof of Proposition 4

We have to minimize $L = KL + \mu(e' p'_1)$ with

$$\begin{aligned} KL &= \Phi_e + e' \Phi_{p_1} + \bar{e}' \Phi_{q_1} - \\ &(e' p'_1 + \bar{e}' q'_1) \log p_2 - (e' \bar{p}'_1 + \bar{e}' \bar{q}'_1) \log q_2 \end{aligned}$$

and a Lagrange multiplier μ . L takes the constraints C_1 and C_2 into account. More explicitly, C_2 is given by

$$e' p'_1 = \alpha + \bar{\alpha} e p_1. \quad (5)$$

Differentiating L with respect to q'_1 and setting the resulting expression equal to zero yields

$$q'_1 = \frac{q_1 p_2}{q_1 p_2 + \bar{q}_1 q_2}.$$

Note that $q'_1 > q_1$ iff $p_2 > q_1 p_2 + \bar{q}_1 q_2$ iff $p_2 > q_2$. This completes the proof of claim (iv).

To determine e' and p'_1 , we proceed analogously. Setting $u := p_1 p_2$, $v := \bar{p}_1 q_2$ and $w := q_1 p_2 + \bar{q}_1 q_2$ (as in Proposition 2), we obtain

$$\begin{aligned} e' &= \frac{e(u + vx)}{e u + (e v + \bar{e} w) x} \\ p'_1 &= \frac{u}{u + vx}, \end{aligned}$$

with $x := \exp(\mu)$. Inserting these expressions into eq. (5) and solving for x yields

$$x = \frac{e u \bar{\alpha}}{e v + \bar{e} w} \cdot \frac{1 - e p_1}{\alpha + e p_1 \bar{\alpha}}.$$

With this, we can calculate e' and p'_1 :

$$\begin{aligned} e' &= \frac{e v + \bar{e} w (\alpha + e p_1 \bar{\alpha})}{e v + \bar{e} w} \\ p'_1 &= \frac{(e v + \bar{e} w) (\alpha + e p_1 \bar{\alpha})}{e v + \bar{e} w (\alpha + e p_1 \bar{\alpha})} \end{aligned}$$

We use these expressions to proof the remaining claims. First, we note that

$$e' - e = \frac{\bar{e} [(1 - e p_1) w \alpha - e \bar{p}_1 q_1 (p_2 - q_2)]}{e v + \bar{e} w}.$$

Hence, $e' > e$ iff $\alpha > \alpha_c$ with

$$\alpha_c := \frac{e \bar{p}_1 q_1 (p_2 - q_2)}{(1 - e p_1) w}.$$

Note that if $q_2 \approx 0$, then $w \approx q_1 p_2$ and therefore $\alpha_c \approx e \bar{p}_1 / (1 - e p_1)$. If also $e \approx 0$ holds, then we obtain $\alpha_c \approx e \bar{p}_1$. This completes the proof of claim (i).

Second, we note that

$$p'_1 - p_1 = \frac{e \bar{e} p_1 \bar{p}_1 q_1 (p_2 - q_2) + A \alpha}{e v + \bar{e} w (\alpha + e p_1 \bar{\alpha})},$$

with

$$A = (1 - e p_1)(e v + \bar{e} \bar{p}_1 w) > 0.$$

Hence, $p'_1 > p_1$ if $p_2 > q_2$. This completes the proof of claim (ii).

Third and finally, we calculate

$$\begin{aligned} \frac{\partial e'}{\partial \alpha} &= \frac{\bar{e} w}{e v + \bar{e} w} \cdot (1 - e p_1) > 0 \\ \frac{\partial p'_1}{\partial \alpha} &= \frac{e v (e v + \bar{e} w) \cdot (1 - e p_1)}{(e v + \bar{e} w (\alpha + e p_1 \bar{\alpha}))^2} > 0. \end{aligned}$$

This completes the proof of claim (iii). ■

References

- Adams, E. W. (1975). *Logic of Conditionals*. Dordrecht: Reidel.
- Collins, P. J., Krzyżanowska, K., Hartmann, S., Wheeler, G., & Hahn, U. (2020). Conditionals and testimony. *Cognitive Psychology*, 122, 101329.
- Cruz, N., Baratgin, J., Oaksford, M., & Over, D. E. (2015). Bayesian reasoning with ifs and ands and ors. *Frontiers in Psychology*, 6, 192.
- Douven, I., & Dietz, R. (2011). A puzzle about Stalnaker's hypothesis. *Topoi*, 30(1), 31–37.
- Eva, B., & Hartmann, S. (2018). Bayesian argumentation and the value of logical validity. *Psychological Review*, 125(5), 806–821.
- Eva, B., Hartmann, S., & Rafiee Rad, S. (2020). Learning from conditionals. *Mind*, 129(514), 461–508.
- Evans, J. S. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(2), 321–335.
- Hartmann, S. (2020). Bayes nets and rationality. In M. Knauff & W. Spohn (Eds.), *Handbook of Rationality*. Cambridge, MA: MIT Press.
- Krzyżanowska, K., Collins, P. J., & Hahn, U. (2017). The puzzle of conditionals with true clauses. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Annual Conference of the Cognitive Science Society*. Austin, TX: Cognitive Science Society.
- Krzyżanowska, K., Wenmackers, S., & Douven, I. (2013). Inferential conditionals and evidentiality. *Journal of Logic, Language and Information*, 22(3), 315–334.
- Mirabile, P., & Douven, I. (2020). Abductive conditionals as a test case for inferentialism. *Cognition*, 200, 104232.
- Salmon, W. C. (1992). Scientific Explanation. In M. H. Salmon et al. (Eds.), *Introduction to the Philosophy of Science* (p. 7-41). Englewood Cliffs, NJ: Prentice-Hall.
- Stalnaker, R. C. (1970). Probability and conditionals. *Philosophy of Science*, 37(1), 64–80.