A New Probabilistic Explanation of the Modus Ponens–Modus Tollens Asymmetry

Benjamin Eva (Benjamin.Eva@uni-konstanz.de)
Department of Philosophy, Univeristy of Konstanz, 78464 Konstanz (Germany)

Stephan Hartmann (S.Hartmann@lmu.de)
Munich Center for Mathematical Philosophy, LMU Munich, 80539 Munich (Germany)

Henrik Singmann (Henrik.Singmann@warwick.ac.uk)
Department of Psychology, University of Warwick, Coventry, CV4 7AL (UK)

Abstract

A consistent finding in research on conditional reasoning is that individuals are more likely to endorse the valid modus ponens (MP) inference than the equally valid modus tollens (MT) inference. This pattern holds for both abstract task and probabilistic task. The existing explanation for this phenomenon within a Bayesian framework (e.g., Oaksford & Chater, 2008) accounts for this asymmetry by assuming separate probability distributions for both MP and MT. We propose a novel explanation within a computational-level Bayesian account of reasoning according to which "argumentation is learning". We show that the asymmetry must appear for certain prior probability distributions, under the assumption that the conditional inference provides the agent with new information that is integrated into the existing knowledge by minimizing the Kullback-Leibler divergence between the posterior and prior probability distribution. We also show under which conditions we would expect the opposite pattern, an MT-MP asymmetry.

Keywords: conditional reasoning; probabilistic reasoning; Bayesian model; computational-level account

Introduction

Conditionals of the form "If A, then C" – for example, "If global warming continues, then London will be flooded" – are ubiquitous in everyday language and scientific discourse. One research question that has attracted a lot of attention is how individuals reason with conditionals. Usually four conditional inferences are studied, each consisting of the conditional as the major premise, a categorical minor premise, and a putative conclusion:

- *Modus Ponens* (MP): If A then C. A. Therefore, B.
- *Affirmation of the Consequent* (AC): If A then C. C. Therefore, A.
- *Denial of the Antecedent* (DA): If A then C. Not A. Therefore, not B.
- *Modus Tollens* (MT): If A then C. Not C. Therefore, not A.

According to classical logic MP and MT are valid (i.e., truth preserving) inferences and AC and DA are not valid. Early research with conditional inferences has emulated the inference process of classical logic; in the *abstract task*, inferences are presented with abstract content, participants are asked to treat the premises as true, and are asked to only accept necessary conclusions. Results generally showed that even untrained participants are able to distinguish valid from invalid inferences (i.e., they accept more valid than invalid inferences). However, their behavior is clearly not in line with the norms of classical logic. Whereas participants tend to unanimously accept the valid MP, the acceptance rates for the equally valid MT inference scheme is considerably lower. In a meta-analysis of the abstract task, Schroyens, Schaeken, and d'Ydewalle (2001) found acceptance rates of .97 for MP compared to acceptance rates of .74 for MT. This *MP-MT asymmetry* will be the main focus of the present manuscript.[1]

Research in recent years has moved away from the abstract task and its focus on logical validity towards tasks more akin to real-life reasoning within a *probabilistic framework* (Oaksford & Chater, 2007; Over, 2009). In the *probabilistic task*, inferences employ everyday content for which participant posses relevant background knowledge and they are usually asked for their subjective degree of belief in the putative conclusions. The degree of belief in the conclusion of course depends on the actual content (i.e., the probabilistic relationships among premises and conclusion), but there is still ample evidence for an MP-MT asymmetry that goes beyond what would be expected from existing probabilistic accounts. For example, Oaksford, Chater, and Larkin (2000) created materials for which their Bayesian model of conditional reasoning predicted participants to posses similar beliefs in MP and MT. Their results showed that, whereas this reduced the asymmetry, there were still differences such that participants expressed stronger beliefs in MP than MT. Similarly, Singmann, Klauer, and Over (2014) asked participants for their subjective degrees of belief, first in both premises, and then in the conclusion and showed that those only formed a coherent probability distribution "above chance" for MP, but not for MT. Essentially the same results were obtained by Evans, Thompson, and Over (2015). Together, these findings suggest a clear limit for simple probabilistic accounts of conditional reasoning.

Existing Accounts

To describe existing accounts and our new explanation, let us formalize the probabilistic structure of the reasoning problem. We consider an agent who entertains the propositions A (the antecedent) and C (the consequent) of a conditional

[1] Participants also tend to erroneously accept the invalid inferences AC and DA. Schroyens et al. (2001) report acceptance rates of .64 for AC and .56 for DA.

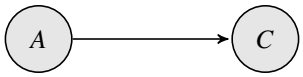


Figure 1: The Bayesian Network representation of the relation between A and C.

"If A, then C". To proceed, we introduce binary propositional variables A and C (in italic script) which have the values A and ¬A, and C and ¬C (in roman script), respectively. A prior probability distribution P is defined over these variables. It is represented by the Bayesian Network in Figure 1. For now, the exact parameterization of P is not yet relevant, however there exist several with three free parameters. In addition, we define the (absolute) *endorsement* of MP as $E_1 := P'(\mathrm{C})$. Similarly, we define the (absolute) endorsement of DA as $E_2 := P'(\neg\mathrm{C})$, the (absolute) endorsement of AC as $E_3 := P'(\mathrm{A})$, and the (absolute) endorsement of MT as $E_4 := P'(\neg\mathrm{A})$.

The original Oaksford et al. (2000) model makes two assumptions. First, it assumes that belief in the conclusion reflects the conditional probability of the conclusion given the minor premise. For example, $E_1 = P(\mathrm{C}|\mathrm{A})$ and $E_4 = P(\neg\mathrm{A}|\neg\mathrm{C})$. Second, it assumes that P is fixed throughout the reasoning process; that is, P is the same for responses to the four conditional inferences (Oaksford and Chater call this the "invariance assumption"). In other words, this model assumes that reasoning amounts to consulting ones fixed probability distribution and responding in line with it (e.g., by sampling from memory; Costello & Watts, 2014) . As shown by Oaksford and Chater (2007, ch.5) this model does a good job in accounting for many of the existing data from the abstract task, but underestimates the MP-MT asymmetry.

To account for the MP-MT asymmetry, the solution first proposed in Oaksford and Chater (2007, ch.5) and subsequently defended in (Oaksford & Chater, 2008, 2013) is to give up on the second of their assumptions, that P is fixed for responses to all four inferences. Specifically, they argue (e.g., Oaksford & Chater, 2013) that MP represents a special case that does not require changing P as it basically reflects the probabilistic information already present in the conditional (i.e., $P(\mathrm{C}|\mathrm{A})$). Thus, presenting MP does not allow the agent to learn new information about P. However, the other three inferences, MT, AC, and DA, present new information and thus require an updated probability distribution P', which individuals "learn". Practically, they did not specify many restriction of P', other than that $P(\mathrm{C}|\mathrm{A}) > P'(\mathrm{C}|\mathrm{A})$, which was primarily motivated by fitting their model to the extant data. From a statistical point of view, it not too surprising that a model that then essentially has one free parameter for fitting E_1 and three free parameters for fitting the remaining three observations (i.e., responses to the other three inferences, E_2, E_3, and E_4) does a relatively good job in accounting for the existing data.

Therefore, there are two main theoretical shortcomings in Oakford and Chater's approach. First, their revised model assumes that the endorsement for MP, E_1, comes from one probability distribution, P, whereas the endorsement for the other inferences, E_2 to E_4, comes from the updated probability distribution P'. This seems somewhat unsatisfactory from a rational Bayesian perspective and more of an ad-hoc solution than a principled argument. Second, the actual processes in which the agent updates P to arrive at P' are not specified well enough. What does it entail for the agent to learn the new information presented in MT? How can we characterize the cognitive processes involved in making a probabilistic MP or MT inference?

Our answer to these questions is based on Eva and Hartmann's (2018) recent Bayesian account of reasoning according to which "argumentation is learning". In line with Oaksford and Chater (2013), learning is specified as updating an agent's prior belief state, represented by P, in light of new information resulting in the posterior belief state P'. Specifically, the premises of an inference will affect specific parts of P (e.g., for MP, the agent learns the new values of both $P(\mathrm{C}|\mathrm{A})$ and $P(\mathrm{A})$). The novel assumption is that as a consequence, the agent needs to incorporate this new information into their existing beliefs which requires her to update potentially all parts of P. According to Eva and Hartmann (2018), this updating follows a well-defined Bayesian rule which generalizes conditonalization and Jeffrey conditionalization and requires that a suitably defined distance (or divergence) between P' and P is minimized. Eva and Hartmann (2018) argue that these divergencies should be members of the family of f-divergences. One important member of this family is the Kullback-Leibler (KL) divergence (Diaconis & Zabell, 1982), which we will use in the remainder. In this way, updating satisfies the constraints provided by the new information and is conservative (i.e., the changes are as minimal as possible). We will show that from this assumption, the typically found MP-MT asymmetry must appear for certain P. However, in some situations the opposite pattern (i.e., $E_4 > E_1$) should also be observed.

The Model

Our new explanation for the MP-MT asymmetry is based on the Bayesian Network in Figure 1 representing the prior probability distribution P. In addition, we assign

$$P(\mathrm{A}) = a, \tag{1}$$

for the prior probability of the antecedent and the conditional probabilities of the consequent C, given the values of its parent:

$$P(\mathrm{C}|\mathrm{A}) = p \quad , \quad P(\mathrm{C}|\neg\mathrm{A}) = q \tag{2}$$

With this, the joint prior probability distribution P over the variables A and C is given by

$$\begin{aligned} P(\mathrm{A},\mathrm{C}) &= a\,p \quad , \quad P(\mathrm{A},\neg\mathrm{C}) = a\,\overline{p} \\ P(\neg\mathrm{A},\mathrm{C}) &= \overline{a}\,q \quad , \quad P(\neg\mathrm{A},\neg\mathrm{C}) = \overline{a}\,\overline{q}, \end{aligned} \tag{3}$$

where we have used the shorthand notation $P(\mathrm{A},\mathrm{C})$ for $P(\mathrm{A}\wedge\mathrm{C})$ which we will use throughout this paper. We also use the shorthand $\overline{x}$ for $1-x$ and assume that $a,p,q\in(0,1)$.

Following the slogan "argumentation is learning", the agent then learns the premises of the argument. More specifically, she learns the major premise "If A, then C" and sets the new probability of $P'(\mathrm{C}|\mathrm{A})=p'=1$ in turn. This is the first constraint on P'. She also learns a minor premise: A in the case of MP, and ¬C in the case of MT. For completeness, we also consider AC and DA. In the case of AC, she additionally learns C, and in the case of DA she additionally learns ¬A. Following Eva and Hartmann (2018), we model this by assuming that the probability of the minor premise increases. This is the second constraint on P'. More specifically, we assume that the agent changes the probabilities of the minor premise in the following way:

$$\begin{aligned} P'_{MP}(\mathrm{A}) &= \lambda+\overline{\lambda}P(\mathrm{A}) \quad , \quad P'_{DA}(\mathrm{A}) = \overline{\lambda}P(\mathrm{A}) \\ P'_{AC}(\mathrm{C}) &= \lambda+\overline{\lambda}P(\mathrm{C}) \quad , \quad P'_{MT}(\mathrm{C}) = \overline{\lambda}P(\mathrm{C}) \end{aligned} \tag{4}$$

Here $\lambda\in(0,1]$ measures to what extent the agent changes the probability of the minor premise. For $\lambda\to 0$, the new probability of the minor premise does not change at all, and for $\lambda=1$ it goes to its maximal value, i.e. to 1.

To find the full new probability distribution P', we then minimize the KL-divergence between P' and P. This allows us to compute the new probability of the conclusion of the corresponding argument. For example, in the case of MP ($\mathrm{A},\mathrm{A}\to\mathrm{C}$, therefore C) the conclusion is C and the new probability of C, i.e. $P'(\mathrm{C})$ measures to what extent the agent endorses the corresponding inference pattern. More specifically, we define the (absolute) *endorsement* of MP as $E_1 := P'(\mathrm{C})$. As described above, we define the (absolute) endorsement of DA as $E_2 := P'(\neg\mathrm{C})$, the (absolute) endorsement of AC as $E_3 := P'(\mathrm{A})$, and the (absolute) endorsement of MT as $E_4 := P'(\neg\mathrm{A})$. Furthermore, we define the relative endorsement of inferences i and j as $\Delta_{ij} := E_i - E_j$ with $i<j$. These quantities will be calculated in the next section.

It is worth pausing here to note that these endorsement quantities should be conceptually distinguished from the corresponding acceptance rates discussed in the introduction. While the former are interpreted as representations of the extent to which a single idealised Bayesian agent will endorse an inference in a probabilistic reasoning task, the latter represent the relative frequency with which those inferences are accepted at the population level. There is no a-priori reason to expect a close correspondence between these two different quantities. In what follows, we try to explain the MP-MT asymmetry in terms of individual endorsement rates.

The Results

Our formal results can be summarized in the following two propositions (all proofs are in the Appendix):

Proposition 1 *An agent considers the binary propositional variables A and C with a probability distribution P defined over them. She then learns (i) the major premise of an argument and sets $P'(\mathrm{C}|\mathrm{A})=1$ and (ii) the minor premise and sets its new probability to a value according to eqs. (4) with $\lambda\in(0,1]$. To find the full new probability distribution P', we minimize the KL-divergence between P' and P. The (absolute) endorsements are then given by $E_1=\lambda+\overline{\lambda}P(\mathrm{A}\vee\mathrm{C})$, $E_2=\lambda P(\neg\mathrm{C}|\neg\mathrm{A})+\overline{\lambda}P(\neg\mathrm{A},\neg\mathrm{C})$, $E_3=\lambda P(\mathrm{A}|\mathrm{C})+\overline{\lambda}P(\mathrm{A},\mathrm{C})$ and $E_4=\lambda+\overline{\lambda}P(\neg\mathrm{A}\vee\neg\mathrm{C})$.*

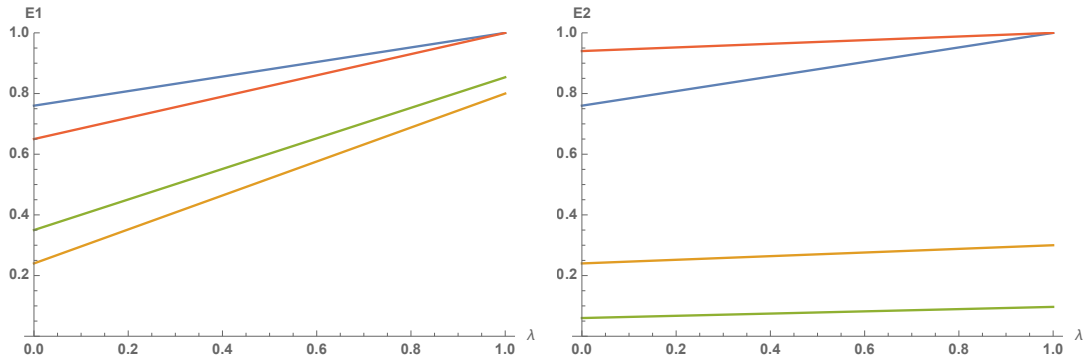
Proposition 2 *Proposition 1 implies the following statements: (i)* MP > AC. *(ii)* MT > DA. *(iii) If $P(\mathrm{A})\geq 1/2$, then* MP > DA *(iv) If $P(\mathrm{A},\mathrm{C})\geq P(\neg\mathrm{A},\neg\mathrm{C})$, then* MP > MT, AC > DA *and $E_1+E_2<E_3+E_4$. (v) If $P(\mathrm{A},\mathrm{C})\leq 1/2$, then* MT > AC. *(vi)* MP > MT *iff* AC > DA. *(vii) If $P(\mathrm{A}\vee\mathrm{C})\geq 1/2$, then* MP > DA.

Here we have used the notation MP > AC for $\Delta_{13}>0$ etc. Note that the assumptions stated in the various if-sentences in Proposition 2 are only sufficient conditions. It turns out that the respective consequents also hold in a large range of other contexts. These depend, however, on the value of both P and λ as shown in Figure 2.

The two left panels of Figure 2, panels (a) and (c), show a situation in which the probability of the antecedent is relatively high (i.e., large a), the conditional expresses a relationship with reasonable confidence (i.e., the conditional probability of the consequent given the antecedent, p, is at least .5), and exceptions are somewhat uncommon (i.e., relatively low conditional probability of the consequent given that the antecedent, q, does not hold). In this situation we see the typical MP-MT asymmetry pattern (as long as $\lambda<1$), when comparing the blue (MP) and red (MT) line. We also see that the degree of the MP-MT asymmetry crucially depends on λ and increases with decreasing λ. Furthermore, the degree of the MP-MT asymmetry also depends on the specific parameters of P. If the conditional expresses a more certain relationship, as in panel (c), the MP-MT asymmetry is larger than if the relationship expressed by the conditional is more uncertain, as in panel (a).

An interesting pattern is observed if the prior probability of the antecedent is low (i.e., $a<.5$), as shown in panels (b) and (d). We can see that in this case the sign of the MP-MT asymmetry flips. Now, we expect stronger endorsement to MT than to MP. However, as for the case in which the prior probability of the antecedent is relatively large, we see that the extent of this reversed asymmetry also depends on λ and the other parameters of P.

Figure 2 also shows the predicted endorsement for the other two inferences, AC and DA. Their ordering (i.e., whether endorsement is expected to be larger for AC or DA) follows the same general pattern also observed for MP and MT. For panels (a) and (c) we expect larger endorsement for AC and DA (as is commonly observed in the literature). However, if the prior probability of the antecedent is low, we expect the same flip; larger endorsement for DA than for AC. In addition, the figure shows another interesting empirical prediction. For certain values of P, see panel (c), we expect either AC > MT or MT > AC, depending on the value of λ



(a) $a = .7, p = .5$, and $q = .2$

(b) $a = .2, p = .3$, and $q = .7$

(c) $a = .8, p = .7$, and $q = .3$

(d) $a = .2, p = .7$, and $q = .3$

Figure 2: The absolute endorsements E_1 (MP, blue), E_2 (DA, orange), E_3 (AC, green) and E_4 (MT, red) as a function of λ for different prior probability distributions P.

(the qualitatively similar also holds between MP and DA, see panel (d)).

Discussion

These results show that the MP-MT asymmetry is predicted by the behavior of a rational agent who updates her belief after encountering new information that is part of the premises of a conditional inference under certain conditions. In contrast to previous probabilistic accounts (Oaksford & Chater, 2008, 2013), we do not need to assume two different probability distributions for MP and the other inferences. Instead, we describe a rational account of how agents update their beliefs in light of new information and use this updated probability distribution as the basis for her endorsement to the four conditional inferences. With this model, we can also make specific predictions when we would expect the opposite pattern, a MT-MP asymmetry.

Disabling Conditions

So far we have assumed that the agent only considers two propositions, i.e. A and C. In many cases, however, there are other relevant propositions and learning new information might affect them. This might have implications for the endorsement of the various inference patterns we have discussed. Consider the following case (Oaksford & Chater, 2008): Let A be the proposition "you turn the key of your car" and let C be the proposition "the car starts". You then learn the premises of a MT inference, i.e. $\mathrm{A} \rightarrow \mathrm{C}$ and $\neg\mathrm{C}$. In that case it seems reasonable to not infer $\neg\mathrm{C}$, but rather that the car is broken or, more generally, that a disabler is present (D). To model this situation, we consider the Bayesian Network in Figure 3 and assume that

$$P(\mathrm{A}) = a \quad , \quad P(\mathrm{D}) = d, \tag{5}$$

where a is large (you will be pretty certain that you turned the key of your car if you did so) and d is somewhat smaller, but it seems reasonable to take the possibility that the car might be broken into account before actually turning the key of the car.

Furthermore, we have to specify the likelihoods

$$\begin{aligned} P(\mathrm{C}|\mathrm{A},\mathrm{D}) &= \alpha \quad , \quad P(\mathrm{C}|\mathrm{A},\neg\mathrm{D}) = \beta \\ P(\mathrm{C}|\neg\mathrm{A},\mathrm{D}) &= 0 \quad , \quad P(\mathrm{C}|\neg\mathrm{A},\neg\mathrm{D}) = 0. \end{aligned} \tag{6}$$

Here we have assumed that the car does not start if the key is not turned. Note that the context suggests that $\beta > \alpha \approx 0$ although we will not need the left inequality. All we will need is that α is fairly small.

The agent then learns the conditional $\mathrm{A} \rightarrow \mathrm{C}$ which imposes the constraint $\beta' > \beta$ on P'.[2] The agent furthermore

[2] β' could be 1, but we will see that this does not matter. All we need is that $\beta, \beta' > 0$.

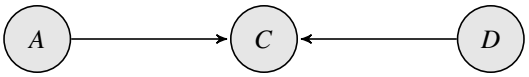


Figure 3: The Bayesian Network representation of the relation between A, C and D.

learns that $P'(\mathrm{C}) = 0$. (We set this value to 0 as it will be hard to doubt that the car did not start if in fact it did not start.) We then note that

$$P'(\mathrm{C}) = a'\,(d'\,\alpha' + \overline{d'}\,\beta') = 0, \tag{7}$$

where we have assumed that P' can be parameterized analogously to P. Given that $\beta' > \beta > 0$, eq. (7) has two solutions: (i) $a' = 0$ and (ii) $\alpha' = 0$ and $d' = 1$. Obviously, solution (i) corresponds to the proper MT inference. However, this inference is implausible in the present case. To explore this issue further, let us consider the KL-divergence between P' and P:

$$KL = \Phi_a + \Phi_d + a'\,d'\,\Phi_\alpha + a'\,\overline{d'}\,\Phi_\beta \tag{8}$$

We have to minimize KL with the constraint (7). Let us consider solution (i) first. Then $KL_1 = -\log\overline{a} + \Phi_d$. This expression minimizes for $d' = d$ and therefore $KL_1^{min} = -\log\overline{a}$. Next, consider (ii). Here $KL_2 = \Phi_a - a'\log\overline{\alpha} - \log d$. Minimizing this expression with respect to a' yields $a' = a\,\overline{\alpha}/(a\,\overline{\alpha} + \overline{a})$ and $KL_2^{min} = -\log((a\,\overline{\alpha} + \overline{a})\,d)$. Hence, $KL_2^{min} < KL_1^{min}$ iff $(a\,\overline{\alpha} + \overline{a})\,d > \overline{a}$ or $a\,\overline{\alpha} > \overline{a}\,\overline{d}$. This condition is fulfilled in the present case as $\alpha \approx 0$ and $a \approx 1$. (Note that the value of d does not matter too much here, but it should not be too low. If it is very low and the inequality is violated, then the agent should make a MT inference and infer $\neg$A.)

Conclusions

Our main goal was to provide a novel probabilistic explanation for the MP-MT asymmetry found in both the traditional abstract task as well as in probabilistic tasks with conditional inferences. In contrast to previous explanations within a probabilistic framework (Oaksford & Chater, 2007, 2013, 2008), our explanation is based on a principled approach of how agents update a probability distribution P in light of new information provided by the premises of a conditional inference resulting in a updated probability distribution P'. Following the idea that "argumentation is learning" (Eva & Hartmann, 2018), we propose that agents update their probability distributions in light of new information by minimizing the KL-divergence between the posterior and prior probability distribution. In this conceptualization, reasoning does not only amount to a read-out from memory, but requires the agent to actively integrate the new knowledge with the existing one. The exact cognitive processes how this is achieved (e.g., by creating new memory traces or overwriting existing ones), is an open question for future work. Our work provides a full *computational-level account* (in the sense of Marr, 1982) of conditional reasoning.

The theoretical results presented here provide evidence that the MP-MT asymmetry is a direct consequence from this Bayesian conceptualization of conditional reasoning. Specifically, it occurs if the prior probability of the conditional probability of C given A (i.e., the relationship expressed in the conditional) and the the prior probability of the antecedent is at least .5. In the case that these conditions do not hold, we expect the opposite pattern, an inverted MP-MT asymmetry.

Minimizing the KL-divergence, as proposed here, is one rational way for an agent to update her prior probability distribution in light of new information which implies Jeffrey conditionalization (Diaconis & Zabell, 1982). Importantly, the results shown here do not only apply to updating via minimizing the KL-divergence, but for updating based on minimizing the distance between P' and P for any divergence measure that is a member of the family of f-divergences. All these divergence metrics are rational in the same sense and also predict the MP-MT asymmetry under the same circumstances. This is an important aspect of our results in light of the findings of Singmann, Klauer, and Beller (2016). They have investigated the empirical adequacy of conditional reasoning based on KL-minimization between P' and P in a two-step conditional reasoning task – which allowed to obtain estimates of both P and P' in an independent manner – and found that it did not provide a very adequate account. However, as soon one is willing to give up the assumption that $P'(\mathrm{C}|\mathrm{A}) = p' = 1$ and assumes that $P'(\mathrm{C}|\mathrm{A}) < 1$ (as done in Singmann et al.'s study), different members of the family of f-divergences make different predictions. Preliminary work suggests that a more empirically adequate account of conditional reasoning is provided if we assume reasoners update their probability distribution by minimizing the inverse-KL divergence between prior and posterior distribution.

Proof of Proposition 1

We use the parameterization of the prior probability distribution P according to eqs. (1) and (2) and begin with MP and DA. Here we set the new value of the probability of the antecedent to a'. Disregarding constant terms, the KL-divergence is then given by $KL = \overline{a'}\,\Phi_q$ with

$$\Phi_x := x' \log\frac{x'}{x} + \overline{x'} \log\frac{\overline{x'}}{\overline{x}}\,. \tag{9}$$

Differentiating KL with respect to q' and setting the resulting expression equal to zero yields $q' = q$. Hence, $P'(\mathrm{C}) = a' + \overline{a'}\,q$. We now insert the appropriate values of c' from eqs. (4) and use the definition of the respective (absolute) endorsements to obtain

$$\begin{aligned} E_1 &:= P'(\mathrm{C}) \\ &= \lambda + \overline{\lambda}(a + \overline{a}\,q) \\ &= \lambda + \overline{\lambda} P(\mathrm{A} \vee \mathrm{C}) \\ E_2 &:= P'(\neg\mathrm{C}) \\ &= \lambda\,\overline{q} + \overline{\lambda}\,\overline{a}\,\overline{q} \\ &= \lambda P(\neg\mathrm{C}|\neg\mathrm{A}) + \overline{\lambda} P(\neg\mathrm{A}, \neg\mathrm{C}). \end{aligned}$$

Let us now consider AC and MT. In this case, learning the minor premise amounts to the constraint

$$a' + \overline{a'}\, q' = c', \tag{10}$$

with c' specified in eqs. (4). We therefore have to minimize the function

$$L = \Phi_a + a' \log \frac{1}{p} + \overline{a'}\, \Phi_q + \mu(a' + \overline{a'}\, q' - c'),$$

with the Lagrange multiplier μ.

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

$$q' = \frac{1}{q + \overline{q}x}, \tag{11}$$

with $x := \exp(\lambda)$. Hence,

$$L = \Phi_a + a' \log \frac{1}{p} + \overline{a'} \log \frac{1}{q + \overline{q}x} + \mu \overline{c'}.$$

Differentiating this expression with respect to a' and setting the resulting expression equal to zero yields

$$a' = \frac{a\,p}{a\,p + \overline{a}\,(q + \overline{q}x)}. \tag{12}$$

From eqs. (10), (11) and (12), we then obtain

$$a' = \frac{a\,p\,c'}{a\,p + \overline{a}\,q}. \tag{13}$$

We now insert the appropriate values of c' from eqs. (4) in eq. (13) and use the definitions of the respective (absolute) endorsements to obtain

$$\begin{aligned}
E_3 &:= P'(\mathrm{A}) \\
&= \lambda \frac{a\,p}{a\,p + \overline{a}\,q} + \overline{\lambda}\, a\,p \\
&= \lambda P(\mathrm{A}|\mathrm{C}) + \overline{\lambda} P(\mathrm{A},\mathrm{C}) \\
E_4 &:= P'(\neg \mathrm{A}) \\
&= \lambda + \overline{\lambda}(1 - a\,p) \\
&= \lambda + \overline{\lambda} P(\neg \mathrm{A} \vee \neg \mathrm{C}).
\end{aligned}$$

This completes the proof of Proposition 1. ∎

Proof of Proposition 2

We use Proposition 1 to compute the relative endorsements:

$$\begin{aligned}
\Delta_{12} &= \lambda q + \overline{\lambda}\,[2(a + \overline{a}\,q) - 1] \\
\Delta_{13} &= \lambda \frac{\overline{a}\,q}{a\,p + \overline{a}\,q} + \overline{\lambda}(a\,\overline{p} + \overline{a}\,q) \\
\Delta_{14} &= \overline{\lambda}(a\,p - \overline{a}\,\overline{q}) \\
\Delta_{23} &= -(a\,p - \overline{a}\,\overline{q}) \cdot \left[\lambda \frac{q}{a\,p + \overline{a}\,q} + \overline{\lambda}\right] \\
\Delta_{24} &= -\lambda q - \overline{\lambda}(a\,\overline{p} + \overline{a}\,q) \\
\Delta_{34} &= -\lambda \frac{\overline{a}\,q}{a\,p + \overline{a}\,q} + \overline{\lambda}(2a\,p - 1)
\end{aligned}$$

From these results, the statements made in the proposition follow. For example, the third statement in (iv) follows by noting that $\Delta_{14} + \Delta_{23} < 0$. ∎

References

Costello, F., & Watts, P. (2014). Surprisingly rational: Probability theory plus noise explains biases in judgment. *Psychological Review*, *121*(3), 463–480. doi: 10.1037/a0037010

Diaconis, P., & Zabell, S. L. (1982). Updating Subjective Probability. *Journal of the American Statistical Association*, *77*(380), 822–830. doi: 10.1080/01621459.1982.10477893

Eva, B., & Hartmann, S. (2018). Bayesian Argumentation and the Value of Logical Validity. *Psychological Review*, *125*(5), 806–821. doi: 10.1037/rev0000114

Evans, J. S. B. T., Thompson, V. A., & Over, D. E. (2015). Uncertain deduction and conditional reasoning. *Cognition*, *6*, 398. doi: 10.3389/fpsyg.2015.00398

Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. San Francisco: W.H. Freeman.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: the probabilistic approach to human reasoning*. Oxford; New York: Oxford University Press.

Oaksford, M., & Chater, N. (2008). Probability logic and the Modus Ponens – Modus Tollens asymmetry in conditional inference. In N. Chater & M. Oaksford (Eds.), *The Probabilistic Mind:: Prospects for Bayesian cognitive science* (pp. 97–120). Oxford University Press.

Oaksford, M., & Chater, N. (2013). Dynamic inference and everyday conditional reasoning in the new paradigm. *Thinking & Reasoning*, *19*(3-4), 346–379. doi: 10.1080/13546783.2013.808163

Oaksford, M., Chater, N., & Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(4), 883–899. doi: 10.1037/0278-7393.26.4.883

Over, D. E. (2009). New paradigm psychology of reasoning. *Thinking & Reasoning*, *15*(4), 431–438. doi: 10.1080/13546780903266188

Schroyens, W. J., Schaeken, W., & d'Ydewalle, G. (2001). The processing of negations in conditional reasoning: A meta-analytic case study in mental model and/or mental logic theory. *Thinking & Reasoning*, *7*(2), 121–172. doi: 10.1080/13546780042000091

Singmann, H., Klauer, K. C., & Beller, S. (2016). Probabilistic conditional reasoning: Disentangling form and content with the dual-source model. *Cognitive Psychology*, *88*, 61–87. doi: 10.1016/j.cogpsych.2016.06.005

Singmann, H., Klauer, K. C., & Over, D. E. (2014). New normative standards of conditional reasoning and the dual-source model. *Frontiers in Psychology*, *5*, 316. doi: 10.3389/fpsyg.2014.00316