# Putting The Irrelevance Back Into the Problem of Irrelevant Conjunction & An Easy Way Out of the Problem of Old Evidence


Branden Fitelson

Department of Philosophy
Stanford University

fitelson@stanford.edu
philosophy.wisc.edu/fitelson


November 19, 2001



## Preliminaries I: Some General Bayesian Background

- Bayesianism (*i.e.*, Bayesian *epistemology*) assumes that the degrees of belief (or credence) of rational agents are *probabilities.*
- $\Pr(H)$ denotes the (rational) degree of belief in a proposition $H$. This is called the *un*conditional (or *prior*) probability of $H$.
- $\Pr(H \mid E)$ denotes the (rational) degree of belief in $H$, *given* $E$ (the degree of belief a rational agent would assign to $H$ after learning $E$). This is the *conditional* (or *posterior*) probability of $H$ (on $E$).
- There is much controversy over whether rational degrees of belief really are probabilities, and over the objective status and origins of *prior* probabilities. Presently, all such questions will be bracketed.
- To simplify and focus the discussion, I will assume (*arguendo!*) that all rational agents share a *single* probability function Pr.





## Preliminaries II: Background to Confirmation Theory

- In Bayesian confirmation theory, evidence $E$ *confirms* (or *supports*) a hypothesis $H$ (roughly) if learning $E$ *raises the probability of* $H$.
- If learning $E$ *lowers* the probability of $H$, then $E$ *dis*confirms (or *counter*-supports) $H$, and if learning $E$ does not change the probability of $H$, then $E$ is confirmationally *irrelevant* to $H$.
- Within (Kolmogorov!) probability theory, there are many logically equivalent ways of saying that $E$ confirms $H$. Here are a few:
  - $E$ confirms $H$ if $\Pr(H \mid E) > \Pr(H)$.
  - $E$ confirms $H$ if $\Pr(E \mid H) > \Pr(E \mid \bar{H})$.
  - $E$ confirms $H$ if $\Pr(H \mid E) > \Pr(H \mid \bar{E})$.
- By taking differences, ratios, *etc.*, of the left/right sides of these alternative inequalities, a *plethora* of possible *quantitative measures* of the *degree* to which $E$ confirms (or supports) $H$ can be formed.





## Four Popular and Representative Contemporary Measures

- *Dozens* of such measures have been used in the philosophical literature. Here are four popular contemporary choices:
  - The *Difference*: $d(H,E) =_{df} \Pr(H \mid E) - \Pr(H)$
  - The Log-*Ratio*: $r(H,E) =_{df} \log\left[\frac{\Pr(H \mid E)}{\Pr(H)}\right]$
  - The Log-*Likelihood-Ratio*: $l(H,E) =_{df} \log\left[\frac{\Pr(E \mid H)}{\Pr(E \mid \bar{H})}\right]$
  - The "Normalized" Difference: $s(H,E) =_{df} \Pr(H \mid E) - \Pr(H \mid \bar{E})$
- Logs of the ratios are taken to ensure ($i$) that they are $+/-/0$ when $E$ confirms/disconfirms/is irrelevant to $H$, and ($ii$) that they are *additive* in various ways. The use of logs is just a helpful convention; it does not lead to any loss of generality in our argumentation.
- The first part of our story concerns the disagreement exhibited by these measures, and its ramifications for confirmation theory …

## Disagreement Between Alternative Measures

- What kind of disagreement between measures is important?
- Mere *numerical* (or *conventional* or *syntactical*) differences between measures are not important, since they need not effect *ordinal* judgments of what is more/less well confirmed than what.
- *Ordinal* differences are crucial, since they can effect the cogency of many arguments surrounding Bayesian confirmation theory.
- For instance, it is part of Bayesian lore that the observation of a black raven ($E_1$) confirms the hypothesis ($H$) that all ravens are black more strongly than the observation of a white shoe ($E_2$) does.
- But, given the standard assumptions in Bayesian accounts of Hempel's ravens paradox, this conclusion [$\mathfrak{c}(H, E_1) > \mathfrak{c}(H, E_2)$] follows only for *some* measures of confirmation $\mathfrak{c}$ (and *not* others).
- Such arguments are said to be *sensitive to choice of measure.*





## The Problem of Measure Sensitivity

- A detailed study of the literature shows that *virtually every argument* involving quantitative Bayesian confirmation theory is sensitive to choice of measure in this sense! Some examples:
  - The Popper-Miller Argument *Against* Bayesianism
  - Rosenkrantz and Earman on "Irrelevant Conjunction" (see below)
  - Eells and Sober on Goodman's "Grue" Paradox
  - Horwich *et al.* on Hempel's Ravens Paradox
  - Horwich *et al.* on the Confirmational Value of Varied Evidence
  - Quantitative resolutions of the problem of old evidence (see below)
- One needn't gerrymander or comb the historical literature for Bayesian measures which fail to undergird these arguments.
- Each of these arguments is valid with respect to *only some* of the four most popular contemporary measures $d$, $r$, $l$, and $s$.





## The Traditional Problem of "Irrelevant Conjunction"

- According to the Hypothetico-Deductive (H–D) account of confirmation, $E$ confirms $H$ (roughly) if $H$ entails $E$.
- As a result, H–D-confirmation has the following property:

(1) If $E$ confirms $H$, then $E$ confirms $H \,\&\, X$, for *any* $X$.

- But, the $X$'s in (1) can be *utterly irrelevant* (or, for that matter, $X$ could be highly *negatively* relevant!) to $H$ (and $E$).
- While (1) is *not* a consequence of Bayesian confirmation theory, the following special case (in which $E$ is *deductive evidence* for $H$) *is*:

(2) If $H$ entails $E$, then $E$ confirms $H \,\&\, X$, for *any* $X$.

- Below, I (*i*) discuss two existing Bayesian approaches to "irrelevant conjunction", and (*ii*) provide a new and improved Bayesian way of thinking about and resolving the "irrelevant conjunction" problem.





## The Canonical Bayesian Approach to "Irrelevant Conjunction"

- Earman gives a Bayesian resolution of the traditional problem which uses the following fact about the difference measure.[a]

(3) If $H$ entails $E$, then $d(H \,\&\, X, E) < d(H, E)$.

- What (3) says is that, while "irrelevant conjunctions" $H\&X$'s will be (Bayesian) confirmed by deductive evidence for $H$, such conjunctions will be confirmed *less strongly* than $H$. Closer scrutiny reveals:

(a) The "irrelevance" of $X$ is *irrelevant* to the decrease in support. Indeed, (3) is true for ***all*** $X$ — even *highly relevant* $X$'s!

(b) (3) is not true for $r$. This is often used as an argument *against* $r$.

(c) (3) only applies to cases of *deductive* evidence — arguably, not an important case for *Bayesianism.* So, perhaps this is *not* bad news for $r$. And, we'd like a *general* account of "irrelevant conjunction".

[a] We assume that $\Pr(X) < 1$, else $H \,\&\, X \approx H$ (*i.e.*, $X$ is not really a *conjunct*).

## Rosenkrantz on the Problem of "Irrelevant Conjunction" I

- Rosenkrantz provides a Bayesian resolution of the problem of Irrelevant Conjunction (*a.k.a.*, the Tacking Problem) which trades on:

(4) If $H$ entails $E$, then $d(H \,\&\, X, E) = \Pr(X \mid H) \cdot d(H, E)$.

- Rosenkrantz does try to address *some* of the problems with the canonical account. In particular, he seems sensitive to (a). He says:

> … I hope you will agree that the two extreme positions on this issue are equally unpalatable, (i) that a consequence $E$ of $H$ confirms $H \,\&\, X$ not at all, and (ii) that $E$ confirms $H \,\&\, X$ just as strongly as it confirms $H$ alone. … In general, intuition expects intermediate degrees of confirmation that depend on the degree of compatibility of $H$ with $X$.

- Adopting $\Pr(X \mid H)$ as the measure of "degree of compatibility" of $H$ with $X$, and $d$ as the measure of confirmation yields the kind of result that Rosenkrantz wants: (4). Is this an *improvement* on Earman?





## Rosenkrantz on the Problem of "Irrelevant Conjunction" II

- In a way, Rosenkrantz is *trying* to address (a) here. He seems to be thinking of $\Pr(X \mid H)$ as a kind of measure of "degree of relevance" between $H$ and $X$. This is a *very peculiar* way for a Bayesian to think of "relevance"! Moreover, $\Pr(X \mid H)$ can tell us nothing about the "degree of relevance" between $X$ and $E$ or between $X$ and $H \,\&\, E$.
- Unfortunately, Rosenkrantz's approach is *even worse* than Earman's when it comes to the *measure sensitivity* issue (b). Rosenkrantz's approach works *only* for $d$ and $s$. It does *not* work for $r$ or $l$. This is especially bad, in light of my arguments in favor of $l$ over $d$ and $s$.
- Finally, Rosenkrantz is still only addressing the *deductive* case. So, his account also lacks *inductive generality* in exactly the same way that the canonical approach does. Thus, Rosenkrantz has not addressed (c).
- I suggest that we *re-think* this problem. Let's see how Bayesianism handles *irrelevant* conjunctions, in the general, *inductive* case …





## The *Bayesian* Problem of *Irrelevant* Conjunction I

- First, we need to say what it *means* for a conjunct $X$ to be *irrelevant* in a confirmational context involving hypothesis $H$ and evidence $E$.[a]
**Definition.** *$X$ is an irrelevant conjunct with respect to the pair $\langle H, E\rangle$ iff $X$ is confirmationally irrelevant to $H$, $E$, and $H \,\&\, E$.*
- Do we even *have* a *Bayesian* problem of *irrelevant* conjunction? Yes.
**Theorem 1.** *If $E$ confirms $H$, and $X$ is an irrelevant conjunct with respect to the pair $\langle H, E\rangle$, then $E$ also confirms $H \,\&\, X$.*
- Can we say anything general and interesting concerning *deleterious effects* of tacking irrelevant conjuncts onto hypotheses $H$? Yes.
**Theorem 2.** *If $E$ confirms $H$, and $X$ is an irrelevant conjunct with respect to $\langle H, E\rangle$, then $\mathfrak{c}(H \,\&\, X, E) < \mathfrak{c}(H, E)$, where $\mathfrak{c}$ may be **any** measure of degree of confirmation, except $r$ $[r(H \,\&\, X, E) = r(H, E)]$.*

[a] This is a *stronger* definition of irrelevant conjunct than we will actually need for our present purposes. All we will *need* is $X \perp H$ and $X \perp H \,\&\, E$. We will *not* need $X \perp E$.





## The *Bayesian* Problem of *Irrelevant* Conjunction II

Here are some advantages of our new analysis of this problem:

- Our analysis *makes use of the irrelevance of $X$*. Moreover, our notion of a confirmationally irrelevant conjunct is not some peculiar one (like Rosenkrantz's), but just the standard Bayesian notion of *independence*.
- Our resolution is not restricted to the (*in*ductively *un*interesting) case of *deductive evidence*; it explains why *irrelevant* conjuncts are confirmationally disadvantageous in *all* contexts (deductive or not).
- Our resolution is as robust (*i.e.*, measure-*in*sensitive) as any other existing resolution (*e.g.*, Earman's), and more robust than any other existing account that tries to be sensitive to the "irrelevance" of the conjunct $X$ (*e.g.*, Rosenkrantz's). Our account is based on a result which holds for *all* Bayesian measures of confirmation currently being used or defended, *except* the ratio measure $r$. This *is bad news* for $r$.

## An Easy Way Out of the Problem of Old Evidence I

- The problem of old evidence (and new theories) derives from the assumption that Bayesians learn by strict conditionalization (and that the probability function Bayesians should use in *all* judgments is their own, current, most well-informed degree of belief function).
- So, if $\Pr_t^a(H)$ is a Bayesian agent $a$'s degree of belief in $H$ at time $t$, then if $a$ learns $E$ at $t$, then for all $t' > t$, $\Pr_{t'}^a(E) = 1$, and $E$ is therefore powerless to change $a$'s degree of belief in $H$ from $t'$ on.
- In *Bayes or Bust?* Earman considers, and quickly dismisses, a possible "easy way out" of the problem of old evidence which rejects the assumption that evidence is learned by strict conditionalization (and so insists that $\Pr_{t'}^a(E) < 1$). He says that this just replaces the *qualitative* problem of old evidence with a *quantitative* problem of old evidence.
- It is instructive to look carefully at what Earman claims here …





## An Easy Way Out of the Problem of Old Evidence II

> … denying that $\Pr_{t'}^a(E) = 1$ only serves to trade one version of the old-evidence problem for another. Perhaps it is not certain in November 1915 that the true value of the anomalous advance (of the perihelion of Mercury) was roughly 43″ of arc per century, but most members of the scientific community were pretty darn sure, e.g., $\Pr_{1915}^a(E) = .999$. Assuming that Einstein's theory does entail $E$, we find that the confirmatory power $\mathfrak{c}(H, E)$ of $E$ is … less than .0001002. This is counterintuitive …

- A careful reconstruction of Earman's argument reveals that it trades on the following property of the difference measure $d$:

(5) If $H$ entails $E$, then $\Pr(E) \approx 1$ implies $d(H, E) \approx 0$.

- This argument has two flaws. First, (5) does hold for $d$ and $r$, but it does *not* hold for $l$ or $s$ (contrary to what Earman says in a footnote). Second, this argument only applies to the case of *deductive evidence*.





## An Easy Way Out of the Problem of Old Evidence III

- We can fix the second problem, since the following is also a theorem:

(6) If $\Pr(E) \approx 1$, then $d(H, E) \approx 0$ and $r(H, E) \approx 0$.

- However, the first problem cannot be fixed, because:

(7) *Even if* $H$ entails $E$ and $\Pr(E) \approx 1$, $l(H, E)$ can be *arbitrarily large*.

- As Earman himself admits, this is not a "crazy" way out, since:

> There are both historical and philosophical reasons for such a stance … (regarding his Mercury example) … the literature of the period contained everything from 41″ to 45″ of arc per century as the value of the anomalous advance … and even the weaker proposition that the value lies somewhere in this range was challenged by some astronomers. … Bayesians are hardly at a loss here, since Jeffrey has proposed a replacement for strict conditionalization which allows for uncertain learning.

- Indeed, this is precisely my idea for an "easy way out" …





## An Easy Way Out of the Problem of Old Evidence IV

- Earman's rejection seems premature. A more charitable reading of the proposed resolution seems only to depend on the following two claims:

(8) A rational Bayesain agent should not assign probability 1 to contingent empirical propositions (*e.g.*, evidential propositions $E$).

(9) The (log) likelihood-ratio measure $l$ is an adequate (preferable) way for rational Bayesian agents to gauge confirmational power.

- Claim (8) seems quite reasonable. Indeed, this is the cornerstone of Jeffrey's "radical probabilism". Others, including Carnap, have also endorsed principles consistent with (8) (*e.g.*, *regularity* of Pr).
- Claim (9) is non-trivial. In my dissertation, I have argued that $l$ should be preferred to $d$ (Earman's favorite measure). I.J. Good does too …
- Earman's reaction to this: "I like it! It is neat and simple and avoids the contortions required by many of the proposed solutions."