FEW-2004

Commentary on Michael Strevens' paper: "The Wrong Problem: Relevance and Irrelevance in Bayesian Confirmation Theory"

Luca Moretti

Philosophy, Probability, and Modeling Research Group

Center for Junior Research Fellows

University of Konstanz

P.O. Box M 682

D-78457 Konstanz, Germany

E-mail: Luca.Moretti@uni-konstanz.de

Phone: 0049 (0)7531 88 4736

Fax: 0049 (0)7531 88 4898

1. Introduction

In his acute and interesting paper, Michael Strevens argues that there are at least two *prima facie* problems that the existence of irrelevant conjunctions raises to Bayesian methodology. He however thinks that just one of the two problems is the serious one.

Consider a hypothesis *H*, an evidential statement *E*, and a statement *J* intuitively irrelevant to *H* and *E*. The first apparent problem is that, in many cases, if *E* confirms *H*, *E* also confirms the *conjunction H* & *J*. (For the sake of simplicity, I will generally omit referring to background knowledge in my considerations about relevance and confirmation, but I will always refer to it implicitly). Strevens briefly argues that this is no problem at all. The reason is one already put forward by Maher 2004. If confirmation is a *probabilistic* relationship, *A* confirms *B* if *A* raises the probability of *B*, and – as Strevens emphasises – "it is quite normal for evidence that raises the probability of a part to thereby raise the probability of the whole" (Strevens 2004, p. 10). Thus it is quite normal for a piece of evidence *E*, which confirms *H*, to confirm *H* & *J*, even though *J* is in some sense irrelevant to *H* and *E*. I believe that Strevens is right in this case.

The second problem of irrelevant conjunction for Bayesian methodology is that we have apparently no guarantee that, if *E* confirms *H*, and then *E* confirms *H* & *J*, confirmation cannot be "transmitted" to the irrelevant *J*. For we have no guarantee that both Hempel's *Converse Consequence Condition* and Hempel's *Special Consequence Condition* might apply *locally* to one and the same irrelevant conjunction to generate the above paradox.[1] If the confirmation of an irrelevant conjunction were "transmitted" to its irrelevant conjunct, Bayesian confirmation theory would surely collapse.

Strevens calls this second problem the "real problem" of irrelevant conjunction and he argues that it cannot be settled by Bayesian means. The reason would be that, on the one hand, a Bayesian solution should provide a genuine *explanation* of why *E* does not confirm

[1] I say more on these *Conditions* below.

the irrelevant J. Consequently, the solution cannot consist in defining the irrelevance of J by *assuming* – trivially – that Pr ($J|E$) = Pr (J) (i.e. that E and J are probabilistically independent). Yet, on the other hand – according to Strevens – Bayesian methodology cannot provide any set of intuitive conditions for the irrelevance of J that entail that Pr ($J|E$) = Pr (J) without already including this very equation among the members of the set, and thus without assuming it.

Although Strevens uses the expressions "the real problem" and "the serious problem", I believe that the problem he focuses on is not *that* real or *that* serious for Bayesian methodology. For it can be settled in each case by imposing the trivial condition that Pr ($J|E$) = Pr (J). Bayesianism is thus far from falling afoul of Strevens' real problem of irrelevant conjunction. This however does not make Strevens' paper uninteresting, as the question whether Bayesian methodology can provide a *non-trivial* solution of the problem is – I believe – interesting and worth investigating.

Against Strevens' opinion, I will now show that there are general, intuitive and *non-trivial* Bayesian conditions that are sufficient for excluding the confirmation of irrelevant conjuncts. Before doing it, let me address two minor objections.

2. Two minor objections

To support the claim that, on Bayesian methodology, E often confirms H & J when E confirms H, Strevens affirms that:

> Increasing the probability of a conjunct will increase the probability of the conjunction, provided that neither E nor H is negatively irrelevant to J. (p. 2).

This is false. Consider the conjunction ($E \vee J$) & J, such that Pr ($J|E$) = Pr (J), Pr (E) > 0 and Pr ($E \vee J$) < 1. In this case, the conjunct H is equivalent to $E \vee J$, and neither E nor H is negatively relevant to J. E is not negatively irrelevant to J because E and J are assumed to be probabilistically independent; H is not negatively relevant to J because the latter entails the former. Besides, since E entails H (i.e., $E \vee J$), E does confirm H, but E cannot confirm the whole conjunction $(E \vee J)$ & J, as the latter is logically equivalent to J, which is, by assumption, probabilistically independent of E.

Nevertheless, the failure of the conditions fixed by Strevens does not prevent E from confirming H & J when E confirms H. For there are many cases in which this happens. Consider for instance that, if H entails E and E confirms H, E will always confirm H & J, provided that Pr $(H \,\&\, J) \neq 0$ and 1.

My second minor criticism is the following: according to Strevens, the real problem of the irrelevant conjunction depends on the simultaneous application of the *Converse Consequence Condition* and the *Special Consequence Condition*, two principles that have only a local validity in Bayesian methodology. Strevens contends that: “The principal victim of irrelevant conjunction is, of course, Hypothetico-deductivism” (p. 3). The reason would be that, if H is a hypothesis and J an irrelevant conjunct, the simultaneous use of the two Hempelian conditions make any evidence E that confirms H confirm any irrelevant J through the confirmation of the conjunctive “hypothesis” E & J. (Suppose E confirms H. Since H & J entails H, if the *Converse Consequence Condition* holds, E confirms H & J. But H & J entails J too. Thus, if also the *Special Consequence Condition* holds, since E confirms H & J, E confirms J). This would be a *reductio* of Hypothetico-deductivism.

Indeed, Hypothetico-deductivism is immune to this *reductio* because, notoriously, it only entails the *Converse Consequence Condition* but not the *Special Consequence Condition*. As a result, no irrelevant conjunct can be confirmed by evidence unless the former entails the latter (but in such a case the conjunct is not irrelevant to the evidence!). In conclusion, Hypothetico-deductivism is *not* a victim of the irrelevant conjunction problem, as Strevens contends.

But Strevens' contention has the form of a conditional: he claims – more precisely – that the *reductio* obtains "*If* the *Special Consequence Condition* is true" (p. 3). But, again, this seems to me incorrect. If the *Special Consequence Condition* is true, the problem is that Hypothetico-deductivism is not a fully adequate theory of confirmation because it does not satisfy that adequacy condition. This does not mean that Hypothetico-deductivism allows irrelevant conjuncts to be confirmed by evidence: hypothetico-deductive confirmation does not permit it in any case.

3. A non-trivial Bayesian solution of the real problem of irrelevant conjunction

Let us now turn to the main objection. According to Strevens, the Bayesian solution of the real problem of irrelevant conjunction should consist of two parts:

> First, a probabilistic definition of what it is for a hypothesis J to be irrelevant relative to a piece of evidence E and another hypothesis H, and second, a demonstration that when E confirms H, and so confirms H & J, it nevertheless does not confirm J. (p. 11).

Why, in Strevens' opinion, cannot there be any Bayesian solution of the real problem of irrelevant conjunction? Let us proceed gradually. First, Strevens indicates that Fitelson's definition of irrelevance is unhelpful because it just assumes, among other conditions, that $\Pr(J|E) = \Pr(J)$. This is correct, as Fitelson irrelevance states that J is irrelevant to H and E if and only if J and E are probabilistically independent, and the same happens with, respectively, J and H, and with J and H & E. (Cf. Fitelson 2002).

Second, Strevens argues that Hawthorne's (and Fitelson's) more recent definition of irrelevance is also unhelpful. Strevens is right in this case too. However, one of the reasons he adduces is incorrect. Let us see why. A statement J is Hawthorne irrelevant to the confirmation of a hypothesis H by evidence E if and only if (cf. Fitelson and Hawthorne 2004):

(1) $\Pr(E|H \,\&\, J) = \Pr(E|H)$.

Fitelson and Hawthorne emphasise that, in the deductive case, a hypothesis is tested by what it says about evidence. It is thus intuitive that J is irrelevant to the confirmation of a hypothesis H by evidence E if and only if "adding J to H … says nothing more about E than H … alone already says" (Fitelson and Hawthorne 2004, p. 7). Notice that the thesis is that J is irrelevant if and only if the *conjunction* J & H says nothing more about E than H already says. Fitelson and Hawthorne think of (1) as a generalisation of *this* intuitive but quite precise notion of deductive irrelevance. Strevens, however, attributes to Fitelson and Hawthorne the thesis that J is irrelevant when "J says nothing more about E than H already says" (Strevens 2004, p. 15).

According to Strevens, appealing to Hawthorne irrelevance does not solve the real problem of irrelevant conjunct for two reasons. The first is that the satisfaction of (1) is not sufficient to prevent that $\Pr(J|E) > \Pr(J)$, so that E can confirm J. This is true, numerical examples can be given. Such examples are indeed unnecessary because we can obtain the same result even in the deductive case: the situation in which adding J to H says nothing more about E than H alone already says is compatible with the possibility that both H alone and J alone entail E. In this case, E does confirm J.[2]

Strevens' second reason why Hawthorne irrelevance is unhelpful is however mistaken. Strevens contends that, even in case Hawthorne irrelevance is satisfied, J could have "quite a lot to say ... about E". For – if I understand Strevens – $\Pr(E|J)$ could be very different from $\Pr(E|H \,\&\, J)$ and then from $\Pr(E|H)$. This, according to Strevens, means that: "the Hawthorne irrelevance of J relative to H and E fails to capture just those cases in which J has

[2] This is no unwanted consequence of Hawthorne's definition irrelevance, as this definition is *not* meant to resolve the real problem of irrelevant conjunction.

nothing to say about *E* beyond what is said by *H*' (p. 16). But the criticism is misplaced. For Hawthorne's specific notion of irrelevance does not focus on what *J alone* says about *E*.

The third step of Strevens' reasoning consists in generalising the negative results just attained to any possible attempt to define irrelevance in Bayesian terms in a non-trivial way. I quote Stevens' words:

> The problem is not with Hawthorne irrelevance in particular, but with probabilistic definitions of irrelevance in general. Such definitions take the form of one or more probabilistic independence claims. There are always two ways that such claims can be satisfied. First, the independent propositions can, in the intuitive sense, genuinely have nothing to do with one another. But second, they may have quite a lot to do with one another, both positively and negatively, but the impact may cancel out. It is because of this second kind of situation that [a] a probabilistic definition of irrelevance will never capture one of our intuitive notions of relevance, and also, that [b] a probabilistic definition of irrelevance will never imply the independence of two propositions unless the fact of their independence is itself a part of the definition. (p. 17).

Not all parts of the above passage are completely clear to me. Nevertheless, it seems that both the conclusions [a] and [b] – precisely formulated by Strevens – can be challenged by the following example.

Let us take a hypothesis *J* to be *irrelevant* to another hypothesis *H* and a piece of evidence *E*, if and only if:

(1) $\Pr(E|H \,\&\, J) = \Pr(E|H)$

(2) $\Pr(H|E \,\&\, J) = \Pr(H|E)$

(3) $\Pr(H|J) = \Pr(H)$.

These three conditions are logically independent of one another. Condition (1) expresses Hawthorne irrelevance, which does have intuitive appeal. Condition (2) accounts for the situation in which adding J to E says nothing more about H than E alone already says. (2) appears to capture a further sense in which J is irrelevant to H and E. Finally, if J and H are irrelevant to one another, (3) must very plausibly hold (notice that this condition is also part of Fitelson irrelevance). Thus, in opposition to Strevens' conclusion [a], there seem to be a Bayesian definition that *prima facie* captures an intuitive notion of irrelevance: precisely, the one given in terms of (1)–(3).

Further conditions could perhaps be added to make this framework fit our intuitions better. What is important is that conditions (1)–(3) appear *minimal* for the solution of the problem at stake. For, against Strevens' conclusion [b], the intuitive definition of irrelevance given in terms (1)–(3) does not trivially *assume* the probabilistic independence of E and J but it does entail the latter. Let us see how.

By Bayes' theorem, we obtain:

$$(4) \quad P(E/H) = \frac{P(E) \times P(H/E)}{P(H)}$$

and

$$(5) \quad P(E/H \,\&\, J) = \frac{P(E/J) \times P(H/E \,\&\, J)}{P(H/J)}.$$

(Let us assume that Pr (H) and Pr ($H|J$) $\neq 0$).

Given (1), we can identify the right-hand side of (4) with the right-hand side of (5). Thus:

$$(6) \quad \frac{P(E) \times P(H/E)}{P(H)} = \frac{P(E/J) \times P(H/E \,\&\, J)}{P(H/J)}.$$

Given (2), we can replace Pr ($H|E$ & J), in the right-hand side of (6), by Pr ($H|E$). It follows:

$$(7) \quad \frac{P(E) \times P(H/E)}{P(H)} = \frac{P(E/J) \times P(H/E)}{P(H/J)}.$$

Given (3), we can replace Pr ($H|J$), in right-hand side of (7), by Pr (H). We obtain:

$$(8)\ \frac{P(E)\times P(H/E)}{P(H)}=\frac{P(E/J)\times P(H/E)}{P(H)}.$$

(8) entails:

(9) Pr (E) = Pr ($E|J$),

which is equivalent to:

(10) Pr ($J|E$) = Pr (J).

In conclusion, if the conditions for irrelevance (1)–(3) hold, evidence *E* cannot confirm the irrelevant conjunct *J*. As Fitelson and Hawthorne have shown, when (1) is satisfied, if *E* confirms *H*, *E* does confirm *H* & *J*. (Cf. Fitelson and Hawthorne 2004). But, whenever all conditions (1)–(3) hold together, *E* will *not* confirm *J*. This would seem to constitute a genuine and non-trivial Bayesian solution of the real problem of the irrelevant conjunction.

As a final point, Strevens argues that confirmational relevance generally depends on background knowledge and can change as background knowledge changes. I think this is true in most cases. The unpleasant consequence would however be that, to establish whether a statement *J* is *really* irrelevant for a hypothesis *H* and evidence *E*:

> You would have to know something about the probabilistic relevance relations between everything and everything else. But if you know all this, then you already know enough to see whether the evidence will confirm the allegedly irrelevant conjunct. There is no interesting local fact about probabilistic relevance that you can leverage, then, to gain knowledge about the confirmation or otherwise of an irrelevant conjunct. (Strevens 2004, p. 22).

Indeed, I do not see why, to establish the relevance or irrelevance of a conjunct, we must know something about the probabilistic relations between everything and everything else. One could suggest that we only have to know the probabilistic relations that make the conditions (1)–(3) true or false. This does not seem to involve knowing in advance whether the evidence confirms the allegedly irrelevant conjunct. Even if confirmational relevance is background-dependent, a Bayesian explanation of irrelevance and a Bayesian solution of the real problem of irrelevant conjunction appear still possible.

References

Fitelson, B. 2002. "Putting the Irrelevance back Into the Problem of the Irrelevant Conjunction". *Philosophy of Science* **69**(4), pp. 611–622.

Fitelson, B. and J. Hawthorne. 2004. "Discussion: Resolving *Irrelevant Conjunction* with Probabilistic Independence". Forthcoming in *Philosophy of science*.

Maher, P. 2004. "Bayesianism and irrelevant conjunction". Forthcoming in *Philosophy of Science*.

Strevens, M. 2004 (draft of the 9th of May 2004). "The Wrong Problem: Relevance and Irrelevance in Bayesian Confirmation Theory".