NAGELIAN REDUCTION
AND COHERENCE

Philippe van BASSHUYSEN*

Abstract: It can be argued (cf. Dizadji-Bahmani et al. 2010) that an increase in coherence is one goal that drives reductionist enterprises. Consequently, the question if or how well this goal is achieved can serve as an epistemic criterion for evaluating both a concrete case of a purported reduction and our model of reduction: what conditions on the model allow for an increase in coherence? In order to answer this question, I provide an analysis of the relation between the reduction and the coherence of two theories. The underlying model of reduction is a (generalised) Nagelian model (cf. Nagel 1970, Schaffner 1974, Dizadji-Bahmani et al. 2010). For coherence, different measures have been put forward (e.g. Shogenji 1999, Olsson 2002, Fitelson 2003, Bovens & Hartmann 2003). However, since there are counterexamples to each proposed coherence measure, we should be careful that the analysis be sufficiently stable (in a sense to be specified). It will turn out that this can be done.

Keywords: Nagelian reduction, Coherence, Bayesian coherence measures, Bayesian networks, Bayesian analysis

I. MOTIVATION AND OUTLINE.

It can be argued (cf. Dizadji-Bahmani et al. 2010) that one goal that drives reductionist enterprises is the coherence of theories. Suppose we have two theories, T_P and T_F, that share – at least partially – one domain of applicability. There is a danger that T_F and T_P don't cohere. What we would expect from a *reduction* from T_P to T_F is that it gives us a proof of their coherence. If it doesn't – so much the worse for our concept of intertheoretic reduction.

Two questions arise: first, what happens to the coherence of theories in reductions? Moreover, we can use coherence as a touchstone for epistemically evaluating reductions. What conditions has a reduction to satisfy in order to increase the coherence of T_F and T_P?

* Philippe van BASSHUYSEN, London School of Economics and Political Science, Department of Philosophy, Logic and Scientific Method. Email: philippe.v.basshuysen@gmail.com

RRfA

In order to answer these questions, the concept of coherence shall be considered first (§ 2). Various measures of coherence have been proposed and are presented in § 2.1, and a sense shall be given in § 2.2 of how problematic it indeed is to give the concept a precise meaning in agreement with our intuitions. In § 3, I present what Dizadji-Bahmani et al. in (2010) call the „generalised" Nagel-Schaffner model of intertheoretic reduction. In § 4, the two concepts run together: the coherence measures are applied to a reductive context in order to see what the conditions on a reduction are in order for the coherence of the theories to increase. This will shed a light on the logic of reductions.

II. THE CONCEPT OF COHERENCE.

Coherence comes in degrees (unlike, for example, consistency). Consequently, coherence is normally seen as inducing an order on sets of propositions („information sets"); moreover, it is usually treated probabilistically.

Coherence is an important concept, for example for theory choice in science (cf. Kuhn 1977, Bovens & Hartmann 2003, 53 et sqq.), or, as claimed here, as one goal that drives reductionist enterprises. However, I claim that it is not a concept whose everyday meaning is to be made precise. There are numerous proposals to measure the degree of coherence of information sets, each of them giving some insight into the notion, but there is not one „correct" coherence measure: for each of them, one can construct examples in which the measure in question yields counterintuitive results.

I suggest the situation is rather like the case of Bayesian confirmation measures, and should receive an analogous treatment Fitelson gave for confirmation measures (1999). This gives the following picture: different coherence measures may yield different results in different situations; this is not a dilemma as long as it is secured that either the respective application is sufficiently stable (i.e. insensitive to the choice of the measure) or, if this is not the case, arguments must be given why the chosen coherence measure(s) are to be preferred to other measures in the application in question. This should be kept in mind when investigating what happens to the coherence of theories in reductions.

II.1. COHERENCE MEASURES.

In what follows I give a list of coherence measures (which I don't claim is exhaustive). Despite the advertised pluralism of coherence measures, I think that Bovens' and Hartmann's (cf. 2003) approach is distinct in that it yields more philosophical insight into the nature of coherence: rather than tentatively putting forward a measure, they develop it axiomatically. The presentation of their measure will therefore occupy more space than the other ones.

Call a statement R_i a partially reliable[1] source i provides us with an *information item*. The information set $S = \{R_1,...,R_m\}$ consists of information items from m independent sources; the same for n independent sources in $S' = \{R'_1,...,R'_m\}$. Let P be a probability distribution over information sets; let C be the relation „not less coherent than". The following measures of coherence $m_x(...)$ have been proposed.

<u>Shogenji (cf. 1999).</u> SCS' iff

$$m_{Sh}(S) = \frac{P(R_1,\ldots,R_m)}{\prod_{i=1}^{m}P(R_i)} \geq \frac{P(R'_1,\ldots,R'_n)}{\prod_{i=1}^{n}P(R'_i)} = m_{Sh}(S').$$

<u>Olsson (cf. 2002).</u> iff

$$m_O(S) = \frac{P(R_1,\ldots,R_m)}{P(R_1 \vee \ldots \vee R_m)} \geq \frac{P(R'_1,\ldots,R'_n)}{P(R'_1 \vee \ldots \vee R'_n)} = m_O(S'),$$

$\vee$ standing for inclusive disjunction.

<u>Fitelson (cf. 2003).</u> Consider the two-item information sets $\{R_i, R_j\}$ and $\{R'_i, R'_j\}$. Fitelson uses the Kemeny-Oppenheim measure of factual support for his measure. For example,

$$F(R_i, R_j) = \frac{P(R_j \mid R_i) - P(R_j \mid \neg R_i)}{P(R_j \mid R_i) + P(R_j \mid \neg R_i)}$$

(defined for $P(R_j) < 1$ and $P(R_i) > 0$) denotes the factual support R_j provides for R_i. What Fitelson then measures is the average factual support between the propositions of the information set: $\{R_i, R_j\}C\{R'_i, R'_j\}$ iff

$$m_F(\{R_i, R_j\}) = \frac{F(R_i, R_j) + F(R_j, R_i)}{2} \geq \frac{F(R'_i, R'_j) + F(R'_j, R'_i)}{2} = m_F(\{R'_i, R'_j\}).$$

In general, for all information sets S, the measure takes the mean value of the factual support all $R_i \in S$ receive from all $x \in \{\mathcal{P}(S): R_i \notin x\}$. In the applications of Fitelson's measure the calculations tend to become quite large: there are $n \times (2^{n-1} - 1)$ cases to consider for information sets with n elements. For example, for $S = \{R_1, R_2, R_3\}$ there are 9 cases to consider ($\{F(R_1, R_2, R_3), F(R_1, R_3), F(R_1, R_2 \& R_3), ...\}$) whose mean value defines the coherence of S.

„Weak Bayesian Coherentism" (cf. 2003, *NA*). Bovens and Hartmann claim that what has gone wrong in all the tentative approaches to measure coherence is that they invoke certain correct intuitions that usually only account for one aspect of this complex notion. For example, the Olsson measure is sensitive only to the relative agreement between the propositions in

---

[1] A feature that will be of importance in the Bovens & Hartmann approach. In the formalism I closely follow their exposition in 2003, 30 et sqq.

RRfA

question (to be explained below) but neglects the fact that more information can increase the coherence while even diminishing the relative agreement of the propositions. The idea of Bovens' and Hartmann's axiomatic approach is to instead look at the *role* that coherence plays: the role of increasing our confidence in an information set –- the subjective probability, that is. The first axiom of „Weak Bayesian Coherentism" is thus:

• For all information sets S, S', if S is no less coherent than S', then our degree of confidence that the content of S (i.e. the conjunction of the propositions in S) is true is no less than our degree of confidence that the content of S' is true, *ceteris paribus*.

In order to measure coherence, it must be taken into account that there are other factors that play a role in determining our confidence in information sets, namely, how expected the transmitted information is and how reliable its sources are. We need to abstract from these factors; the ceteris paribus-clause thus says that *how expected the information is and how reliable the sources are* does not vary from information set to information set.

The second axiom takes into account the probabilistic nature of coherence:

• A partial coherence-order over a set of information sets is fully determined by the probabilistic features of its information sets.

The order is partial because of an impossibility result: it turns out that some pairs of information sets are not comparable with respect to their coherence as defined in the B&H way without reference to the reliability of the sources. Bovens' and Hartmann's reaction is to constrain the application of their coherence measure to pairs of information sets for which the measure yields a definite result for *all possible values* of the reliability of the sources (hence, the addendum „weak" to their „Bayesian Coherentism").

Their strategy then is the following. Our confidence that some information set is true is maximal when the information is *maximally coherent* – only containing logically equivalent propositions –, *ceteris paribus*. This leads B&H to define the coherence of an information set as the ratio of the confidence boost we receive from it over the confidence boost we would receive from the same information but in a maximally coherent information set (cf. 2003, 33).

This can be made precise: let R_n denote the propositional variable that can take two values, viz. R_n and $\neg R_n$. Let u_i be the sum of all the joint probabilities of the instances of i negative values and $n - i$ positive values of the variables $R_1, ..., R_n$ (cf. 2003, 17). For example, for the information set $S = \{R_1, R_2\}$, we have $u_0 = P(R_1, R_2)$, $u_1 = P(R_1, \neg R_2) + P(\neg R_1, R_2)$, and $u_2 = P(\neg R_1, \neg R_2)$.

A maximally coherent information set $S = \{R_1, ..., R_n\}$ can formally be defined as an information set with the specific distribution of u:

$$< u_0 = r, u_1 = 0, ..., u_{n-1} = 0, u_n = 1 - r >$$

– since its propositions are logically equivalent, either all the propositions are true or all the propositions are false.

R Rℱ A

It can be shown (cf. Bovens & Hartmann 2003, Appendix A.1), given the constraints that the sources be equally partially reliable and independent, that under the following definition of the coherence measure the partial order of coherence defined with the help of it satisfies the axioms (i) and (ii):

SCS' iff

$$m_{B\&H}(S) = \frac{u_0 + (1-u_0)(1-r)^m}{\sum_{i=0}^{m} u_i(1-r)^i} \geq \frac{u_0' + (1-u_0')(1-r)^n}{\sum_{i=0}^{n} u_i'(1-r)^i} = m_{B\&H}(S'),$$

for all values of $r \in (0,1)$.

$(1 - r)$ is the reliability parameter; it tells us how reliable the sources of the information are: the bigger $(1 - r)$ the more reliable the source of the information is (so the smaller $r \in (0,1)$ the more reliable the source is). What the reliability parameter measures can formally be made precise.[2] However, this is not important for our purpose. For when comparing information sets with the same number of information items there is a sufficient condition for the comparability of these information sets in which $(1 - r)$ cancels out; and in the analysis of coherence in reductive contexts, this will be enough (for the information sets in question – the set of two scientific theories before and the set of the very same scientific theories after a reduction – will always have the same cardinality). In Appendix B.1 in (Bovens & Hartmann 2003) it is proved that the following proposition follows from the above definition:

For two information sets S and S' with the same number of information items it is a sufficient condition for $S'CS$ that

(a) $u_0 \leq u_0'$ and $u_i \geq u_i'$ for all $i \in 1,..., n- i$ or

(b) $u_0 \geq u_0'$ and $\dfrac{u_i}{u_i'} \geq \dfrac{u_0}{u_0'}$ for all $i \in 1,..., n- i$ (cf. Boven & Hartmann 2003, 37).

For information sets containing only two information items, (a) or (b) are also necessary for $S'CS$.

II.2. DIGRESSION: AN EXAMPLE.

To get a feeling for the measures, but especially to underline my thesis that each of the measures yields counterintuitive results in some situations, let's consider a simple example.[3] In Bovens & Hartmann 2003, there are

---

[2] In the line of the following (cf. Bovens & Hartmann 2003). Let p denote the probability that the source gives a report to the effect that R_i is the case, given that $R_{i'}$ and let q denote the probability that the source gives a report to the effect that R_i is the case given that $\neg R_{i'}$ for $i = 1,..., n$. It is assumed that the sources are not fully reliable, i.e. $p \neq 1$ but that they are more or less reliable, i.e. $p > q$. Then we define $r := \dfrac{q}{p}$. By our assumptions $p \in (0,1)$ and $p > q$, $r \in (0,1)$.

[3] For readers not interested in the conceptual riddles of coherence, this section can be skipped.

RRfA

counterexamples to all the coherence measures except their own; so let's fill the gap.[4]

Someone killed Francois. There are 10 suspects for the crime, some of which definitely R_1 = had a motive to kill Francois. The detective knows this about five of them. He also knows that five suspects R_2 = are Francophile; and that only one is both R_1 and R_2: Francophile and interested in Francois's death. Nothing else is known about any of the suspects.

Now consider <u>Situation 1:</u> the detective receives two different reports from two partly and equally reliable witnesses the first one claiming that the killer (k) was one of the suspects we already know have a motive to kill Francois and the second one claiming that the killer was Francophile. The information set $S = \{R_1k, R_2k\}$ is, of course, not very coherent under our assumption that the overlap of Francophiles and suspects interested in Francois getting killed (i.e. the relative agreement of R_1k and R_2k) is very small. The Boolean algebra in Fig. 1 shows the situation the detective faces.
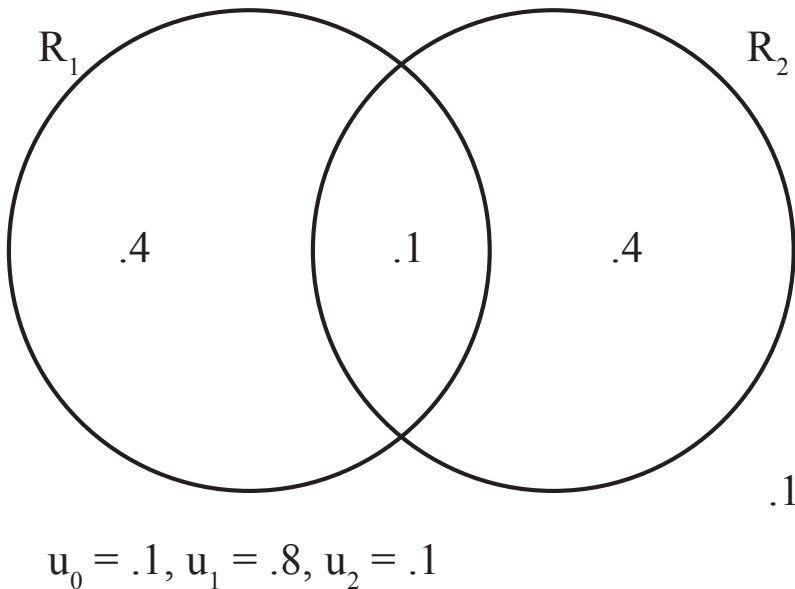

$$u_0 = .1, \ u_1 = .8, \ u_2 = .1$$

Figure 1: Boolean algebra representing
the initial information in <u>situation 1</u>.

---

[4] Finishing this investigation, I found that Douven and Meijs come up with an alleged counter-example to Bovens' and Hartmann's measure in (2005), and Bovens and Hartmann reject their critique in (2005). However, my counterexample is stronger in the following sense: Douven and Meijs state two information sets S' and S one of which is intuitively clearly more coherent than the other one but in the *B&H* measure they are incomparable. In my first counterexample, one information set S'' is intuitively more coherent than another one S''' but the measure yields that $S'''CS''$. In consideration of this counterexample, the measure is harder to salvage.
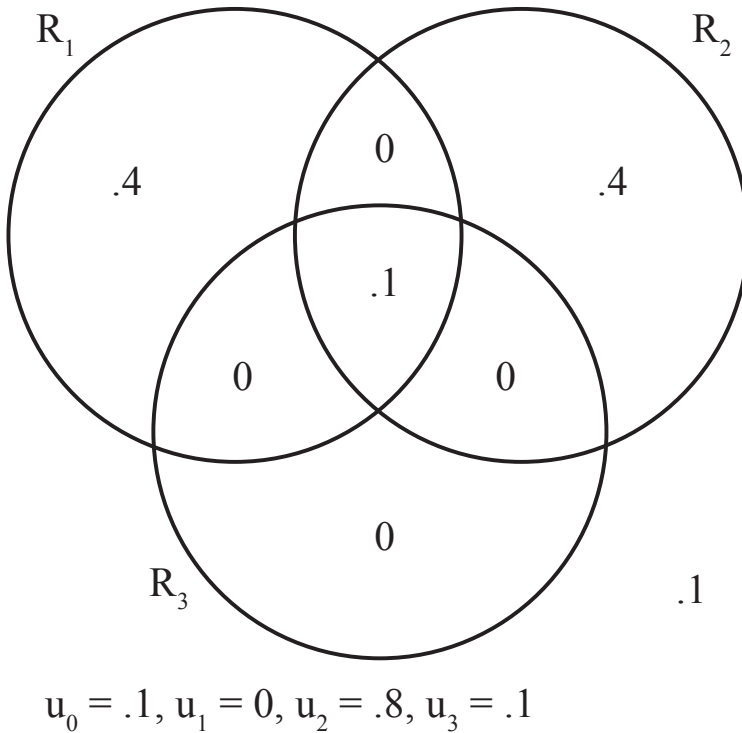
$$u_0 = .1, \ u_1 = 0, \ u_2 = .8, \ u_3 = .1$$

Figure 2: Extending the information set
in <u>situation 1</u>.

During his investigation, the detective finds out that Francois had an affair with one of the suspect's wives (let R_3 = to have a wife Francois had an affair with); namely, the suspect that happens to be both interested in Francois's death and Francophile. A third witness appears, equally reliable as the other two, claiming that the killer is just the suspect whose wife had an affair with Francois. The Boolean algebra in Fig. 2 shows the new information set $S' = \{R_1 k, R_2 k, R_3 k\}$ the detective faces.

Applying the Bovens&Hartmann measure yields that the difference of coherence between S' and S is positive:

$$m_{B\&H}(S') = \frac{.1 + .9 \times (1-r)^3}{.1 + .8 \times (1-r)^2 + .1 \times (1-r)^3} > \frac{.1 + .9 \times (1-r)^2}{.1 + .8 \times (1-r) + .1 \times (1-r)^2} = m_{B\&H}(S)$$

for all $0 < r < 1$. Thus, the extra information the detective receives makes the information set more coherent. This can be depicted as a function on r, as in Fig. 3.
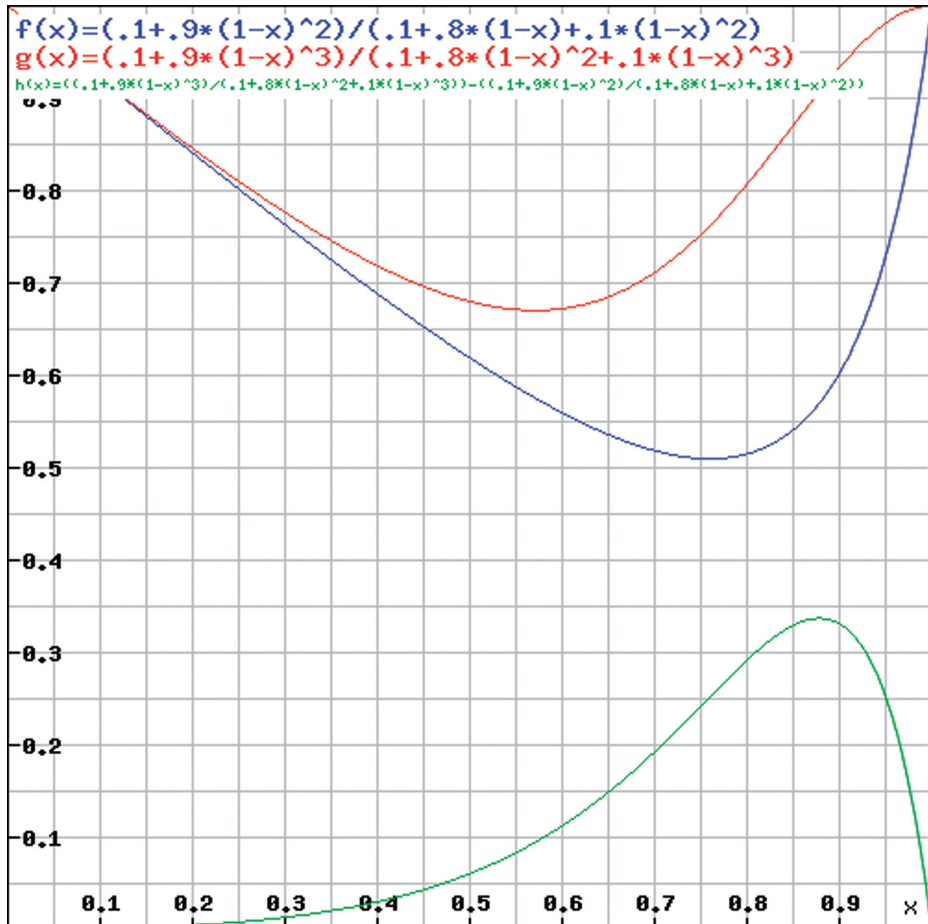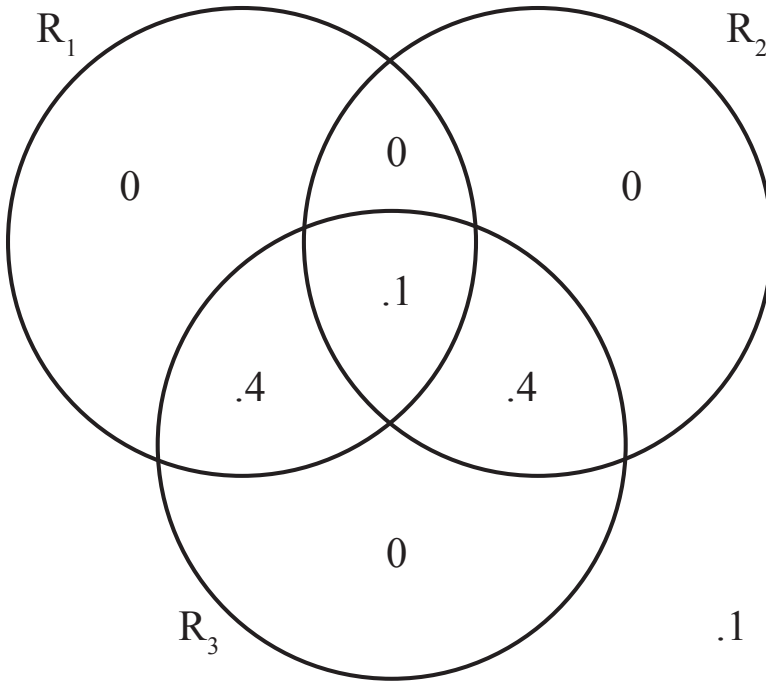
Figure 3: The Bovens&Hartmann measure yields
an intuitive result in <u>situation 1</u>: the value of the extended information set
S' (red graph, $g(x)$) is higher than that
of the original information set S (blue graph, $f(x)$),
for all values of the reliability parameter (here called x).
The green graph is the difference $g(x) - f(x)$,
which is strictly positive for $x \in (0,1)$.

Thus, the measure yields quite an intuitive result in situation 1. But now consider <u>situation 2</u>: suppose Francois had an affair with almost all of the suspects' wives, as in the Boolean algebra in Fig. 4. Just like in situation 1, the detective is in the state of knowledge shown in Fig. 1 at first; he then receives a report to the effect that Francois had an affair with his killer's wife, so the new information set again is $S' = \{R_1k, R_2k, R_3k\}$.

$$u_0 = .1, u_1 = .8, u_2 = 0, u_3 = .1$$

Figure 4: <u>Situation 2</u>: a counterintuitive result
in the Bovens&Hartmann measure.

Intuitively, the information set S' is, as in situation 1, more coherent than S, since it is identical to S but contains another information item, R_3, which is just the union of R_1 and R_2. The new witness report does not add much information (although it does add information, and our detective is of course interested in obtaining this information). One would expect this information set to be more coherent than S and less coherent than S' in situation 1. Applying Shogenji's and Fitelson's measures yields results coinciding with this intuition (I omit to give the calculations here), whereas Olsson's measure is completely insensitive to all three situations. In contrast, applying the Bovens&Hartmann measure, in situation 2 S' has *lower* values than S:

$$m_{B\&H}(S') = \frac{.1 + .9 \times (1-r)^3}{.1 + .8 \times (1-r) + .1 \times (1-r)^3} < \frac{.1 + .9 \times (1-r)^2}{.1 + .8 \times (1-r) + .1 \times (1-r)^2} = m_{B\&H}(S)$$

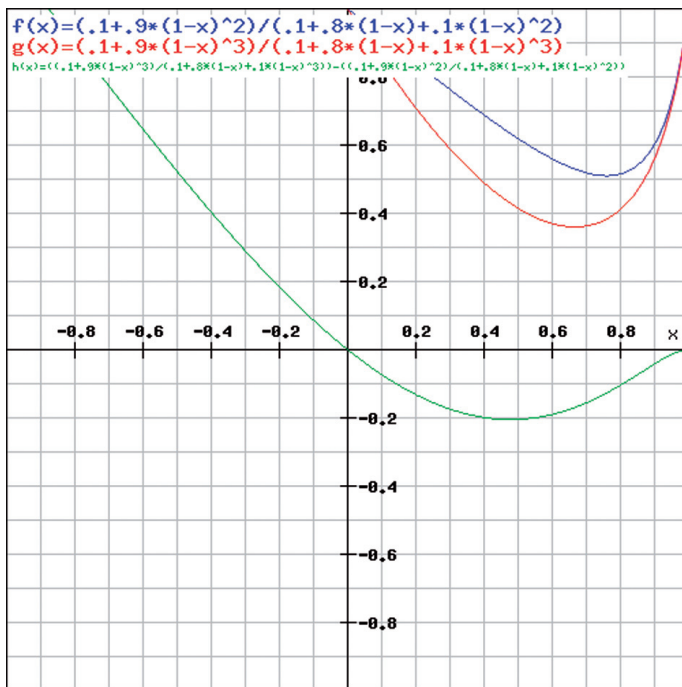for all $0 < r < 1$. This can be seen in the graph in Fig. 5.

f(x)=(.1+.9*(1-x)^2)/(.1+.8*(1-x)+.1*(1-x)^2)
g(x)=(.1+.9*(1-x)^3)/(.1+.8*(1-x)+.1*(1-x)^3)
h(x)=((.1+.9*(1-x)^3)/(.1+.8*(1-x)+.1*(1-x)^3))-((.1+.9*(1-x)^2)/(.1+.8*(1-x)+.1*(1-x)^2))

Figure 5: <u>Situation 2:</u> $g(x) < f(x)$ for $x \in (0,1)$.
A counterintuitive result.

The B&H measure thus yields a counterintuitive result in situation 2. We can also ask what happens when applying the measure to Boolean algebras that assign more and more information to R_3, i.e. if we go towards situation 1 with $R_3 \cap R_1 \setminus R_2 = \varnothing$ and $R_3 \cap R_2 \setminus R_1 = \varnothing$; some of these Boolean algebras are shown in Fig. 6.

$u_0 = .1, u_1 = .7, u_2 = .1, u_3 = .1$

72
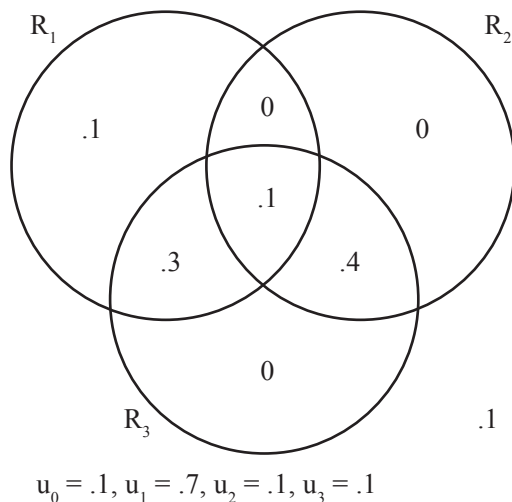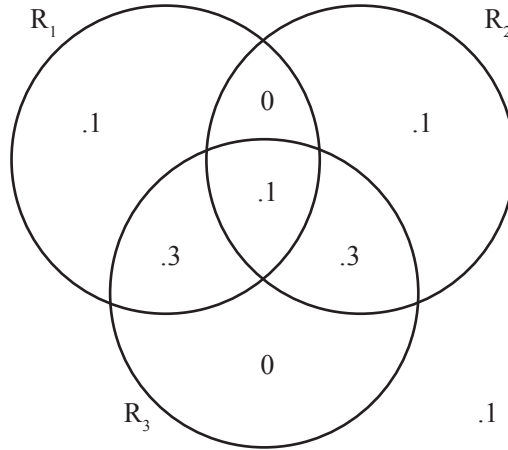
$u_0 = .1, u_1 = .6, u_2 = .2, u_3 = .1$
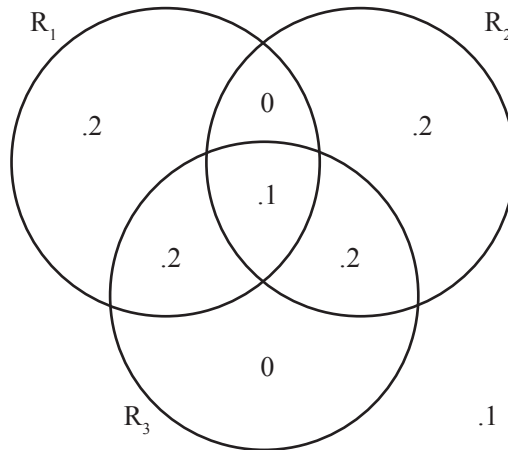
$u_0 = .1, u_1 = .4, u_2 = .4, u_3 = .1$

Figure 6: Situations not comparable to S
in the Bovens&Hartmann definition of coherence.

We would expect that the information set gets more and more coherent until the extreme case of situation 1. This is what the Shogenji and the Fitelson measures indeed tell us. However, in the Bovens&Hartmann measure, what happens is that the coherence of the information sets thus produced compared to each other does increase, but the information sets are not comparable to S: for different values of the reliability parameter r, one curve is sometimes above and sometimes below the other curve. They are incomparable to S until we reach the extreme case of situation 1. This is another counterintuitive result.

RR*f*A

By the way, it is hard to come up with a counterexample to the Bovens & Hartmann measure in which the two information sets in question have the same number of information items. This might hint at where the problem lies.

Needless to say, all this depends on intuitions; and they may vary quite a bit. On the other hand, the goal of the coherence measures is precisely to make our intuitions about coherence precise. There are clear cases where our intuitions abandon us in deciding which of two information sets is more coherent (cf. Bovens & Hartmann 2003), and it is nice to have an impossibility result as in the Bovens&Hartmann measure and thus to impose a *partial* order on sets of information sets. However, the Bovens&Hartmann measure does not seem to correctly draw the line (as seen in the second counterexample).

Again: the conclusion to draw from this is, in my opinion, that we must ensure that in our applications of the measures the results must be sufficiently stable. This is not very precise, but it turns out that in the application to reductions this is clearly the case.

III. NAGELIAN REDUCTION.

My presentation of Nagelian reduction is guided by the model Dizadji-Bahmani et al. proposed in (2010).[5] In the same paper, arguments can be found to the effect that this is indeed the *correct* model of intertheoretic reduction.

The general, well-known idea of Nagelian reduction is that what constitutes a reduction is a logical derivation of the laws of the theory to be reduced from the laws of the reducing theory. However, there are hardly any two scientific theories for which each law of the one can be derived from the laws of the other: an approximation is all there can normally be found.

Let T_F be the set of the law-statements of the fundamental theory in question and let T_P be the set of the law-statements of the phenomenological theory in question. Notice a notational ambiguity: I use T_F to denote both the fundamental theory and the set of its laws; the same for T_P, and T_F^*, T_P^*, which are yet to be defined (this is adopted from Dizadji-Bahmani et al. 2011, 323). If one holds the view that a theory is more than a set of laws there should be two different signs for the two. This is, however, no question with which the present analysis is concerned, and since nothing hinges on it, an innocent sloppiness.

A sufficient condition for a successful reduction from T_F to T_P in the line of Nagel (cf. 1970) is that there is a set of laws T_P^* for which it holds that

---

[5] Bahmani et al. (2010), (2011) call it the *Generalised* Nagel-Schaffner Model of Reduction, GNS. Besides giving a very neat analysis of it I see, however, no substantive differences between GNS and Schaffner's model as in (1974) or even Nagelian reduction of the later Nagel (e.g. 1970, 362 et seqq.), and thus no need for a change of labels.

(i) each of its members is [derivable] from T_F together with boundary conditions, and

(ii) its laws are [strongly analogous] to T_P.

This defines a *homogeneous reduction*. Nagel (cf. 1970) was aware of the fact that the theory to be reduced may contain theoretical concepts that do not occur and are not definable in the reducing theory (for example, „entropy" as a concept of thermodynamics, is not definable in statistical mechanics (cf. Nagel 1970, 912)). If this is the case, [bridge laws] are required to connect the concepts; Nagel called this case *inhomogeneous reduction*. Inhomogeneous reductions can be defined substituting (i) with (i′) in the above definition:

(iii) there is a set T_F^* each of whose members is [derivable] from T_F together with boundary conditions and whose members via [bridge laws] constitute T_P^*.

The establishment of a reduction is illustrated in Figures 7 and 8. In Fig. 7, no connection is known between T_F and T_P. Fig. 8 shows a heterogeneous reduction: we have a mediating theory T_F^* derived from T_F whose essential terms are connected with another mediating theory T_P^* via bridge laws, and T_P^* is strongly analogous to T_P. If the reduction in question is homogeneous, we can conceive of T_F^* and T_P^* to be identical. Thus, the model of inhomogeneous reduction is general enough to cover both homogeneous and inhomogeneous reductions, and the definition {(i′)(ii)} gives a necessary and sufficient condition for the establishment of a Nagelian reduction.

Figure 7: Theories T_F and T_P before reduction.
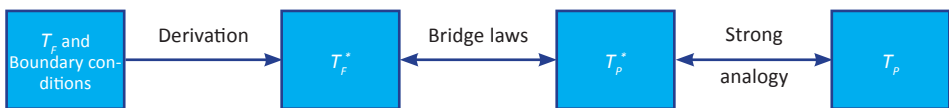

Figure 8: Nagelian reduction as illustrated in (Dizadji-Bahmani et al. 2010).

The connecting notions of [bridge laws] and [strong analogy] are intuitive but not precise and thus, problematic. Nagel (1970) claims that [bridge laws] can either express the identity of entities or relations between properties. Following Nagel and Schaffner (e.g. Nagel 1970), Dizadji-Bahmani et al. (2010) require for a [strong analogy] that T_P^* and T_P share the same essential concepts, and that T_P^* be at least as empirically adequate as T_P. The last desideratum follows from the fact that T_P^* is in fact a *corrected* version of T_P (cf. Nagel 1970, 362 et sqq.).

RRƒA

However, [bridge laws] and [strong analogy] really call for further (formal) refinement. This investigation into what happens to the coherence of two theories when one gets reduced to the other one will shed a light particularly on the critical notion of strong analogy (cf. § 5). If our claim that coherence works as one epistemic touchstone of reductions is correct, it will follow from our analysis that a necessary condition on a successful reduction is that T_P^* and T_P be positively probabilistically related, viz. that $P(T_P \mid T_P^*) > P(T_P \mid \neg T_P^*)$.

III.1. REPRESENTING NAGELIAN REDUCTION PROBABILISTICALLY.

The model of intertheoretic reduction presented in the preceding paragraph is the model of reduction that underlies this investigation; in the following, „reduction" shall just refer to this Nagelian model of intertheoretic reduction. Again, I refer to (Dizadji-Bahmani et al. 2010) for arguments to the effect that Nagel's is in fact the correct model of reduction.

A reduction establishes a particular relation between two scientific theories. It has already been said that this relation is more complex than a mere logical derivation of the laws of one theory from the laws of the other theory. As in other contexts too complex (arguably) to be represented in a purely qualitative framework (e.g. the context of confirmation), it seems to be natural to apply probability theory to model reduction.

For our purpose, it will be convenient to represent the model of Nagelian reduction with the help of Bayesian networks. Bayesian networks are graphical models originally applied to expert systems in contexts of causal reasoning. In our context, no causality is involved, but that doesn't matter: for *any* probability function over a set of random variables can be represented in terms of Bayesian networks (cf. Williamson 2005). They are an effective representation of a joint probability distribution over a set of random variables; moreover, probabilistic (in)dependencies can easily be read off this graphical representation of a probability function.

Bayesian networks are directed acyclic graphs. They consist of three things (for a detailed presentation of Bayesian networks, cf. (Jensen 2000), particularly chapter 2):

(1) a finite set of nodes which represent discrete random variables X, Y, … – in our case propositional variables each of which can take two values, e.g. X and $\neg X$;

(2) a set of arrows between nodes such that no cycles occur; and

(3) a probability table that gives the probabilities of the values of each variable X conditional on all the combinations of all the values of X's parents. For variables in root nodes, the probability table gives the unconditional probabilities of their values.

RRfA

A few more words about Bayesian networks and causality. Usually the arrows are interpreted as causal links between variables: assume X is represented by a parent node; then the value of X has causal impact on the values taken by the variables $Y_1...Y_n$ in X's child nodes. However, for our purpose, an arrow just denotes that a change of the value of X may change the probabilities of the values of $Y_1...Y_n$. This doesn't imply any attitude towards causality.

Now our key problem is to find a Bayesian network to represent the right target probability function to model reduction. In the preceding paragraph we developed graphical models of two theories pre and post reduction; let's try to convert them into Bayesian networks. First, we define the propositional variable T_F: T_F can take the values $T_F = y$ meaning the conjunction of the propositions in T_F is true; and $T_F = n$ meaning the conjunction of the propositions in T_F is false. $X = y/n$ is standard notation when working with Bayesian nets, but let me introduce another innocent sloppiness (cf. § 3) and just write T_F and $\neg T_F$ for $T_F = y$ and $T_F = n$, respectively – this is convenient because we are only concerned with propositional variables. We define the propositional variables T_P and (for the situation after the reduction) T_F^* and T_P^* in the same way as T_F. These propositional variables define the nodes of our networks.

Let P be the probability function pre-reduction, and P' the probability function post-reduction. Let us make a few simplifying assumptions about the probability functions that could be dropped subsequently. Suppose T_F and T_P are probabilistically independent before the reduction. Second, assume that the prior probability of the fundamental theory is the same pre and post reduction. This makes the comparison of the coherence pre and post reduction easier. In the following section, I will at some points also assume that the prior probability of the phenomenological theory be the same before and after the reduction; it will be made explicit when this is the case. Finally, assume that the [bridge laws] state a perfect correlation (see below).

Consider the situation before the reduction first. We assumed that T_F and T_P are probabilistically independent. We can represent this assumption in a Bayesian network with two nodes T_F and T_P and no arrows between the nodes. The probabilities of the theories are two constants $P(T_F) = a$ and $P(T_P) = b$. We assume in the entire course of this investigation that $a, b \in (0,1)$. This is an unproblematic assumption idle to argue for: theories with zero probability are not interesting.

Now consider the situation after the reduction. Clearly, for each connecting notion – [derivation], [bridge laws], and [strong analogy] – there is an arrow because these notions establish a relation between the respective variables which we want to express probabilistically. These are also *all* the arrows because all the conditions that constitute a reduction are taken into account; and we haven't assumed that there are relations apart from these. For example, we haven't assumed a direct relation between T_F and T_P. Now *how* to draw the arrows?

RRfA

First, we need to represent the [derivation] from T_F to T_F^* in our Bayesian network. Since it follows from Kolmogorov's axioms that for any probability function P and propositions A, B, if $A \models B$ then $P(B|A) = 1$, we draw an arrow from T_F to T_F^* and write $P'(T_F^*|T_F) = 1$ and $P'(T_F^*|\neg T_F) = x$ with $x \in [0,1]$.

Second, the bridge laws: we assume that they state a perfect correlation, that is „you can have both or neither" of T_F^* and T_P^*: they are logically equivalent. Logical equivalence can be modeled by an arrow in either direction together with suitable entries in the probability table; for example, Fig. 9 shows a Bayesian network with an arrow from T_F^* to T_P^* and the entries $P'(T_P^*|T_F^*) = 1$ and $P'(T_P^*|\neg T_F^*) = 0$. Dizadji-Bahmani et al. (2011) argue that this is equivalent to an arrow in the opposite direction and the entries $P'(T_F^*|T_P^*) = 1$ and $P'(T_F^*|\neg T_P^*) = 0$, for non-extreme priors. This is not entirely true, for this version produces a different Bayesian network in which T_F and T_P are d-separated (independent, that is) if no evidence for T_F^* is involved. If T_F and T_P are independent there is no change in coherence to the situation before the reduction (nor is there any flow of confirmation). Hence, we are in need of another argument for drawing the arrow from T_F^* to T_P^* since the argument „it doesn't matter, so we can equally draw it to the right" does not apply. The argument might be found in the general fact that T_F explains T_P.[6]

Third, how to model the notion of strong analogy between T_P^* and T_P? Analogy seems to be a symmetrical relation. However, as explained in § 3, T_P^* in fact *corrects* T_P and is *logically stronger* than T_P. Bahmani et al. thus argue that „'analogy' is perhaps not the right word as T_P^* is indeed stronger than

---

[6] There might be an argument already to be found in (Nagel 1970) to the effect that this direction in fact produces the correct Bayesian network. According to Nagel, bridge laws in reductions (call them *reductive bridge laws*) state either necessary and sufficient or *only sufficient* conditions (cf. Nagel 1970, 367). More precisely, reductive bridge laws have the following form. In § 3, it was mentioned that Nagel considered two types of bridge laws, namely relations between properties and identifications of (classes of) entities. Let F be a predicate of T_F and P a predicate of T_P but not of T_F; let a be a name in T_F and b a name in T_P but not T_F; let x range over the domain of objects T_P talks about. According to Nagel (ibid.), a sentence is a reductive bridge law just in case (a) or (b) holds:
(a) „For all x: if F applies to x then P applies to x" / „For all x: if a denotes x then so does b", or
(b) „For all x: F applies to x iff P applies to x" / „For all x: a denotes x iff b does". (There are other forms of bridge laws – necessary conditions – but they are, according to Nagel (ibid.), not reductive: there may be loss of logical strength from T_P to T_F. This might pose a problem for the Nagelian model of reduction; think of examples like light waves (in T_P = physical optics) as a subset of electromagnetic waves (in T_F = electrodynamics), which Nagel explicitly mentions. In these cases, boundary conditions must be specified ("Electromagnetic waves in the spectrum [X, Y] are light waves") for a successful reduction to take place.) Now think of our simplifying assumption that bridge laws state a perfect correlation – as in (b) – as a special case. In normal cases (a), we naturally draw the arrow from the left to the right since a bridge law states a sufficient condition in terms of T_F under which the property P refers to occurs. We can then specify the probability table $P'(T_P^*|\neg T_F^*) = 1$ and $P'(T_P^*|T_F^*) = y$, where $y \in [0,1]$. However, the argument needs to be backed up by case studies showing that (a) is indeed the default case. This might suggest a procedural reading of equations in bridge laws – however, this exceeds the extension of this investigation.

T_p, and so it makes sense to draw an arrow from T_p^* to T_p" (2011, 330). I follow this suggestion. Nothing quantitative is said about the notion of strong analogy, so we put two undefined constants $P'(T_p|T_p^*) = p$ (I shall assume $p \neq 0$) and $P'(T_p|\neg T_p^*) = q$.
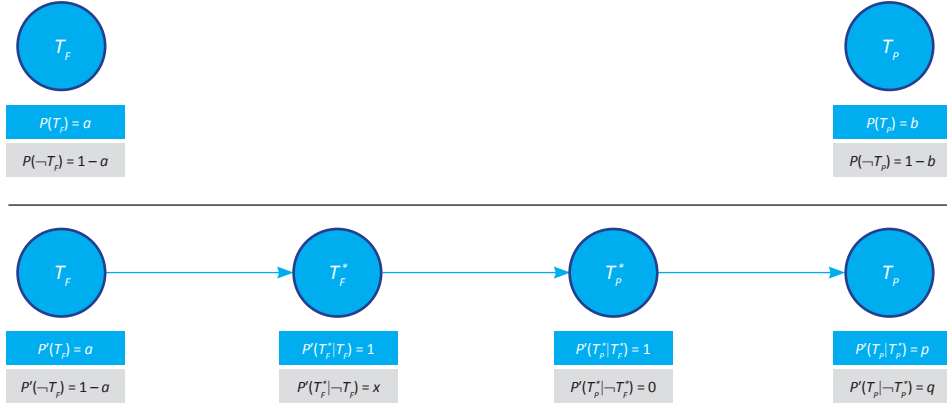

Figure 9: Bayesian networks pre & post reduction.

Fig. 9 shows the two Bayesian networks representing the situation before and after the reduction. The simplifying assumptions are taken into account. For example, the prior probability of T_F is the same before and after the reduction. But note that instead of the prior probability of T_P the probability table states its conditional probability after the reduction.

IV. COHERENCE AND REDUCTION.

We are now able to undertake the actual analysis, i.e. to see what conditions need to be fulfilled in a reduction in order for the coherence of the two theories involved to increase. I apply the measures given in § 2 to the theories T_F and T_P before and after reduction as presented in the Bayesian networks in Fig. 9. So instead of applying the measures to two different information sets (as in § 2), we apply them to one information set $S = \{T_F, T_P\}$ under two different probability functions P and P' Let $m_x(...)$ denote the respective measure under the probability function P and $m'_x(...)$ the respective measure under the probability function P'.

IV.1. SHOGENJI.

T_F and T_P are probabilistically independent before the reduction, thus

$$m_{Sh}(S) = \frac{P(R_1, \ldots, R_m)}{\prod_{i=1}^{m} P(R_i)} = \frac{P(T_F) \times P(T_P)}{P(T_F) \times P(T_P)} = 1.$$

RR*f*A

Note that the Shogenji measure has the feature that for any number of independent theories, their coherence always gets assigned the value 1.

After the reduction, T_F and T_P are probabilistically dependent. Hence,

$$m'_{Sh}(S) = \frac{P'(T_F \mid T_P) \times P'(T_P)}{P'(T_F) \times P'(T_P)} = \frac{P'(T_F \mid T_P)}{P'(T_F)}.$$

So $m'_{Sh}(S) > m_{Sh}(S)$ iff $P'(T_F|T_P) > P'(T_F)$. We specified in the Bayesian network in Fig. 9 that $P'(T_F) = a$; in appendix [B], it is calculated that $P'(T_F \mid T_P) = \dfrac{ap}{p(a+x-ax)+q(1-a-x+ax)}$. Inserting these values, we get

$$\frac{ap}{p(a+x-ax)+q(1-a-x+ax)} > a$$

as a necessary and sufficient condition for an increase in coherence of the two theories in a reduction, when the underlying measure is the Shogenji measure (the fraction is always well-defined; see appendix [B]). This holds iff

$$ap > a(p(a+x-ax)+q(1-a-x+ax)),$$

so iff

$$p > p(a+x-ax)+q(1-a-x+ax)$$

where the expression on the right is just $P'(T_P)$ (this is calculated in appendix [A]). So in the Shogenji measure, S receives a coherence boost in a reduction iff the post reductive probability of T_P is lower than that of T_P given its strongly analogous theory T_P^*. When is this the case?

Proposition (1). S receives an increase in coherence in a reduction under the assumptions specified in the Bayesian network (Fig. 9) and if the underlying coherence measure is the <u>Shogenji measure</u> iff

$$P'(T_P|T_P^*) > P'(T_P|\neg T_P^*).$$

Proof. We have shown that under the Shogenji measure, the coherence of $S = \{T_F, T_P\}$ increases in a reduction iff

$$p(a+x-ax)+q(1-a-x+ax) < p.$$

Since $(a+x-ax)$ and $(1-a-x+ax)$ sum to 1, we can express these terms with the help of a constant, say $\alpha = a+x-ax$, and thus $1-\alpha = 1-a-x+ax$. Note that $\alpha \in (0,1)$ since we assumed that $a \in (0,1)$.

Inserting the constant, the condition on an increase in coherence is

$$p \times \alpha + q \times (1-\alpha) < p;$$

So

$$p\alpha + q - q\alpha < p$$

$$q - q\alpha < p - p\alpha.$$

RR*f*A

This condition is satisfied iff $q < p$, which is just $P'(T_P|T_P^*) > P'(T_P|\neg T_P^*)$. □

This is a plausible result: it seems an appropriate condition on the notion of strong analogy that $P'(T_P|T_P^*) > P'(T_P|\neg T_P^*)$, i.e. that the probability of a theory conditional on a strongly analogous theory be higher than the probability of the theory conditional on the negation of the strongly analogous theory.

IV.2. OLSSON.

Remember that the Olsson measure is based on $m_O(S) = \dfrac{P(R_1,\ldots,R_m)}{P(R_1 \vee \ldots \vee R_m)}$.
Because of the independence of the two theories before the reduction, the measure yields

$$m_O(S) = \frac{P(T_F,T_P)}{P(T_F \vee T_P)}$$
$$= \frac{ab}{a+b-ab}.$$

We don't need to worry about the expressions being undefined because of our assumption that $a, b \neq 0$. This holds for all the fractions in this section.

Note that the Olsson measure is sensitive to prior probabilities in the case of independent information items: the higher the prior probabilities, the higher the value of coherence the measure yields. This does not coincide with our intuitive notion of coherence.

After the reduction, we need the value of

$$m_O'(S) = \frac{P'(T_F,T_P)}{P'(T_F \vee T_P)} = \frac{P'(T_F|T_P) \times P'(T_P)}{P'(T_F) + P'(T_P) - P'(T_F|T_P) \times P'(T_P)}.$$

Calculating (see appendix [A] and [B]) the values of $P'(T_P)$ and $P'(T_F|T_P)$ and inserting all the values, we get

$$= \frac{ap}{a + p(a+x-ax) + q(1-a-x+ax) - ap}$$

So the coherence of the two theories as measured with the Olsson measure increases in a reduction iff

$$\frac{ab}{a+b-ab} < \frac{ap}{a + p(a+x-ax) + q(1-a-x+ax) - ap}.$$

When is this the case? One way to compare the expressions is to fix the probability of T_P:

Proposition (2). If $P(T_P) = P'(T_P)$ then it is a necessary and sufficient condition for an increase in coherence under the <u>Olsson measure</u> that $P'(T_P|T_P^*) > P'(T_P|\neg T_P^*)$.

RRfA

Proof. The coherence increases iff

$$\frac{ab}{a+b-ab} < \frac{ap}{a+p(a+x-ax)+q(1-a-x+ax)-ap}. \quad (1)$$

Assume that $P(T_p) = b = p(a + x - ax) + q(1 - a - x + ax) = P'(T_p)$.

<u>Sufficient condition.</u> Suppose $P'(T_p|T_p^*) = p > q = P'(T_p|\neg T_p^*)$. It was proved in § 4.1 that this is the case iff $p(a + x - ax) + q(1 - a - x + ax) < p$. Since we assumed that $b = p(a + x - ax) + q(1 - a - x + ax) < p$, it is also the case that $b < p$. So $ab < ap$ for $a \in (0,1)$. So the enumerator on the right side of (1) is bigger than the enumerator on the left side.

Thus, if the denominator on the left side is bigger or equal than on the right side then the inequality in (1) holds. Is it the case that $a + b - ab \geq a + p(a + x - ax) + q(1 - a - x + ax) - ap$? a cancels out, so $b - ab > p(a + x - ax) + q(1 - a - x + ax) - ap$. But since we assumed $P(T_p) = b = p(a + x - ax) + q(1 - a - x + ax) = P'(T_p)$, this is the case iff $-ab \geq -ap$. But since $p > b$, this is the case.

<u>Necessary condition.</u> Suppose it is not the case that $P'(T_p|T_p^*) = p > q = P'(T_p|\neg T_p^*)$, so $q \geq p$. By the same line of argument as above,

$$\frac{ab}{a+b-ab} \geq \frac{ap}{a+p(a+x-ax)+q(1-a-x+ax)-ap}. \qquad \square$$

The condition $q < p$ is plausible and in agreement with what we proved in § 4.1 is the result of applying Shogenji's measure. However, the presupposition that $P(T_p) = P'(T_p)$ is an (over)simplification. In general, we would expect that $P(T_p) < P'(T_p)$. Why would we expect this? Dizadji-Bahmani et al. (2011) argue (although they indeed fix the probability of T_p) that evidence which pre reduction confirmed T_F but not T_p may post reduction as well confirm T_p; thus, its probability may be higher. Another factor is coherence: if a theory gets reduced to a well accepted fundamental theory and thus (as argued here) coheres well with it, its probability might be higher.

So let's drop the simplifying assumption. It is then possible to get a sufficient condition for an increase in coherence.

Proposition (2.1). It is a sufficient condition for an increase in coherence under the <u>Olsson measure</u> that $P'(T_p|T_p^*) > P(T_p)$ and $P'(T_p|T_p^*) > P'(T_p|\neg T_p^*)$.

Proof. We further reduce (1):

$$b+a-1 < p + \frac{ap}{p(a+x-ax)+q(1-a-x+ax)} - 1$$

$$a+b < \frac{ap}{p(a+x-ax)+q(1-a-x+ax)} + p.$$

Now assume that $P'(T_p|T_p^*) = p > b = P(T_p)$ and $P'(T_p|T_p^*) = p > q = P'(T_p|\neg T_p^*)$. By the first assumptions, it is a sufficient condition for a coherence boost that

$$a < \frac{ap}{p(a + x - ax) + q(1 - a - x + ax)},$$

so

$$p(a + x - ax) + q(1 - a - x + ax) < p,$$

which means that

$$P'(T_P) < P'(T_P|T_P^*).$$

As shown in § 4.1, this is equivalent to $P'(T_P|T_P^*) > P'(T_P|\neg T_P^*)$, which was our second assumption. □

Note again that $P'(T_P|T_P^*) > P'(T_P|\neg T_P^*)$ iff $P'(T_P|T_P^*) > P'(T_P)$; so we can also state the result thus: if the probability of the theory to be reduced conditional on its strongly analogous theory is both bigger than its prior probability before and after the reduction then the coherence of the theories increases.

To sum up, we have two plausible results if we apply the Olsson measure: under the assumption that the probability of T_P be the same before and after the reduction, it is a necessary and sufficient condition for a coherence boost that $P'(T_P|T_P^*) > P(T_P|\neg T_P^*)$; and it is a sufficient condition for a coherence boost that both $P'(T_P|T_P^*) > P'(T_P|\neg T_P^*)$ and $P'(T_P|T_P^*) > P(T_P)$.

IV.3. BOVENS & HARTMANN.

Remember that u_i denotes the sum of all the joint probabilities of the instances of i negative values and $n - i$ positive values of the variables $R_1,...,$ R_n (cf. § 2 or Bovens & Hartmann 2003, 17). So for our information set $S = \{T_F, T_P\}$, we have $u_0 = (T_F, T_P)$, $u_1 = P(T_F, \neg T_P) + P(\neg T_F, T_P)$ and $u_2 = P(\neg T_F, \neg T_P)$. I use u_i to denote the respective value pre reduction and u_i' to denote the respective value post reduction.

Let's repeat what the Bovens & Hartmann measure states for our special case in which the information sets in question have only two items: for such information sets S and S' it is a necessary and sufficient condition for $S'CS$ that

(i) $u_0 \leq u_0'$ and $u_1 \leq u_1'$ or

(ii) $u_0 \geq u_0'$ and $\dfrac{u_1}{u_1'} \geq \dfrac{u_0}{u_0'}$ (cf. Bovens & Hartmann 2003, 37).

As said before, for our reductive context, we don't apply the measure to two different information sets but to the same information set S under two different probability functions, namely before and after the reduction takes place. Condition (i) then requires for S to be not less coherent after the reduction that the joint probabilities of the theories be no lower after reduction and that the sum of the joint probabilities of one theory being true while the other one being false be no lower before the reduction. In other words, condition (i) is satisfied iff

$$P(T_F|T_P) \leq P'(T_F|T_P)$$

RRfA

and

$$P(T_F, \neg T_P) + P(\neg T_F, T_P) \geq P'(T_F, \neg T_P) + P'(\neg T_F, T_P),$$

The values for the first part are already known; the values for the second are calculated in appendix [C]. Inserting values, a sufficient condition for a coherence increase or equality in a reduction under the Bovens & Hartmann measure is:

(i.a) $ab \leq ap$, so $b \leq p$ and

(i.b) $a + b - 2ab \geq a + p(a + x - ax) + q(1 - a - x + ax) - 2ap$, so $b - 2ab \geq p(a + x - ax) + q(1 - a - x + ax) - 2ap$.

Condition (ii) is satisfied iff

$$P(T_F, T_P) \geq P'(T_F, T_P)$$

and

$$\frac{P(T_F, \neg T_P) + P(\neg T_F, T_P)}{P'(T_F, \neg T_P) + P'(\neg T_F, T_P)} \geq \frac{P(T_F, T_P)}{P'(T_F, T_P)}.$$

Inserting values (calculations are in appendix [C]), condition (ii) requires that

(i.a) $b \geq p$ and

(ii.b) $\dfrac{a(1-b) + (1-a)b}{(p(a+x-ax) + q(1-a-x+ax) - ap) + (a - ap)} \geq \dfrac{ab}{ap}$.

When are conditions (i) or (ii) the case? Again, it is easy to compare the right with the left sides if we fix the probability of T_P.

Proposition (3). If the probabiliy of T_P is the same pre and post reduction it is a necessary and sufficient condition for an increase in coherence under the <u>B&H measure</u> that $P'(T_P | T_P^*) > P'(T_P | \neg T_P^*)$.

Proof. Suppose $P(T_P) = b = p(a + x - ax) + q(1 - a - x + ax) = P'(T_P)$.

First consider condition (i). Since $b = p(a + x - ax) + q(1 - a - x + ax)$, (i.b) holds iff $-2ab \geq -2ap$, so iff $b \leq p$. But just in this case (i.a) is fulfilled as well. So $b \leq p$ is a necessary and sufficient condition for (i) to hold.

What we have proved is that $b \leq p$ is a necessary condition for S being not less coherent post reduction. Is it more coherent post reduction? Suppose it is not, so it is also the case that S is not less coherent pre reduction. Then, by the same line of argument as above, $b \geq p$. So $b = p$. Thus, if $b = p$ then S is in fact equally coherent before and after reduction. An increase in coherence takes place only if $b < p$.[7]

---

[7] Note that in the other measures, I have presupposed that S is strictly more coherent after reduction iff $m_X'(S) > m_X(S)$. This is trivial: if $m_X'(S) > m_X(S)$ then $S_{post} CS_{pre}$ but not $S_{pre} CS_{post}$ is the case; but if $m_X'(S) = m_X(S)$ then both $S_{post} CS_{pre}$ and $S_{pre} CS_{post}$ are the case, so S_{pre} and S_{post} are equally coherent.

Now consider condition (ii). Since $p(a + x − ax) + q(1− a − x + ax) = b$, condition (ii.b) becomes

$$\frac{a+b-2ab}{a+b-2ap} \geq \frac{b}{p}$$

$$(a + b − 2ab)p \geq (a + b − 2ap)b$$

$$p(a + b) − 2abp \geq b(a + b) − 2abp$$

$$p \geq b.$$

But (ii.a) states that $b \geq p$; so $p = b$ is a necessary and sufficient condition for S to be not less coherent after the reduction, by (ii). However, it is easily seen that if $p = b$, by (ii) S is also not less coherent before reduction; i.e., it is equally coherent before and after the reduction in this case. Hence condition (ii) cannot produce a coherence boost in a reduction, and thus can be dismissed.

We conclude that it is a necessary and sufficient condition for an increase in coherence under the B&H measure that $b = p(a + x − ax) + q(1− a − x + ax) < p$, which is a condition already known from § (4.1). As shown there, this is the case iff $P'(T_P|T_P^*) = p > q = P'(T_P|\neg T_P^*)$.

IV.4. FITELSON.

The range of Fitelson's coherence function is $[−1, 1]$. As in the Shogenji measure, sets of independent propositions always get assigned the same value; in the Fitelson measure, this is the value 0. The extreme values $−1$ and 1 are assigned only to information sets with propositions each or which is unsatisfiable (for $−1$) or satisfiable and each logically equivalent (for 1) (cf. Fitelson 2003). We need the values of $F(T_F, T_P)$ and $F(T_P, T_F)$, both before and after the reduction. As with the probability function, let's use $F'(...)$ to distinguish the Kemeny-Oppenheim function after the reduction.

Before reduction, applying Fitelson's measure to the conjunction of the theories yields 0 because of our assumption that T_F and T_P be probabilistically independent. So all that Fitelson's measure requires for the coherence to increase in the reduction is

$$\frac{F'(T_F, T_P) + F'(T_P, T_F)}{2} > 0,$$

so

$$F'(T_F, T_P) + F'(T_P, T_F) > 0$$

The question is, what are the conditions for this to be the case? In appendix [D] it is calculated that

$$F'(T_F, T_P) = \frac{P'(T_P \mid T_P^*) − P'(T_P)}{P'(\neg T_F) \times P'(T_P)};$$

RRƒA

and

$$F'(T_P, T_F) = \frac{P'(T_P \mid T_P^*) - P'(T_P)}{P'(T_P \mid T_P^*) + P'(T_P) - 2 \times P'(T_P \mid T_P^*) \times P'(T_P)}.$$

Both denominators are strictly positive (see appendix [D]); and since the enumerators are equal, it is a necessary and sufficient condition for an increase in coherence that $P'(T_P \mid T_P^*) - P'(T_P) > 0$. We know this result from § 4.1, so:

Proposition (4). A necessary and sufficient condition for an increase in coherence if the underlying measure is <u>Fitelson's</u> is that $p > q$.

V. FURTHER RESEARCH; MERITS AND PROBLEMS OF THE BAYESIAN FRAMEWORK.

We have seen that the coherence measures applied yield a stable result: in all the measures, it is a sufficient condition and in most of the measures also a necessary condition for the coherence of two theories to increase in a reduction that the theory to be reduced be positively related to its strongly analogous theory which, possibly via bridge laws, follows from the reducing theory. The simplifying assumptions underlying this result are that the theories be probabilistically independent before the reduction, that the probability of the fundamental (and partly also of the phenomenological) theory be the same pre and post reduction, and that the [bridge laws] state a perfect correlation. It is to be expected that if these assumptions are given up, besides the positive correlation of T_P^* and T_P there will be additional conditions on an increase in coherence. This is left for another investigation.

It was claimed at the beginning that we can not only look at what happens to two theories in terms of their coherence when one gets reduced to the other one, but also evaluate cases of purported intertheoretic reductions in terms of the coherence of the theories involved: does their coherence increase in the reduction? Second, from this double-edged relation (which is a typical feature of Bayesian analyses) we can draw conclusions about the logic of reduction. It was mentioned in § 3 that the intuitive notion of [strong analogy] is problematic and in need of a refinement. We can now see the line in which this can be done. If coherence is a relevant criterion for analysing reductions, positive correlation seems to be a minimal requirement for the theory to be reduced and the mediating theory to be strongly analogous.

Some lines of further research: in the next step, our questions could be stated cardinally: how much should the coherence increase in order for the reduction to be admissible, and which conditions on a reduction and its critical notions result from this?

Another interesting follow-up question: what about the coherence of two phenomenological theories both of which get reduced to one and the same fundamental theory? Arguably, their coherence should increase as well. Naturally, in answering these further questions the results of applying diverse measures should again sufficiently agree with each other.

Even if they do, we should bear a few caveats in mind when asking what coherence demands for a reduction. First, coherence is only one of various factors (like simplicity, conservatism, confirmation, etc.) that drive reductionist programmes. Only a stubborn coherentist would hold that coherence should be the only one. Certainly, coherence can still be regarded as an important one: if two theories (partly) share a domain of applicability, it is an unfortunate state if they do not cohere, and we wouldn't call a phenomenological theory reduced to a fundamental one (or two phenomenological theories reduced to one and the same fundamental one) if the „reduction" did not give us a *proof* of an increase in coherence. Moreover, the various factors – coherence, simplicity, etc. – are rivals only if they demand features which are mutually inconsistent. Now our result nicely coincides with the one Dizadji-Bahmani et al. (2011, 331 et sqq.) arrive at in the context of confirmation: it is a necessary condition for a reduction to have advantageous confirmatory features that the strongly analogous theories be positively related. These confirmatory features are that the prior probability $P'(T_F|T_P)$, their probability given some evidence, and their degree of confirmation as measured by the difference measure are higher after than before a reduction. So much for coherence and confirmation; other factors should as well be checked.

Finally, we shouldn't remain silent about a problem that occurs when we are confronted with practical cases of reductions; namely the old Bayesian riddle of how to ascribe subjective probabilities to theories and which poses (or so I argue) some particular problems in the model of reduction that underlies this investigation. It is to be expected that – besides the usual problem of assigning prior probabilities – it is particularly problematic to answer the questions:

(a)Before reduction: how to decide if two given theories T_F and T_P are in fact probabilistically independent or not? After all, they share (at least partially) one domain of applicability; however, without having reduced T_P to T_F (or having established some other intertheoretic relation between them), what can be said about their (in)dependence?

(b) After reduction: how are we to fix conditional probabilities? Particularly problematic is the fixing of

    1. $P'(T_P^*|T_F^*)$ and $P'(T_P^*|\neg T_F^*)$: should the assumption that the bridge laws state a perfect correlation be relaxed? and

    2. $P(T_P|T_P^*)$ and $P'(T_P|\neg T_P^*)$: how to determine if T_P^* and T_P are indeed strongly analogous?

It should be noted that mediating theories like T_F^* and T_P^* are not normally stated by scientists; this seems to make the fixing of their probabilities or of other theories' probabilities given mediating theories artificial.

RRfA

Appendix

The procedures for calculating probabilities in Bayesian nets are, for example, explained in (Jensen 2000). They are standard when working with Bayesian networks.

[A] Post-reduction prior probability $P'(T_p)$.

First, we use the second Bayesian network in Fig. 9 to get $P'(T_p)$ – the prior probability of T_p after the reduction. This could be done by just applying the chain rule, but it is instructive for seeing how the value depends on the assumptions specified in the Bayesian net to once go for an iterated marginalisation. We start with marginalising $P'(T_F)$ out of $P'(T_F^*|T_F)$:

$$P'(T_F^*) = P'(T_F^*, T_F) + P'(T_F^*, \neg T_F)$$

We read the conditional probability table off the Bayesian network

$P'(T_F^*\|T_F)$	T_F	$\neg T_F$
T_F^*	1	x
$\neg T_F^*$	1 − 1 (not needed)	1 − x (not needed)

With the help of the conditional probability table we get the joint probability table using the fundamental rule on $P'(T_F) = a$:

$P'(T_F^*, T_F)$	T_F	$\neg T_F$
T_F^*	a	$(1 - a)x$
$\neg T_F^*$	(not needed)	(not needed)

So $P'(T_F^*) = a + (1 - a)x = a + x - ax$.

Next, we need the value $P'(T_P^*)$:

$$P'(T_P^*) = P'(T_P^*, T_F^*) + P'(T_P^*, \neg T_F^*)$$

Conditional probability table:

$P'(T_P^*\|T_F^*)$	T_F^*	$\neg T_F^*$
T_P^*	1	0
$\neg T_P^*$	(not needed)	(not needed)

RR*f*A

Applying the same procedure as before, we get the joint probability table:

$P'(T_P^*, T_F^*)$	T_F^*	$\neg T_F^*$
T_P^*	$a + x - ax$	0
$\neg T_P^*$	(not needed)	(not needed)

So $P'(T_P^*) = a + x - ax$ (because of the perfect correlation between T_F^* and T_P^*, this step could have been omitted but is, of course, vital if the assumption of a perfect correlation is dropped).

Finally, we calculate

$$P'(T_P) = P'(T_P, T_P^*) + P'(T_P, \neg T_P^*)$$

Conditional probability table:

$P'(T_P | T_P^*)$	T_P^*	$\neg T_P^*$
T_P	p	q
$\neg T_P$	(not needed)	(not needed)

Joint probability table:

$P'(T_P, T_P^*)$	T_P^*	$\neg T_P^*$
T_P	$p(a + x - ax)$	$q(1 - (a + x - ax))$
$\neg T_P$	(not needed)	(not needed)

Thus, we have the first value we are interested in – the prior probability of T_P after the reduction:

$$P'(T_P) = p(a + x - ax) + q(1 - a - x + ax)$$

[B] Post-reduction conditional probability $P'(T_F | T_P)$.

Next, we need $P'(T_F | T_P)$. By definition of conditional probability,

$$P'(T_F | T_P) = \frac{P'(T_F, T_P)}{P'(T_P)}.$$

RRfA

Applying the chain rule, we get

$$P'(T_F, T_P) = \sum_{\mathrm{T}_F^*, \mathrm{T}_P^*} P'(T_F, \mathrm{T}_F^*, \mathrm{T}_P^*, T_P)$$

$$= P'(T_F) \times P'(T_F^* \mid T_F) \times P'(T_P^* \mid T_F^*) \times P'(T_P \mid T_P^*)$$

$$= a \times 1 \times 1 \times p = ap.$$

So

$$P'(T_F \mid T_P) = \frac{ap}{p(a + x - ax) + q(1 - a - x + ax)}.$$

For this to be a well-defined expression we require that $p(a + x - ax) + q(1 - a - x + ax) \neq 0$, which always holds if $a > 0$ and $p > 0$ because the first summand is > 0 and the second summand cannot be negative.

[C] Sums of joint probabilities.

For the <u>Bovens&Hartmann measure</u>, we need the sum of the joint probabilities of one theory and the negation of the other, respectively, i.e. the values of $P(T_F | \neg T_P)$ and $P(\neg T_F | T_P)$ and of $P'(T_F | \neg T_P)$ and $P'(\neg T_F | T_P)$.

Before reduction the theories are independent; thus, the sum of these joint probabilities is just

$$u_1 = a(1 - b) + (1 - a)b = a + b - 2ab.$$

After reduction the values are calculated as follows. From appendix [B], we already have the values to calculate (instead of using the chain rule) $P'(\neg T_F, T_P) = P'(\neg T_F | T_P) \times P'(T_P)$, namely

$$= (1 - \frac{ap}{(p(a + x - ax) + q(1 - a - x + ax)})) \times (p(a + x - ax) + q(1 - a - x + ax)).$$

So

$$P'(\neg T_F, T_P) = p(a + x - ax) + q(1 - a - x + ax) - ap.$$

By the chain rule,

$$P'(T_F, \neg T_P) = \sum_{\mathrm{T}_F^*, \mathrm{T}_P^*} P'(T_F, \mathrm{T}_F^*, \mathrm{T}_P^*, \neg T_P)$$

$$= P'(T_F) \times P'(T_F^* \mid T_F) \times P'(T_P^* \mid T_F^*) \times P'(\neg T_P \mid T_P^*)$$

$$= a \times 1 \times 1 \times (1 - p) = a - ap.$$

RR*f*A

The post-reduction sum of the joint probabilities of one negative and one positive value of T_F and T_P, respectively, is thus

$$u_1' = (p(a+x-ax)+q(1-a-x+ax)-ap)+(a-ap)$$

$$= p(a+x-ax)+q(1-a-x+ax)+a-2ap.$$

[D] Values for the Kemeny-Oppenheim measure.

Finally, we need some missing values for the <u>Fitelson measure</u>. It has already been said (cf. § 2.1) that the Kemeny-Oppenheim function is defined thus:

$$F(R_i,R_j) = \frac{P(R_j \mid R_i) - P(R_j \mid \neg R_i)}{P(R_j \mid R_i) + P(R_j \mid \neg R_i)}$$

for $P(R_j) < 1$ and $P(R_i) > 0$. We are interested in $F'(T_F, T_P)$ and $F'(T_P, T_F) - F'$ being the Kemeny-Oppenheim function applied to the theories after reduction. So,

$$F'(T_F,T_P) = \frac{P'(T_P \mid T_F) - P'(T_P \mid \neg T_F)}{P'(T_P \mid T_F) + P'(T_P \mid \neg T_F)}.$$

Inserting values which are already known (from appendix [B]),

$$P'(T_P \mid T_F) = \frac{P'(T_P,T_F)}{P'(T_F)} = \frac{ap}{a} = p$$

and (from appendix [C])

$$P'(T_P \mid \neg T_F) = \frac{P'(T_P,\neg T_F)}{P'(\neg T_F)} = \frac{px - apx + q - aq - qx + aqx}{1-a}.$$

For both expressions to be well-defined, we require that $0 < a < 1$. Inserting the values in the Kemeny-Oppenheim measure, we get

$$F'(T_F,T_P) = \frac{p - \dfrac{px - apx + q - aq - qx + aqx}{1-a}}{p + \dfrac{px - apx + q - aq - qx + aqx}{1-a}}$$

$$= \frac{\dfrac{p - ap - px + apx - q + aq + qx - aqx}{1-a}}{\dfrac{p - ap + px - apx + q - aq - qx + aqx}{1-a}}$$

$$= \frac{p - ap - px + apx - q + aq + qx - aqx}{p - ap + px - apx + q - aq - qx + aqx},$$

RRƒA

which is always well defined for $p > 0$ and $a < 1$ (both of which were already required above and are unproblematic). It looks rather ugly, but we meet old acquaintances in it:

$$= \frac{p - (p(a + x - ax) + q(1 - a - x + ax))}{p \times (1 - a) + px - apx + q - aq - qx + aqx}$$

$$= \frac{P'(T_P \mid T_P^*) - P'(T_P)}{P'(T_P \mid T_F) \times P'(\neg T_F) + P'(\neg T_F, T_P)}$$

$$= \frac{P'(T_P \mid T_P^*) - P'(T_P)}{P'(T_P \mid T_F) \times P'(\neg T_F) + P'(T_P \mid \neg T_F) \times P'(\neg T_F)}$$

$$= \frac{P'(T_P \mid T_P^*) - P'(T_P)}{P'(\neg T_F) \times (P'(T_P \mid T_F) + P'(T_P \mid \neg T_F))}$$

$$= \frac{P'(T_P \mid T_P^*) - P'(T_P)}{P'(\neg T_F) \times P'(T_P)}.$$

For $F'(T_P, T_F)$:

$$F'(T_P, T_F) = \frac{P'(T_F \mid T_P) - P'(T_F \mid \neg T_P)}{P'(T_F \mid T_P) + P'(T_F \mid \neg T_P)}$$

In appendix [B], it was calculated that

$$P'(T_F \mid T_P) = \frac{ap}{p(a + x - ax) + q(1 - a - x + ax)} \, ;$$

in order to get $P'(T_F, \neg T_P)$, we insert values we already have (appendix [C] and [A]) in

$$P'(T_F \mid \neg T_P) = \frac{P'(T_F, \neg T_P)}{P'(\neg T_P)} = \frac{a - ap}{1 - (p(a + x - ax) + q(1 - a - x + ax))}$$

in which the denominator is $\neq 0$ iff $p < 1$ or $q < 1$ (let's suppose the latter). $F'(T_P, T_F)$ then becomes a compound fraction:

$$F'(T_P, T_F) = \frac{\dfrac{ap}{p(a + x - ax) + q(1 - a - x + ax)} - \dfrac{a - ap}{1 - p(a + x - ax) - q(1 - a - x + ax)}}{\dfrac{ap}{p(a + x - ax) + q(1 - a - x + ax)} + \dfrac{a - ap}{1 - p(a + x - ax) - q(1 - a - x + ax)}},$$

which we want to reduce. Note that we don't need to worry about the denominator being $= 0$ since the first summand is positive by the assumptions that $a, p > 0$ and the second summand cannot be negative.

RR𝑓A

The fraction in the enumerator becomes

$$\frac{ap(1-p(a+x-ax)-q(1-a-x+ax))-(a-ap)(p(a+x-ax)+q(1-a-x+ax))}{(p(a+x-ax)+q(1-a-x+ax))\times(1-p(a+x-ax)-q(1-a-x+ax))}$$

$$= ...$$

$$=\frac{ap-ap(a+x-ax)-aq(1-a-x+ax)}{(p(a+x-ax)+q(1-a-x+ax))\times(1-p(a+x-ax)-q(1-a-x+ax))},$$

and the fraction in the denominator

$$\frac{ap(1-p(a+x-ax)-q(1-a-x+ax))+(a-ap)(p(a+x-ax)+q(1-a-x+ax))}{(p(a+x-ax)+q(1-a-x+ax))\times(1-p(a+x-ax)-q(1-a-x+ax))}$$

$$= ...$$

$$=\frac{ap-2ap^2(a+x-ax)-2apq(1-a-x+ax)+ap(a+x-ax)+aq(1-a-x+ax)}{(p(a+x-ax)+q(1-a-x+ax))\times(1-p(a+x-ax)-q(1-a-x+ax))}.$$

Now since both fractions have the same denominator, we reduce the compound fraction to

$$F'(T_P,T_F)=\frac{ap-ap(a+x-ax)-aq(1-a-x+ax)}{ap-2ap^2(a+x-ax)-2apq(1-a-x+ax)+ap(a+x-ax)+aq(1-a-x+ax)}.$$

We then try to find compound values in this expression:

$$=\frac{p-p(a+x-ax)-q(1-a-x+ax)}{p+p(a+x-ax)+q(1-a-x+ax)-2p^2(a+x-ax)-2pq(1-a-x+ax)}$$

$$=\frac{p-(p(a+x-ax)+q(1-a-x+ax))}{p+p(a+x-ax)+q(1-a-x+ax)-2p(p(a+x-ax)+q(1-a-x+ax))}$$

$$=\frac{P'(T_P\mid T_P^*)-P'(T_P)}{P'(T_P\mid T_P^*)+P'(T_P)-2\times P'(T_P\mid T_P^*)\times P'(T_P)}.$$

References

Bovens, Luc and Hartmann, Stephan. 2003. *Bayesian epistemology*. Oxford: Clarendon Press.

Bovens, Luc and Hartmann, Stephan. *NA*. „Coherence, belief expansion and Bayesian networks". *Unknown journal*.

Bovens, Luc and Hartmann, Stephan. 2005. „Coherence and the Role of Specificity: A Response to Meijs and Douven". In *Mind*, Vol. 114.

RRƒA

Dizadji-Bahmani, Foad, Frigg, Roman and Hartmann, Stephan. 2010. „Who is afraid of Nagelian reduction?" In *Erkenntnis* 73, 393 - 412.

Dizadji-Bahmani, Foad, Frigg, Roman and Hartmann, Stephan. 2011. „Confirmation and reduction: a Bayesian account". In *Synthese* 179, 321 - 338.

Douven, Igor and Meijs, Wouter. 2005. „Bovens and Hartmann on Coherence". In *Mind*, New Series, Vol. 114, No. 454, 355 - 363.

Fitelson, Branden. 1999. „The Plurality of Bayesian Measures of Confirmation and the Problem of Measure Sensitivity". In *Philosophy of Science* 66, Supplement Proceedings of the 1998 Biennial Meetings of the Philosophy of Science Association. Part I: Contributed Papers (Sep., 1999).

Fitelson, Branden. 2003. „A probabilistic theory of coherence". In *Analysis* 63, 194 - 199.

Hempel, Carl Gustav and Oppenheim, Paul. 1948. „Studies in the Logic of Explanation". In *Philosophy of Science* 15, No. 2 (Apr., 1948), 135 - 175.

Jensen, Finn V. 2000. *An introduction to Bayesian networks*. London: UCL Press.

Kuhn, Thomas. 1977. „Objectivity, Value Judgment, and Theory Choice". In: *The Essential Tension: Selected Studies in Scientific Tradition and Change*, Chicago: Chicago University Press, 320 - 329.

Nagel, Ernest. 1970. „Issues in the Logic of Reductive Explanations". In *Mind, Science, and History*, H.E. Kiefer & K.M. Munitz (eds.), Albany, NY: SUNY Press, 117 – 137.

Neapolitan, Richard. 2003. *Learning Bayesian networks*. Prentice Hall Series in Artificial Intelligence.

Olsson, Erik J. 2002. „What is the problem of coherence and truth?" In *The Journal of Philosophy* 99, 246 - 272.

Schaffner, Kenneth F. 1974. „Reductionism in Biology: Prospects and Problems". In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association 1974*, 613 - 632.

Shogenji, Tomoji. 1999. „Is coherence truth-conducive?" In *Analysis* 59, 338 - 345.

Williamson, Jon. 2005. *Bayesian Nets and Causality: Philosophical and Computational Foundations*. Oxford: Oxford University Press.

RRfA