On The Evidential Import of Unification

Wayne C. Myrvold
Department of Philosophy
The University of Western Ontario
wmyrvold@uwo.ca

Forthcoming in Philosophy of Science.

June 28, 2016

Abstract

This paper discusses the evidential import of two senses in which a hypothesis may be said to unify evidence. One is the ability of the hypothesis to increase the mutual information of a set of evidence statements; the other is the ability of the hypothesis to explain commonalities in observed phenomena by positing a common origin for them. On Bayesian updating, it is only Mutual Information Unification that contributes to the incremental support of a hypothesis by the evidence unified. This poses a challenge for defenders of a view that explanation ought to be taken as a confirmatory virtue that makes a contribution in its own right to incremental support; in order for such a view to be defensible, its advocates must ground it in some relevant difference between humans and a Bayesian agent. Options for such a defense are considered, and it is concluded that common origin unification has at best a limited heuristic role to play in confirmation. Finally, it is shown how Reichenbachian common cause hypotheses fit into the schema of mutual information unification. Keywords: Unification, explanation, confirmation, Bayesianism, common cause.
1 Introduction

Myrvold (2003) identified what was described therein as “one interesting sense” in which a theory can unify phenomena. This consists of the ability of the theory to render distinct phenomena informative (or more informative) about each other. Call this Mutual Information Unification (MIU). This sense lends itself nicely to a probabilistic explication, and it can be shown that unification in this sense contributes to incremental evidential support of the theory by the phenomena unified.

There is another sense of unification, having to do with hypotheses that posit a common origin for the phenomena in question, be it a common cause or some other type of explanation. Call this Common Origin Unification (COU). As emphasized by Lange (2004) and Schupbach (2005), the two senses are logically independent; neither is a necessary or a sufficient condition for the other, even though, in a number of interesting cases, they are concomitants of each other.

In this paper, the respective roles of these two notions of unification in connection with the bearing of evidence on a theory are discussed. There are, of course, other questions one might ask, and other roles for a notion of unification to play besides contributing to confirmation. Having a common explanation for disparate phenomena can contribute to deeper understanding, which is one goal of scientific research. Insofar as it contributes to such understanding, Common Origin Unification may play the role of a cognitive value. As such, it can play a legitimate role in questions such as that of which research programme to pursue; a theory might be regarded as more worthy of development on account of its potential for affording understanding. This is a different matter from the question at issue in this paper, which is whether unification ought to be regarded as contributing to the evidential support of a theory by the phenomena unified.

On the question of the respective roles of these two notions of unification in theory confirmation, on a Bayesian analysis, the answer is clear: Mutual Information Unification contributes to incremental

\footnote{I am grateful to Michel Janssen for making this suggestion. See Myrvold (2011), and references therein, on the subject of how to incorporate cognitive values into a Bayesian framework.}

\footnote{Cf. Salmon (2001, 130): “the scientist might say that Halley’s hypothesis is worth pursuing, not because it is more likely to be true, but because, if it should turn out to be true, it would be extremely valuable in terms of informational content.”}
evidential support, and there is no scope, within Bayesian updating, for Common Origin Unification to add to the evidential support of the theory (see §4, below).

There may be some who do not take this to settle the normative issue, and will maintain that, despite the Bayesian verdict, we ought to take explanatory power of a hypothesis as a confirmatory virtue. Advocates of such a view would have to reject the idea that we should take consideration of a Bayesian agent updating via conditionalization as normative for those of us who are not such agents. This presents a challenge for such advocates. If it is rational, or reasonable, or otherwise well and good for us to do what is impossible for a Bayesian agent updating its credences via conditionalization, that is, to take Common Origin Unification to be something that makes a contribution to evidential support, above and beyond what it contributes to Mutual Information Unification, then this must be grounded in some relevant difference between us and Bayesian agents. It is incumbent on an explanationist to give an account of what that difference is.

In the following, these points are first illustrated by means of a simple example that, despite its artificiality, shares some salient features with cases of actual scientific interest. Next, in §3, are presented the probabilistic measures of MIU introduced in Myrvold (2003), and in §4 their impact on evidential support is exhibited. In §5 are outlined possible reactions to the Bayesian verdict regarding the respective confirmatory roles of the two types of unification. In §6 the question is addressed whether there is still a role for Common Origin Unification to play in hypothesis assessment, in assessing priors rather than in assessing incremental evidential support (the answer is no). Finally, in §7 it is shown how Reichenbachian common causes fit into the schema of Mutual Information Unification.

2 Two Kinds of Unification

Consider the following toy example, of no use except for introducing the issues at hand, though it does share some salient features with a multitude of real-world cases of genuine scientific interest. You are about to be presented with two data streams, \(S_1\) and \(S_2\), each of which will be sequences of ten Heads or Tails. You know that they have been produced by coin flipping, but you aren’t sure of exactly the procedure used, or whether the coin or coins involved are fair.
Suppose that you have nonzero credences in both of the following hypotheses:

$H_1$: A fair coin was flipped ten times, and the results of this series of coin flips are reported in both data streams.

$H_2$: Two fair coins were flipped ten times each, and each data stream reports the results of one of these series of coin flips.

I invite you to consider the effect of the evidence on these two hypotheses. That evidence consists of specification of the two data streams:

$S_1$: HHHTTHTTHTT

$S_2$: HHHTTHTHHTT

Let $E_1$ be the proposition that $S_1$ is the string given above, and $E_2$ the corresponding proposition about $S_2$.

Now, if you have nonnegligible prior credence that the strings might have been produced by radically unfair coins, $E_1$ and $E_2$ might boost your confidence in the fairness of the coins, and hence conditionalizing on each of $E_1$ and $E_2$, separately, might boost your credence in both $H_1$ and $H_2$. But, when taken together, $E_1$ and $E_2$ strongly favor $H_1$ over $H_2$.

There are two features of this example that I would like to draw your attention to. The first feature is that $H_1$, if true, renders $E_1$ informative about what data stream $S_2$ will be. Conditional on $H_1$, knowing $E_1$ permits one to anticipate the truth of $E_2$. That is, $H_1$ exhibits Mutual Information Unification ($MIU$) with respect to the evidence set $\{E_1, E_2\}$. A hypothesis has this property, with respect to a set of evidential propositions, if conditionalizing on that hypothesis increases the mutual informativeness of the set. Obviously, this is the sort of thing that comes in degrees. In our toy example, conditional on $H_1$, knowledge of $E_1$ permits one to anticipate, with certainty, all details of $E_2$. In more interesting cases the increase of informativeness will be less than maximal. Probabilistic measures of degree of this sort of unification will be introduced below.

The second feature is that $H_1$ posits a common origin of the two data streams, and thus is ripe to be the subject of what Janssen (2002) has called a COI story, for Common Origin Inference. In addition to MIU, $H_1$ also exhibits Common Origin Unification ($COU$).

The two concepts are of a manifestly very different sort. One belongs to a cluster of concepts involving information, states of knowledge, and the like; the other is related to concepts of cause and ex-
As already mentioned, they are logically independent. A hypothesis can posit a common origin for two (or more) evidential propositions without making them mutually informative about each other, as the propositions could be about independent aspects of their posited common origin; thus, we can have COU without MIU. Furthermore, once two or more evidential propositions are known, that is, have been absorbed into one’s background knowledge, they are no longer informative about each other, though any common origin they might have remains, and again we have COU without MIU. One can also trivially construct hypotheses that exhibit MIU without COU.

With respect to our toy example, consider the hypothesis,

\[ H_3: \] Two fair coins were flipped ten times each, each data stream reports the results of one of these series of coin flips, and the results of each series of flips just happened to be the same.

This hypothesis, if true, also renders one data stream informative about the other. Of course, prior to the evidence, one would expect credence in \( H_3 \) to be low, lower than credence in \( H_2 \) by a factor of 1,024.

Though artificial, our toy example has a multitude of parallels in actual science. One is the case of heliocentric vs. geocentric world systems, discussed by Janssen (2002) and Myrvold (2003). The analog of \( H_1 \) is what was called \( h_C \) in Myrvold (2003), that is, the heliocentric hypothesis that all planets have circular or nearly circular orbits centred at or near the sun, and the analog of \( H_2 \) is the bare-bones geocentric hypothesis \( h_P \), which posits that, for each planet, there is a deferent circle centered near the earth, and that the planet travels on an epicycle whose center travels on the deferent, with no assumption made about any connections between the motions of different planets or between planetary motions and the motion of the sun. The analog of \( H_3 \) is the geocentric hypothesis conjoined with the sun-planet parallelism condition; this is the hypothesis \( h_{SP} \), or the strengthened Ptolemaic hypothesis.

\[ ^3 \] Similar remarks apply to probabilistic measures of explanatory power such as those proposed by Popper (1954, 1959), Good (1960), Schupbach and Sprenger (2011), and Crupi and Tentori (2012). Glymour (2015) has argued that it would be a grave mistake to take any of these probabilistic notions as an explication of explanatory power. This seems to be generally accepted by recent authors; Schupbach and Sprenger, for example, are clear that what is proposed is a measure of strength of explanation between propositions bearing an antecedently identified explanatory relation to each other.
One can find analogs in cases in which a hypothesis turns disparate, prima facie unrelated phenomena into agreeing measurements of some theoretical parameter. The classic case is Perrin’s argument for the existence of atoms. Perrin (1913, §119; 1916, §120) adduces 13 distinct phenomena that, on the atomic hypothesis, count as measurements of Avogadro’s number. The analog of \( H_1 \) is that atoms exist, and hence there is a common origin explanation of the agreement of these measurements; the analog of \( H_2 \) would be the hypothesis that matter is continuously divisible, and the analog of \( H_3 \) would be the hypothesis that adds to \( H_2 \) the stipulation that Perrin’s 13 phenomena yield values that just happen to agree within experimental error, even though they are not agreeing measurements of any physically meaningful quantity. Another example is the quantum hypothesis, which turns disparate phenomena into agreeing measurements of Planck’s constant; see Kao (2015).

3 Probabilistic Measures of Unification

Consider a Bayesian agent whose credences are represented by a probability function \( C_r \). We define the mutual information of a pair of propositions, \( \{p_1, p_2\} \), relative to background \( b \), by

\[
I(p_1, p_2|b) = \log_2 \left( \frac{C_r(p_2|p_1b)}{C_r(p_2|b)} \right) = \log_2 \left( \frac{C_r(p_1 p_2|b)}{C_r(p_1|b)C_r(p_2|b)} \right). \tag{1}
\]

If \( p_1 \) and \( p_2 \) are probabilistically independent on \( b \), then \( I(p_1, p_2|b) \) is zero; it is positive if conditionalizing on one boosts credence in the other, negative, if conditionalizing on one lowers credence in the other.

For a larger set, \( p = \{p_1, p_2, \ldots, p_n\} \), we add up the information yielded by \( p_1 \) about \( p_2 \), the information yielded by \( p_1 p_2 \) about \( p_3 \), and so on, up to the information about \( p_n \) yielded by the conjunction of

\[\text{\footnotesize A note on notation. We will use concatenation for conjunction, and the overbar } \bar{\overline{p}} \text{ for the negation of } p. \text{ We use boldface letters to denote sets of propositions. Note that these are sets and are not replaceable by a single proposition that is their conjunction. Thus, } \{p_1, p_2\} \text{ is not the same set as } \{p_1 p_2, T\}, \text{ where } T \text{ is the logically true proposition, though the conjunction of their members is the same. This matters because we will be concerned with the mutual informativeness of members of a set of propositions; } p_1 \text{ and } p_2 \text{ may be mutually informative though the logically true proposition is not informative about their conjunction or anything else.}\]
all the others.\(^5\)

\[
I(p_1, \ldots, p_n|b) = I(p_1, p_2|b) + I(p_1 p_2, p_3|b) + \ldots + I(p_1 \ldots p_{n-1}, p_n|b)
\]

\[
= \sum_{k=1}^{n-1} I\left( \bigwedge_{i=1}^{k} p_i, p_{k+1}|b \right). \tag{2}
\]

Although the form of (2) does not make this obvious, this quantity is independent of the order in which the elements of the set \( p \) are taken, and we have,

\[
I(p_1, \ldots, p_n|b) = \log_2 \left( \frac{Cr(p_1 p_2 \ldots p_n|b)}{Cr(p_1|b) Cr(p_2|b) \ldots Cr(p_n|b)} \right)
\]

\[
= \log_2 \left( \frac{Cr(\bigwedge_{i=1}^{n} p_i|b)}{\prod_{i=1}^{n} Cr(p_i|b)} \right). \tag{3}
\]

With a slight abuse of notation, we will write \( I(p|b) \) for \( I(p_1, \ldots, p_n|b) \). We will also drop, as irrelevant, the base of the logarithm, since changing base is only a matter of a constant multiplicative factor.

\( I(p|b) \) is the logarithm of the quantity that appears in Keynes’ *Treatise on Probability* (1921, §XIV.8) as the *coefficient of dependence*, with an attribution to unpublished work by W.E. Johnson.\(^6\) It was called a measure of similarity by Wayne (1995) and Myrvold (1996), and taken by Shogenji (1999) as a measure of coherence of a set of propositions.

We will say that a hypothesis \( h \) MIUnifies a set \( e = \{e_1, \ldots, e_n\} \), relative to background \( b \), if and only if

\[
I(e|h \ b) > I(e|b). \tag{4}
\]

This suggests a way to measure the degree to which a hypothesis MIUnifies a set of evidential propositions.\(^7\)

\[
MIU_1(e; h | b) = I(e|h b) - I(e|b). \tag{5}
\]

\(^5\)Obviously, a single number cannot capture all the informational relations there could be between elements of a set of more than two members. This would require a specification of all \( I(q, q'|b) \), where \( q \) and \( q' \) range over all conjunctions of elements of \( p \). But it is this quantity that will be useful for the purposes at hand.

\(^6\)I am indebted to Brössel (2015) for pointing this out.

\(^7\)This quantity is the logarithm of a quantity that was referred to as an “interaction term” in Myrvold (1996), and is called *focussed correlation* in Wheeler (2009), Schlosshauer and Wheeler (2011), and Wheeler and Scheines (2013). What we are calling \( MIU_1 \) was called \( U \) (for unification) in Myrvold (2003). \( MIU_2 \) was discussed therein, though not given its own name.
We might also be interested in whether a hypothesis does a better job of unifying a set of propositions than its negation. Define

\[
MIU_2(e; h|b) = MIU_1(e; h|b) - MIU_1(e; \overline{h}|b) = I(e|h b) - I(e|\overline{h} b).
\] (6)

The two are not ordinally equivalent, and, indeed, need not agree as to sign. Suppose a hypothesis \( h \) unifies a body of evidence, relative to background \( b \). That is, suppose the evidence is more mutually informative conditional on \( hb \) than on \( b \) alone. Then \( MIU_1(e; h|b) \) is positive. But whether \( MIU_2(e|b) \) is negative or positive depends on whether or not \( \overline{h} \) unifies the evidence more. If \( I(e; h|b) \) is greater than \( I(e; h|b) \), then, even if \( MIU_1(e; h|b) \) is positive, \( MIU_2(e; h|b) \) is negative. In fact, all four combinations of signs of \( MIU_1 \) and \( MIU_2 \) are possible, though it is easy to show that, unless \( e_1 \) and \( e_2 \) are, when taken individually, oppositely relevant to \( h \) (that is, unless one of them is positively relevant and the other negatively relevant), if \( MIU_1(e_1, e_2; h|b) \) is positive, \( MIU_2(e_1, e_2; h|b) \) is also positive. See Appendix for details.

Both of these quantities are special cases of a comparative measure of unification,

\[
MIU_c(e; h_1, h_2|b) = I(e|h_1 b) - I(e|h_2 b).
\] (7)

On McGrew’s account of consilience, \( h_1 \) is said to be more consilient than \( h_2 \) with respect to \( e \) to the extent that \( I(e|h_1 b) > I(e|h_2 b) \), or, equivalently, to the extent that \( MIU_c(e; h_1, h_2|b) > 0 \) (McGrew, 2003, 562).

Readers are asked to kindly refrain from engaging in a battle of the intuitions over whether \( MIU_1 \) or \( MIU_2 \) is the One True Measure of degree of unification. They are simply measuring different things, and if you have intuitions that are incompatible with properties that one or another of these quantities possesses, then your intuitions are about some other concept.\(^8\)

4 The Evidential Value of Unification

To some readers, it might seem obvious that what counts when it comes to confirmation is Common Origin Unification, with Mutual

\(^8\) And if your intuitions find it repugnant to use the word “unification” in connection with either of these, then feel free to use a different word.
Information Unification being a poor cousin that hardly merits the illustrious family surname. This view is expressed by Marc Lange, who writes,

the examples I have given suggest that insofar as theories that unify in the stronger,\(^9\) ontological-explanatory sense derive greater support in virtue of the unification they achieve, they do so not solely in virtue of their achieving unification in the weaker, creating-mutual-positive relevance sense. The stronger sense of unification is epistemically significant. In the case of the light-quantum hypothesis, \(h_C\) and \(h_L\) both supply unity in the weaker sense, but Einstein took \(h_L\) to receive greater support from the phenomena than \(h_C\) by virtue of \(h_L\)'s unifying those phenomena in an ontological-explanatory sense (Lange 2004, 212).

Here \(h_L\) is Einstein’s light quantum hypothesis, and \(h_C\) is the hypothesis that \(h_L\) is false but nevertheless, by sheer coincidence, light behaves as if it were quantized. According to Lange, \(h_L\) receives greater support from the phenomena unified than does \(h_C\).

It is not entirely clear whether incremental or absolute support is meant, where incremental support has to do with an increase in credibility lent to a hypothesis by the evidence, and absolute support with the credibility of the hypothesis, taking all known considerations into account. If absolute, this suggests that the case of \(h_C\) is analogous to that of our toy example’s \(H_3\), which is accorded a low prior because it posits an improbable coincidence. One the other hand, if the claim is to be a counterexample to anything in Myrvold (2003), incremental support must be what is meant. Let us therefore consider the position that, when it comes to incremental support, it is COUnification, not MIUnification, that counts.

A Bayesian analysis renders the opposite verdict: when it comes to incremental support of a hypothesis, it is MIUnification, rather than COUnification, that matters.

One popular measure of the degree to which an evidential proposition \(e\) lends incremental confirmation to a hypothesis \(h\), relative to background \(b\), is the ratio of posterior probability of \(h\) to its prior probability. This is, of course, ordinarily equivalent to its logarithm.

\(^9\)This is a slip; the two senses are, as Lange emphasizes, logically independent.
Let us define

$$R(h; e|b) = \log \left( \frac{Cr(h|eb)}{Cr(h|b)} \right).$$  \hfill (8)

Another is the ratio of the posterior odds of $h$ to its prior odds, or, equivalently, the logarithm of this, called \textit{weight of evidence} by Good (1950). Define

$$W(h; e|b) = \log \left( \frac{Cr(h|eb)/Cr(h|b)}{Cr(e|h)/Cr(e|b)} \right) = \log \left( \frac{Cr(e|h)}{Cr(e|\bar{h})} \right).$$  \hfill (9)

As Myrvold (2003) pointed out, on either way of measuring incremental confirmation, we have a contribution of unification to confirmation.\footnote{Equation (10) corresponds to (6) of Myrvold (1996) and to (12) of Myrvold (2003); (11) corresponds to (13) of Myrvold (2003). Closely related results appear already in Keynes (1921, 151–154); in particular, our equation (11) is essentially the same as Keynes' (48).} The incremental support, as measured by $R$, of $h$ by $e$ can be decomposed into a sum of increments due to the individual members of $e$, plus an additional term that is the degree of MIUnification (positive or negative) of $e$ by $h$, as measured by $MIU_1$.

$$R(h; \bigwedge_{i=1}^{n} e_i | b) = \sum_{i=1}^{n} R(h; e_i | b) + MIU_1(e; h|b).$$  \hfill (10)

The result for $W$ takes the same form, with $MIU_2$ in place of $MIU_1$.

$$W(h; \bigwedge_{i=1}^{n} e_i | b) = \sum_{i=1}^{n} W(h; e_i | b) + MIU_2(e; h|b).$$  \hfill (11)

These relations can be readily verified by the reader.

Since the MIU-term is not the only contribution to the increment of confirmation, it would be incorrect to gloss these results as saying that hypotheses that are more unifying receive more confirmation. Although it would not be incorrect to say that \textit{ceteris paribus}, a hypothesis that achieves a higher degree of MIUnification of the evidence is accorded greater incremental support, this is strictly weaker than what is conveyed in equations (10) and (11), and there is no advantage in making the \textit{ceteris paribus} claim when it is a trivial matter to say how things stand when all else is \textit{not} equal.

Imagine, now, a Bayesian agent that had numerical credences, which it\footnote{I say “it,” because a being with precise numerical credences would be far from human.} updated by conditionalizing on new items of evidence.
Then, depending on how we measured degree of incremental confirmation, the confirmational boost accorded to \( h \) by a set \( e \) of evidential propositions would be given by either (10) or (11). In each case the additional confirmational boost, beyond that attributable to the items of evidence taken singly, is given by the MIUnification term.

Applied to our toy example: The fact that \( H_1 \) and \( H_3 \) make \( E_1 \) and \( E_2 \) informative about each other is reflected in the likelihoods, \( Cr(E_1E_2|H_1) \) and \( Cr(E_1E_2|H_3) \), which are higher than \( Cr(E_1E_2|H_2) \) by a factor of 1,024. Thus, relative to \( H_2 \), credence in \( H_1 \) and \( H_3 \) is boosted:

\[
\frac{Cr(H_1|E_1E_2)}{Cr(H_1)} = \frac{Cr(H_3|E_1E_2)}{Cr(H_3)} = 1,024 \times \frac{Cr(H_2|E_1E_2)}{Cr(H_2)}.
\] (12)

It doesn’t follow, of course, that \( H_3 \) gets final credence comparable to that of \( H_1 \). Since \( H_3 \) posits an improbable coincidence, it is accorded a lower prior probability, lower than that of \( H_2 \) by a factor of 1,024; the additional confirmational boost it receives is just enough to bring it up to posterior credence equal to that of \( H_2 \) (which, of course, must be the case, since, given the evidence, \( H_3 \) is true if and only \( H_2 \) is).

There is a close parallel between this case and the case of geocentric vs heliocentric world systems, and also the case of the light quantum, considered by Lange.

In the case of planetary motion, on both the heliocentric hypothesis and the strengthened Ptolemaic hypothesis, features of one planet’s apparent motion are informative about features of others’ (see Janssen 2002 and Myrvold 2003 for discussion). In the case of the heliocentric hypothesis, \( H_C \), these have a common origin in the motion of our vantage point as observers on earth; for \( H_{SP} \), they are the consequence of the posited sun-planet parallelism. Against a background that includes little or no information about observed planetary motions, both of these get a confirmational boost from the celestial phenomena, due to the \( MIU \)-component of incremental confirmation. It doesn’t follow that they end up with equal posterior credence. Arguably, \( H_{SP} \), on that background, should be accorded markedly lower prior credence than \( H_P \), as it posits a relation that \( H_P \) by itself would not lead one to anticipate. \( H_C \) and \( H_{SP} \) get the same incremental confirmation on the evidence. Therefore, posterior credence in \( H_C \) will be markedly higher than posterior credence in \( H_{SP} \) unless prior credence in \( H_C \) is markedly lower than prior credence in \( H_P \).

Something similar can be said in regards to Lange’s case of the
light quantum hypothesis. Let us grant that the light quantum hypotheses plays a unificatory role. Lange asserts that Einstein took the observed phenomena to lend greater support to the light quantum hypothesis than the hypothesis that, by sheer coincidence, all observable phenomena are as if the light quantum hypothesis is true. The suggestion is that that such a judgment is the right one, given the evidence available to Einstein in 1905. In order for this claim to be relevant to the issue at hand, this must mean that the phenomena lend greater incremental support to the light-quantum hypothesis than to the coincidence hypothesis. One might also regard $h_C$ as so implausible as to be dismissed out of hand. But this would mean according it a low prior, which is consistent with the Bayesian account of the virtue of unification.

5 Possible Reactions to the Bayesian Verdict

Bayesian updating leaves no room for an additional confirmatory boost to be attached to hypotheses with greater explanatory power: the contribution to incremental support comes via the MIUnification term. There is a tension between this Bayesian verdict and the thought that COUnification should play a role in incremental confirmation above and beyond its contribution to MIUnification. We have here an exact parallel with van Fraassen’s argument against those who would take explanatory power of a theory to yield an extra confirmatory boost, beyond that yielded by conditionalization on the evidence (van Fraassen, 1989, 166–169).

One reaction might be to downplay the distinction, focussing on cases in which explanationist and Bayesian judgments agree. One might be tempted to declare that hypotheses that provide ‘lovelier’ explanations are precisely those that bestow higher likelihood on a hypothesis. This is not tenable as a general thesis. Although, in many interesting cases, explanation and likelihood go together, the connection is not so tight that they never come apart. The interesting question is what the explanationist will say about the cases in which they do come apart.

One possible reaction, in my opinion the correct one, is to use the Bayesian verdict to correct any intuitions one might have that are in tension with it. The ability of a theory to unify disparate phenomena
by positing a common origin plays a confirmatory role only insofar as the posited common origin renders distinct phenomena informative (or more informative) about each other. A temptation to assign it a stronger role in confirmation might be ascribed, in part, to a conflation of distinct questions (a conflation encouraged by philosophers’ overuse of the phrase “theory choice,” a phrase that conflates distinct sorts of choices). Certainly, a hypothesis’ power to explain, if true, can contribute to making it worthwhile to pursue a project of developing a theory that includes that hypothesis, and it can contribute to the value of accepting the hypothesis, if true; we should only be wary of thinking that everything that contributes to making a hypothesis pursuit-worthy also lends it greater credibility. The temptation might also be ascribed, in part, to not distinguishing between incremental confirmation and overall credibility in the light of all evidence. The most obvious examples that exhibit MIUnification without COUnification are those such as our $H_3$, that achieve it by brute fiat, by tacking on an improbable conjunct, and we rightly regard these as implausible.

This suggests one way in which an explanationist might retrench; the import of COUnification might be relegated to informing priors. While, certainly, common-origin considerations sometimes play a role in assessing prior credibility, I am skeptical that anything beyond a very limited role can be defended; more on this in the next section.

The only other avenue of defense for an advocate of an explanationist thesis would be to deny that considerations of how a Bayesian agent would update have normative force for the judgments of human scientists.

A line of defense along these lines of thought would have to ground it in some relevant difference between us and Bayesian agents. We are certainly different from Bayesian agents in a number of ways. We do not have precise numerical degrees of belief; our judgments about how likely or unlikely a hypothesis is tend to be vague. Moreover, as an abundance of empirical evidence shows, routinely our qualitative judgments of the relative credibility of various propositions are not even compatible with the existence of numerical credences satisfying the axioms of probability, and our changes in credences are often not in accord with Bayesian conditionalization.

The usual understanding of facts of this sort is that they are due to cognitive limitations, and that some of them can be understood as resulting from usually reliable heuristics, of the sort that any agent
with limited cognitive capacities would be well-advised to employ as an alternative to spending excessive time on cogitation. In taking such limitations into account, one does not *ipso facto* abandon the domain of normativity for descriptive psychology. From a decision-theoretic point of view, deployment of such heuristics can be regarded as *rational* behavior for a cognitively limited agent. This involves what I. J. Good (1971, 1976) called “Type II Rationality”: decision-making that takes into account the cost in time and cognitive effort of the act of deliberation.

Peter Lipton has offered a limited defense of explanationism along these lines. We are often not very good, he notes, at judging likelihoods correctly.

My thought is this. In many real life situations, the calculation that the Bayesian formula would have us make does not, in its bare form, meet the requirement of epistemic effectiveness; it is not a recipe we can readily follow. . . . My suggestion is that explanatory considerations of the sort to which Inference to the Best Explanation appeals are often more accessible than those probabilistic principles to the inquirer on the street or in the laboratory, and provide an effective surrogate for certain components of the Bayesian calculation. On this proposal, the resulting transition of probabilities in the face of new evidence might well be just as the Bayesian says, but the process that actually brings about the change is explanationist (Lipton 2004, 113-114; see also Lipton 2001, 110–111).

On such a view, when a judgment needs to be made on the fly, it is better to invoke an explanationist heuristic than to spend time thinking through likelihoods; this will, one hopes, provide judgments that are not too far off, either most of the time or in the most significant cases. Though Lipton suggests that the division of labor between Bayesian and explanationism maps onto the distinction between normative and descriptive accounts, he also uses language that suggests that we cognitively limited agents are well-advised to employ explanationist considerations as a surrogate for doing a Bayesian calculation: “explanatory considerations help us to perform what is in effect a Bayesian calculation” (Lipton 2004, 120). This suggests that considerations of Type II rationality are in play.

Using a heuristic of this sort as a surrogate for a considered evalua-
tion of likelihoods carries with it a risk of error, in those cases in which
COU and MIU come apart. Presumably, Lipton would agree that, in such cases, if an accurate appraisal of the import of the evidence matters, one should correct the explanationist judgment by reference to the Bayesian one. On Lipton’s view, the role of explanationist considerations is severely constrained.

Can a stronger defense of explanationism be mounted? It is doubtful. Since such a defense would have to be grounded in some difference between cognitively limited humans and Bayesian agents, it’s hard to see any role for explanationist consideration beyond the limited heuristic role envisaged by Lipton.

6 A Prior Preference for Unifying Hypotheses?

We have considered cases (in the toy example, $H_1$ and $H_3$, in the case of planetary motion, $H_C$ and $H_{SP}$, and in the light quantum case, $h_L$ and $h_C$), in which each of a pair of hypothesis possesses the same ability to render items of evidence informationally relevant to one another, but they do so in different ways. In each of these cases one does it by virtue of positing a common origin for prima facie unrelated phenomena, the other, by brute fiat, in positing an unexplained correlation between the phenomena. In each of these cases, the hypothesis that involves a common origin is, arguably, less implausible than the one that posits brute coincidence.

One might be tempted to generalize, positing, that, whenever we have MIU without COU, there will be a corresponding hypothesis that achieves precisely the same MIUnification via COUnification, and we should accord much less prior credence to the hypothesis that exhibits MIU without COU than to the one that achieves it via COU. This would mean that there is a role for COU, not in incremental confirmation, but in setting priors.

Anything so sweeping would be a mistake, I think. There are patterns in the world of all sorts, some due to some sort of common origin, some not. We should not demand that a common origin be found for every similarity between two phenomena. Given any pattern in the phenomena, however, it will be possible to cook up an artificial MIUnifying hypothesis. We ought not seek a common origin lurking behind every such hypothesis!
Perhaps, then, the generalization should be that, when we do have a pair of hypotheses that both induce the same informational relevance relations among a body of phenomena, one doing it via COUnification and the other by brute fiat, we should attach higher prior credence to the COUnifying hypothesis.

This is still too sweeping. When we have a case of two hypotheses \( h_1 \) and \( h_2 \) of roughly equal prior credibility, and create a third \( h_3 \) by tacking on to \( h_2 \) some conjunct with low prior plausibility, then, indeed, in such a case, we should place lower credence in \( h_3 \) than in \( h_1 \). But not all cases will be like that, and a COUnifying hypothesis might be deemed implausible on other grounds. Take, for example, Ptolemy’s attitude towards heliocentric hypotheses. Since Ptolemy recognized that in the observed phenomena there is a connection between the apparent motion of the sun and that of the other planets, he was in a position to appreciate the COUnifying power of heliocentrism. But, since he accepted Aristotelian physics for terrestrial phenomena, he thought that terrestrial phenomena ruled out a diurnal rotation of the earth (see Ptolemy 1984, Bk. I, §7); for him, it was reasonable to place low credence in heliocentric theories that posited such a rotation.

One can exhibit plenty of hypothesis pairs in which the less unifying, less explanatory hypothesis has less prior credibility, because the less explanatory hypothesis posits an implausible coincidence. But the emphasis should be on the credibility-diminishing role of coincidence, rather than any prior conviction that nature is unified. What \( H_3 \), the strengthened Ptolemaic hypothesis, and Lange’s \( h_C \) have in common is that, in each case, we have a hypothesis to which is tacked on some additional condition that one would not expect to hold in the absence of evidence that it does, and hence we have a hypothesis that ought to be accorded low prior credence. Rather than a sweeping preference for COUnification, I suggest that the methodological adage that underwrites low prior credence in such hypotheses is:

Place little prior credence in things you take to be improbable.

This is, I hope, unobjectionable! It is, of course, utterly empty, but I am skeptical that anything stronger could be defended as a maxim of more than very limited scope.

It would be a mistake to raise this bland but unobjectionable maxim into a global rejection of hypotheses that posit coincidences.
Improbable things do happen, after all. Moreover, in some cases it is reasonable to accept hypotheses that posit an improbable coincidence. The evidence available to you in the toy example strongly suggests a common cause. But, if you were to obtain strong evidence that the two data streams were the results of independent tosses of two fair coins, then it would be reasonable to accord high credence to $H_3$. For a real-world case: Ptolemy propounded a geocentric system with an unexplained sun-planet parallelism, because he thought he had strong evidence to rule out hypotheses that involved a moving earth.

7 Unification and Reichenbachian Common Causes

Among unifying hypotheses are those that posit a Reichenbachian common cause to explain some observed statistical correlation (Reichenbach, 1956, §19). This type of hypothesis fits well within the schema of the Bayesian account of unification, but, since this might not be obvious, it is worth showing how it fits.

Consider two sequences of propositions, \{A_i, i = 1, \ldots, n\}, and \{B_i, i = 1 \ldots, n\}. Given such sequences, let $n(A)$ be the number of true instances of the $A_i$s, and let $f(A) = n(A)/n$ be the relative frequency of true instances of the $A_i$. Define $f(B)$ and $f(AB)$ similarly. Let $E_1$ be a proposition expressing which of the $A_i$s are true, and which are false. For example, in our toy example, $A_i$ could be the proposition that the $i$th element of $S_1$ is Heads, and $E_1$ would be

$$A_1A_2A_3\bar{A}_4\bar{A}_5A_6\bar{A}_7A_8A_9\bar{A}_{10}.$$ 

Let $E_2$ be the evidence statement specifying the $B$-sequence.

A statistically significant difference between $f(AB)$ and the product $f(A)f(B)$ is thought to call for explanation. A Reichenbachian Common Cause of an observed correlation between $A$ and $B$ is a third sequence $C_i$ that screens off their correlation. That is,

$$Pr(A_iB_i|C_i) = Pr(A_i|C_i)Pr(B_i|C_i);$$

$$Pr(A_iB_i|\bar{C}_i) = Pr(A_i|\bar{C}_i)Pr(B_i|\bar{C}_i).$$

(13)

A hypothesis that posits a common cause of this sort, if it leads one to expect correlations close to those observed, clearly, can be supported
by evidence in which there is an observed statistical correlation between two sequences of events. Such a hypothesis can be a MIUniﬁxing hypothesis, in the sense of making the evidence statements \( E_1 \) and \( E_2 \) mutually informative.

This might seem paradoxical. A common cause screens off the correlations between the \( A_i \)s and \( B_i \)s; how can it be that, at the same time, there is a conﬁrmational boost associated with rendering them informative about each other?

The answer to this is: the hypothesized common causes \( C_i \) screen off the correlations, but a hypothesis \( H_{cc} \) that posits that there are common causes of the right sort can render the truth or falsity of \( A_i \) informative about the truth or falsity of \( B_i \), and hence render \( E_1 \) and \( E_2 \) mutually informative. That is, a hypothesis that there is a common cause of the right sort will lead one to expect correlations between the \( A_i \)s and \( B_i \)s, and so count as MIUniﬁxing with respect to the evidence set \( \{E_1, E_2\} \), relative to a background against which the observed correlations are unexpected.

Moreover, each event \( C_i \) can count as a common origin of \( A_i \) and \( B_i \). Let \( H_{cc} \) be some hypotheses according to which there exists a sequence \( \{C_i\} \) satisfying (13). Suppose that, on the supposition of \( H_{cc} \), \( C_i \) is a probability raiser for both \( A_i \) and \( B_i \), as a cause should be, and suppose that, according to \( H_{cc} \), for each \( i \), \( C_i \) and \( \bar{C}_i \) both have nonzero probability. Then, even though, for each \( C_i \), the truth or falsity of \( C_i \) screens off informational relations between \( A_i \) and \( B_i \), the supposition of \( H_{cc} \) leads one to expect correlations between the \( A_i \)s and the \( B_i \)s.

\[
Pr(A_i B_i|H_{cc}) > Pr(A_i|H_{cc}) Pr(B_i|H_{cc}). \tag{14}
\]

Let us now see in more detail how this works. We consider the bearing of the statistical evidence stemming from observation of the \( A \)-sequence and the \( B \)-sequence on members of a family of hypotheses, each of which posits the existence of a Reichenbachian common cause. For simplicity, we consider only hypotheses on which distinct \( A_i \)s are independent and identically distributed, as are \( \{B_i\} \) and \( \{A_i B_i\} \). The statistical data can be accounted for on a hypothesis positing \( C_i \)s that are also independently and identically distributed. Any hypothesis positing a common cause of this sort can be characterized by five
parameters:
\begin{align*}
p &= Pr(C_i), \\
a_1 &= Pr(A_i|C_i), \quad a_0 = Pr(A_i|\bar{C}_i), \\
b_1 &= Pr(B_i|C_i), \quad b_0 = Pr(B_i|\bar{C}_i).
\end{align*}

(15)

Probabilities for the \(A_i\)s, \(B_i\)s, conditional on a hypothesis of this sort, are
\begin{align*}
Pr(A_i|H_{cc}) &= p a_1 + (1 - p) a_0, \\
Pr(B_i|H_{cc}) &= p b_1 + (1 - p) b_0,
\end{align*}

(16)

and their covariance is,
\begin{align*}
\text{Cov}(A_i, B_i|H_{cc}) &= Pr(A_i B_i|H_{cc}) - Pr(A_i|H_{cc})Pr(B_i|H_{cc}) \\
&= p(1 - p)(a_1 - a_0)(b_1 - b_0).
\end{align*}

(17)

As pointed out by Reichenbach (1956, 159–161), and as can be readily seen from (17), if \(p \in (0, 1)\) and \(a_1 - a_0\) and \(b_1 - b_0\) are both positive, then, conditional on the hypothesis \(H_{cc}\), the \(A_i\)s are positively correlated with the \(B_i\)s. Obviously, the same conclusion follows if \(a_1 - a_0\) and \(b_1 - b_0\) are both negative; also, the \(A_i\)s are negatively correlated with the \(B_i\)s if \(a_1 - a_0\) and \(b_1 - b_0\) have opposite sign, and they are uncorrelated if the \(C_i\)s are irrelevant to either the \(A_i\)s or the \(B_i\)s, that is, if \(a_1 = a_0\) or \(b_1 = b_0\).\(^{12}\) The family of all such hypotheses, thus, includes as a special case those that posit no common cause for \(A_i\) and \(B_i\).

We inquire into the degree of support lent to common-cause hypotheses, with various values of the parameters, by the pair \({E_1, E_2}\). Let \(H_{cc}\) be some hypothesis of the form considered above. We have, from (10),
\begin{equation}
R(H_{cc}; E_1E_2) = R(H_{cc}; E_1) + R(H_{cc}; E_2) + MIU_1(\{E_1, E_2\}; H_{cc}).
\end{equation}

(18)

Since we’re interested in comparing degrees of support for different hypotheses on a fixed body of evidence, it is useful to compare log-likelihoods, as, for two different hypotheses, the differences between

\(^{12}\)These probabilistic facts were familiar in the statistical literature well before Reichenbach’s use of them; see Yule (1911, §§IV.6–7).
their $R$-values will be the same as the differences between the respective log-likelihoods. The log-likelihoods can be partitioned in a manner parallel to our partitioning of $R$:

$$\log Pr(E_1 E_2 | H_{cc}) = \log Pr(E_1 | H_{cc}) + \log Pr(E_2 | H_{cc}) + I(E_1, E_2 | H_{cc}).$$  (19)

The first two terms of this are

$$\log Pr(E_1 | H_{cc}) = n(A) \log Pr(A_i | H_{cc}) + n(\bar{A}) \log Pr(\bar{A}_i | H_{cc});$$

$$\log Pr(E_2 | H_{cc}) = n(B) \log Pr(B_i | H_{cc}) + n(\bar{B}) \log Pr(\bar{B}_i | H_{cc}).$$  (20)

These are maximized by a hypothesis $H_{cc}$ that has $Pr(A_i | H_{cc}) = f(A)$ and $Pr(B_i | H_{cc}) = f(B)$. That is, these terms are largest for hypotheses that posit probabilities for the $A_i$s and $B_i$s that are equal to the observed relative frequencies.

The mutual information of $E_1$ and $E_2$, conditional on a hypothesis $H_{cc}$, is

$$I(E_1, E_2 | H_{cc}) = n(AB) I(A_i, B_i | H_{cc}) + n(\bar{A}B) I(\bar{A}_i, B_i | H_{cc}) + n(\bar{A}\bar{B}) I(\bar{A}_i, \bar{B}_i | H_{cc}).$$  (21)

Once $Pr(A_i | H_{cc})$ and $Pr(B_i | H_{cc})$ are fixed, this is maximized by taking

$$Pr(A_i B_i) = f(AB).$$  (22)

Thus, in the expression (19) for the log-likelihood, we see that the first two terms reward hypotheses whose probabilities for $A_i$ and $B_i$ are close to the observed relative frequencies of these, and the last term, which corresponds to unification in the Mutual Information sense, rewards hypotheses with theoretical correlations close to the observed statistical correlations. What goes for log-likelihoods goes also for the evidential support $R$. Thus, when there is a difference between $f(AB)$ and $f(A)f(B)$, a common-cause hypothesis on which this difference is expected, by virtue of appropriate values of the parameters, counts as a MIUnifying hypothesis, and thereby achieves greater support.

For example, consider a case in which we have two sequences $\{A_i\}$, $\{B_i\}$, with a significant positive correlation between them: $f(AB)$ is much larger than $f(A)f(B)$. Consider two hypotheses, $H_{cc}$ and $H'_{cc}$, which posit the existence of sequences $\{C_i\}$ and $\{C'_i\}$, respectively, such that

$$Pr(A_i | H_{cc}) = Pr(A_i | H'_{cc}) \approx f(A);$$
$$Pr(B_i | H_{cc}) = Pr(B_i | H'_{cc}) \approx f(B).$$  (23)
Suppose, now that, $H_{cc}$ correctly predicts the correlations, but $H'_{cc}$ doesn’t. That is, $Pr(A_iB_i|H_{cc})$ is close to $f(AB)$, but $Pr(A_iB_i|H'_{cc})$ is not. In such a case we will have

$$MIU_1(\{E_1, E_2\}; H_{cc}) > MIU_1(\{E_1, E_2\}; H'_{cc}).$$

(24)

Thus, for appropriate values of the parameters, the hypothesis $H_{cc}$ affords MIUnification to the evidence set $\{E_1, E_2\}$, even though, in individual cases, the supposition $C_i$ does not render $A_i$ informative about $B_i$.

This does not prevent $C_i$ from being regarded as a common origin of $A_i$ and $B_i$. To take an example used by Lange in §3 of his paper, suppose that we take the clinical evidence to establish that some disease $C$ can cause symptoms $A$ and $B$. Then, if we observe $A$ and $B$ in some patient, this will raise our credence that $C$ also occurs in that patient, even if the symptoms $A$ and $B$ are independent, conditional on $C$. In such a case, the support provided by the symptoms $A$ and $B$ to the hypothesis that the patient has disease $C$ is just the sum of the supports given to the hypothesis by the individual items by themselves.

Lange raises the question of whether we should place more credence in a hypothesis that posits a single disease than in one that posits two independent origins of the symptoms $A$ and $B$. Suppose there are two other diseases $D_1$ and $D_2$, such that $A$ but not $B$ is a symptom of $D_1$, and $B$ but not $A$ is a symptom of $D_2$, and suppose further that the chance that a patient with $D_1$ exhibits symptom $A$ is the same as that of a patient with $C$, and that the chance that a patient with $D_2$ exhibits symptom $B$ is the same as that of a patient with $C$. Then, upon observation of both symptoms, the confirmational boost afforded to the hypothesis that the patient has $C$ is the same as the boost afforded to the hypothesis that the patient has both $D_1$ and $D_2$. The issue then comes down to priors. Is the joint occurrence of $D_1$ and $D_2$ much rarer than the occurrence of $C$? If the answer is yes—as would be the case if the three diseases are equally rare, and $D_1$ and $D_2$ uncorrelated—then we should place more credence in the hypothesis that the patient has $C$. If not—if the disease $C$ is so rare, and $D_1$ and $D_2$ so common that more patients contract both $D_1$ and $D_2$ than $C$—then our credences should favor the two-disease hypothesis. It would clearly be a mistake for one’s credences to favor the $C$-hypothesis merely on the basis of a preference for common origin explanations.
8 Conclusion

Mutual Information Unification is not the same as common origin explanation, and is neither a necessary nor sufficient condition for a hypothesis to play an explanatory role. Nevertheless, in a host of interesting cases, MIUnification is a concomitant of common origin explanation. Moreover, when a hypothesis that renders an otherwise puzzling coincidence comprehensible by providing a common origin explanation does receive an incremental confirmational boost from a body of evidence, beyond that provided by the individual items of evidence, that boost stems from MIUnification.

So, at least, is the verdict delivered by a Bayesian analysis; there is no room in Bayesian conditionalization for an extra confirmatory boost that is due to Common Origin Unification. A proponent of an explanationist thesis, to the effect that we ought to take hypotheses that involve common origin explanations to receive greater incremental support than hypotheses that achieve the same degree of Mutual Information Unification without explanation, should be in a position to explain why what is impossible for a Bayesian agent is rational for us. As we have seen, there is a limited heuristic role for considerations of Common Origin Unification, based on considerations of Type II rationality. It is doubtful whether any stronger explanationist thesis can be defended.

9 Appendix

Given a probability function $Pr$, and propositions $h, e_1, e_2$, define,

$$U_1 = \frac{Pr(e_1e_2|h)}{Pr(e_1|h)Pr(e_2|h)} \frac{Pr(e_1)Pr(e_2)}{Pr(e_1e_2)}; \quad (25)$$

$$U_2 = \frac{Pr(e_1e_2|h)}{Pr(e_1|h)Pr(e_2|h)} \frac{Pr(e_1)Pr(e_2)}{Pr(e_1e_2)} \frac{Pr(e_1|h)Pr(e_2|h)}{Pr(e_1e_2)} \frac{Pr(e_1)Pr(e_2)}{Pr(e_1e_2)}. \quad (26)$$

Then we have

$$MIU_1(e_1, e_2; h) = \log U_1; \quad (27)$$

$$MIU_2(e_1, e_2; h) = \log (U_1/U_2). \quad (28)$$

Thus, $MIU_1(e_1, e_2; h)$ is positive iff $U_1 > 1$, negative iff $U_1 < 1$, and zero iff $U_1 = 1$, and $MIU_2(e_1, e_2; h)$ is positive iff $U_1 > U_2$, negative iff $U_1 < U_2$, and zero iff $U_1 = U_2$. 

22
We want to show that each of the following four alternatives can be realized by some probability function.

1. $MIU_1 > 0$ and $MIU_2 > 0$; that is, $U_1 > 1$ and $U_1 > U_2$.
2. $MIU_1 > 0$ and $MIU_2 < 0$; that is, $1 < U_1 < U_2$.
3. $MIU_1 < 0$ and $MIU_2 > 0$; that is, $U_2 < U_1 < 1$.
4. $MIU_1 < 0$ and $MIU_2 < 0$; that is, $U_1 < 1$ and $U_1 < U_2$.

It is easy to show (see Lemma 1, below), that, if either $e_1$ or $e_2$ is irrelevant to $h$, then, if $U_1 > 1$, $U_2 < 1$, and vice versa. Thus, it is easy to construct examples that satisfy conditions 1 and 4. Take $Pr(e_1|h) = Pr(e_2)$. Then, on any probability function with $U_1 > 1$, we will have $U_2 < 1 < U_1$, and condition 1 will be satisfied. Similarly, if $Pr(e_2|h) = Pr(e_1)$, on any probability function with $U_1 < 1$, we will have $U_1 < 1 < U_2$, and condition 4 will be satisfied.

For condition 2, we need to have both $U_1$ and $U_2$ greater than 1. As is shown in Lemma 1, below, this is possible only if $e_1$ and $e_2$ are relevant to $h$ in opposite directions; that is, only if $R(h; e_1)$ and $R(h; e_2)$ have opposite sign. Here’s one way to do it. Take, for simplicity, $Pr(h) = Pr(e_1) = Pr(e_2) = 1/2$, and take $Pr(e_1 e_2) = 1/4$. Take $Pr(e_1|h) = 0.7$, $Pr(e_2|h) = 0.3$, and $Pr(e_1 e_2|h) = 0.24$. The reader can readily verify that these are consistent, and that they determine the full probability function on boolean combinations of \{h, e_1, e_2\}. In particular, they entail that $Pr(e_1|h) = 0.3$, $Pr(e_2|h) = 0.7$, and $Pr(e_1 e_2|h) = 0.26$. We thus have $U_1 = 24/21$ and $U_2 = 26/21$, satisfying the desired conditions.

For condition 3, we can take the probability assignment described in the previous paragraph and create a new one by interchanging $e_2$ and $\bar{e}_2$. We have, once again, $Pr(h) = Pr(e_1) = Pr(e_2) = 1/2$, $Pr(e_1 e_2) = 1/4$, $Pr(e_1|h) = 0.7$, and $Pr(e_1\bar{h}) = 0.3$. We also have $Pr(e_2|h) = 0.7$, and $Pr(e_1 e_2|h) = 0.46$. These further entail that $Pr(e_2|\bar{h}) = 0.3$, and $Pr(e_1 e_2|\bar{h}) = 0.04$. We thus have $U_1 = 46/49$, and $U_2 = 4/9$, and so $U_2 < U_1 < 1$, and condition 4 is satisfied.

Having shown that all four alternatives are possible, we now prove the Lemma alluded to above.

**Lemma 1.** Let \{h, e_1, e_2\} be logically independent propositions, and let $Pr$ be a probability function on the boolean algebra generated by this set. We assume that the denominators of the relevant fractions are nonzero, and define $U_1$ and $U_2$ as above.

23
a) If \( Pr(h|e_1) = Pr(h) \) or \( Pr(h|e_2) = Pr(h) \), then, if \( U_1 > 1 \), \( U_2 < 1 \), and vice versa.

b) If \( U_1 \) and \( U_2 \) are both less than one, then either \( e_1 \) and \( e_2 \) are both positively relevant to \( h \), or they are both negatively relevant to \( h \).

c) If \( U_1 \) and \( U_2 \) are both greater than one, then one of \( \{e_1, e_2\} \) is positively relevant to \( h \), and the other negatively relevant.

Proof. Let

\[
p = Pr(h); \quad q = Pr(\bar{h}) = 1 - p;
\]
\[
\alpha_1 = Pr(h|e_1)/Pr(h); \quad \alpha_2 = Pr(h|e_2)/Pr(h);
\]
\[
\beta_1 = Pr(\bar{h}|e_1)/Pr(\bar{h}); \quad \beta_2 = Pr(\bar{h}|e_1)/Pr(\bar{h}).
\] (29)

This allows us to write

\[
U_1 = \frac{1}{\alpha_1 \alpha_2} \frac{Pr(e_1 e_2|h)}{Pr(e_1 e_2)}; \quad U_2 = \frac{1}{\beta_1 \beta_2} \frac{Pr(e_1 e_2|\bar{h})}{Pr(e_1 e_2)}.
\] (30)

Once \( p, \alpha_1, \alpha_2, \beta_1, \) and \( \beta_2 \) are fixed, this yields a constraint on \( U_1 \) and \( U_2 \):

\[
p \alpha_1 \alpha_2 U_1 + q \beta_1 \beta_2 U_2 = 1. \] (31)

It is convenient to write this in terms of a weighted average of \( U_1 \) and \( U_2 \). Define

\[
w_1 = \frac{p \alpha_1 \alpha_2}{p \alpha_1 \alpha_2 + q \beta_1 \beta_2}; \quad w_2 = \frac{q \beta_1 \beta_2}{p \alpha_1 \alpha_2 + q \beta_1 \beta_2}.
\] (32)

Then (31) becomes,

\[
w_1 U_1 + w_2 U_2 = \frac{1}{p \alpha_1 \alpha_2 + q \beta_1 \beta_2}, \] (33)

with \( w_1 \) and \( w_2 \) both nonnegative, and

\[
w_1 + w_2 = 1. \] (34)

It is instructive to rewrite the right-hand side of (33), using the fact that \( p \alpha_1 + q \beta_1 = p \alpha_2 + q \beta_2 = 1 \). A bit of algebraic manipulation yields,

\[
w_1 U_1 + w_2 U_2 = 1 - \frac{pq(\alpha_1 - \beta_1)(\alpha_2 - \beta_2)}{p \alpha_1 \alpha_2 + q \beta_1 \beta_2}.
\] (35)
From (35) it is readily apparent that, if either $e_1$ or $e_2$ is irrelevant to $h$—that is, if $\alpha_1 = \beta_1$ or $\alpha_2 = \beta_2$, then

$$w_1 U_1 + w_2 U_2 = 1,$$  \hspace{1cm} (36)

and in such a case, if $U_1 > 1$, then $U_2 < 1$, and vice versa. If we want to construct a case in which $U_1$ and $U_2$ are both greater than one, this requires the right-hand side of (35) to be greater than one, which means that $\alpha_1 - \beta_1$ and $\alpha_2 - \beta_2$ must have opposite sign: one of $\{e_1, e_2\}$ must be positively relevant to $h$, and the other negatively relevant. If we want to construct a case in which $U_1$ and $U_2$ are both less than one, then $\alpha_1 - \beta_1$ and $\alpha_2 - \beta_2$ must have the same sign: $e_1$ and $e_2$ are either both positively relevant, or both negatively relevant, to $h$.

10 Acknowledgments

I thank Michel Janssen, Marc Lange, Bill Harper, and Molly Kao for helpful discussions. I am grateful to Clark Glymour for raising the question, addressed in §7, of how common-cause explanations fit into the framework. This work was supported, in part, by a grant from the Social Sciences and Humanities Research Council of Canada (SSHRC).
References


