Assessing Theories

first draft!!!

Franz Huber
Centre for Junior Research Fellows
University of Konstanz, Germany
franz.huber@uni-konstanz.de

prepared for:
Formal Epistemology Workshop
organized by Branden Fitelson and Sahotra Sarkar
May 21-23, 2004, Berkeley, USA

invited paper for:
New Waves in Epistemology
ed. by Vincent F. Hendricks/Duncan Pritchard
Ashgate, 2005

Contents

Abstract

The problem adressed in this paper is “the main epistemic problem concerning science”, viz. “the explication of how we compare and evaluate theories [...] in the light of the available evidence” (van Fraassen 1983, 27).

The first part presents the general, i.e. paradigm independent, loveliness-likeliness theory of theory assessment. In a nutshell, the message is (1) that there are two values a theory should exhibit: informativeness (loveliness) and truth (likeliness) – measured respectively by a strength indicator (loveliness measure) and a truth indicator (likeliness measure); (2) that these two values are conflicting in the sense that the former is an increasing and the latter a decreasing function of the logical strength of the theory to be assessed; and (3) that in assessing a given theory one should weigh between these two conflicting aspects in such a way that any surplus in love(like)liness succeeds, if only the difference in like(love)liness is small enough.

Particular accounts of this general theory arise by inserting particular strength indicators and truth indicators. The theory is spelt out for the Bayesian paradigm; it is then compared with standard (incremental) Bayesian confirmation theory. Part 1 closes by asking whether it is likely to be lovely, and by discussing a few problems of confirmation theory in the light of the present approach.

The second part discusses the question of justification any theory of theory assessment has to face: Why should one stick to theories given high assessment values rather than to any other theories? The answer given by the Bayesian version of the account presented in the first part is that one should accept theories given high assessment values, because, in the medium run (after finitely many steps without necessarily halting), theory assessment almost surely takes one to the most informative among all true theories when presented separating data. The comparison between the present account and standard (incremental) Bayesian confirmation theory is continued.

The third part focuses on a sixty year old problem in the philosophy of science – that of a logic of confirmation. We present a new analysis of Carl G. Hempel’s conditions of adequacy (Hempel 1945), differing from the one Carnap gave in §87 of his (1962). Hempel, so it is argued, felt the need for two concepts of confirmation: one aiming at true theories, and another aiming at informative theories. However, so the analysis continues, he also realized that these two concepts are conflicting, and so he gave up the concept of confirmation aiming at informative theories. It is finally shown that one can have the cake and eat it: There is a logic of confirmation – or rather: theory assessment – that takes into account both of these two conflicting aspects.

1 The Problem

The problem adressed in this paper is this:

> the main epistemic problem concerning science ... is the explication of how we compare and evaluate theories ... in the light of the available evidence ... (van Fraassen 1983, 27)

In other and more mundane words, the question is: What is a good theory, and when is one theory better than another theory, given these data and those background assumptions.

Let us call this the problem of a theory of theory assessment. Its *quantitative* version can be put as follows:

- One is given a scientific hypothesis or theory T, a set of data – the evidence – E, and some background information B.
- The question is: How "good" is T given E and B? That is, what is the "*value*" of T in view of E and B?
- An answer to this quantitative question consists in the definition of a (set $\mathcal{A}$ of) function(s) a such that (for each $a \in \mathcal{A}$:) $a\,(T, E, B)$ measures the value of T in view of E and B, i.e. how good T is given E and B.

Given this formulation of our problem, a theory of theory assessment need not account for the way in which scientists arrive at their theories nor how they (are to) gather evidence nor what they may assume as background information. Furthermore, the purpose of this evaluation is, of course, that we *accept* those theories (among the ones we can or have to choose from) which score highest in the assessment relative to the available data. This makes clear that a proper treatment of the whole problem not only *explicates* how we evaluate theories in the light of the available evidence (sections 2-5); a proper treatment also *justifies* this evaluation by answering the question why we should accept those theories that score highest (sections 6-8).

The term 'theory assessment' is chosen in order to have a neutral term for the problem described above. In the literature, people usually speak about the confirmation or corroboration of, or the support for theory T by evidence E relative to background information B. If one understands the problem of confirmation in the broad sense described above, then our question may also be put as follows: How much does E confirm (support, corroborate) T relative to B? An answer to this question consequently consists of a (set $\mathcal{C}$ of) function(s) c such that (for each $c \in \mathcal{C}$:) $c\,(T, E, B)$ measures the degree to which E confirms T relative B.

However, as will become clear later on, confirmation is usually meant to cover just one of the aspects of a good theory, viz. its being (likely to be) true.

Having made clear what the following is about – and what it is not about – let us now address and solve both the explicatory or descriptive, and the justificatory or normative part of "the main epistemic problem concerning science".

2 Theory, Evidence, and Background Information

In order for the above characterisation to be precise one has to make clear what is meant by theory, evidence, and background information. In what follows it is assumed that for every scientific theory T, every piece of evidence E, and every body of background information B there exist finite axiomatizations (in a first-order language including identity and function symbols) A_T, A_E, and A_B, respectively, which formulate T, E, and B, respectively.

In general, not all finite sets of statements are formulations of a piece of evidence or a scientific theory. Scientific theories, for instance, do not speak about particular objects of their domain of investigation, but express general regularities or patterns. Data, on the other hand, only speak about finitely many objects of the relevant domain – we are damned *qua* humans to be able to observe only finitely many objects.

However, for the general, i.e. paradigm independent, theory outlined below (and its Bayesian version) it suffices that these be finitely axiomatizable. As theory assessment turns out to be closed under equivalence transformations, T, E, and B can and will be identified with one of their formulations A_T, A_E, and A_B, respectively.

3 Conflicting Concepts of Confirmation

Though some take theory assessment to be *the* epistemic problem in philosophy of science, there is no established branch adressing exactly this problem. What comes closest is what is usually called confirmation theory. So let us briefly look at confirmation theory, and see what insights we can get from there concerning our problem.

Confirmation has been a hot topic in the philosophy of science for more than sixty years now, starting with such classics as Carl Gustav Hempel's *Studies in the Logic of Confirmation* (1945)[1], Rudolph Carnap's work on Inductive Logic and Probability[2], and various other contributions by Nelson Goodman, Olaf Helmer, Janina Hosiasson-Lindenbaum, John G. Kemeny, R. Sherman Lehmann, Paul Oppenheim, Abner Shimony, and others.[3]

Roughly speaking, the last century has seen two main approaches to the problem of confirmation:

- On the one hand, there is the qualitative theory of *Hypothetico-Deductivism* HD (associated with Karl R. Popper), according to which a scientific theory T is confirmed by evidence E relative to background information B if and only if the conjunction of T and B logically implies E in some suitable way – the latter depending on the version of HD under consideration.

- On the other hand, there is the quantitative theory of probabilistic *Inductive Logic* IL (associated with Rudolf Carnap), according to which T is confirmed by E relative to B to degree r if and only if the probability of T given E and B is greater than or equal to r. One obvious way of defining a corresponding qualitative notion of confirmation is to say that E confirms T relative to B just in case the probabiliy of T given E and B is greater than or equal to some fixed value $r \geq 1/2$.[4]

However, despite great efforts there is still no generally accepted definition of (degree of) confirmation. One reason for this is that there are at least two conflicting concepts of confirmation: A concept of confirmation aiming at *informative* theories, which one might call the *loveliness* concept of confirmation[5]; and a concept

[1]Cf. also Hempel (1943) and Hempel/Oppenheim (1945).

[2]Cf. Carnap (1962) and (1952).

[3]Cf. Goodman (1946) and (1947), Helmer/Oppenheim (1945), Hosiasson-Lindenbaum (1940), Kemeny (1953) and (1955), Kemeny/Oppenheim (1952), Lehmann (1955), Shimony (1955).

[4]This is *not* the way Carnap defined qualitative confirmation in chapter VII of his (1962). There he required that the probability of T given E and B be greater than that of T given B in order for E to qualitatively confirm T relative to B.

Nevertheless, the above stipulation seems to be the natural qualitative counterpart for the quantitative notion of the degree of *absolute* confirmation, i.e. $\Pr(T \mid E \wedge B)$. The reason is that later on the difference between $\Pr(T \mid E \wedge B)$ and $\Pr(T \mid B)$ – in whatever way it is measured (cf. Fitelson 2001) – was taken as the degree of *incremental* confirmation, and Carnap's proposal is the natural qualitative counterpart of this notion of incremental confirmation. See sections 9-10.

[5]The term 'loveliness' is borrowed from Lipton (1993).

of confirmation aiming at *true* theories, which we may call the *likeliness* concept of confirmation. The first concept, *loveliness*, expresses our approval of theories that are *informative*, whereas the second concept, *likeliness*, expresses our approval of theories that are probable or likely to be true.

These two concepts of confirmation are conflicting in the sense that the former is an increasing and the latter a decreasing function of the logical strength of the theory T to be assessed. That is, if T logically implies T', then T is at least as (lovely or) informative relative to any E and any B as T', and T' is at least as likely or probable relative to any E and any B as T. Hence

Definition 1 *A relation* $|\!\sim\ \subseteq \mathcal{L} \times \mathcal{L}$ *on a language (set of propositional or first-order sentences closed under negation and conjunction)* $\mathcal{L}$ *is a* loveliness *or* informativeness relation *if and only if for all* $E, H, H' \in \mathcal{L}$*:*

$$E \mid\!\sim H, \quad H' \vdash H \quad \Rightarrow \quad E \mid\!\sim H'.$$

$|\!\sim\ \subseteq \mathcal{L} \times \mathcal{L}$ *is a* likeliness relation *on* $\mathcal{L}$ *just in case for all* $E, H, H' \in \mathcal{L}$*:*

$$E \mid\!\sim H, \quad H \vdash H' \quad \Rightarrow \quad E \mid\!\sim H'.$$

According to HD, E HD-confirms H relative to B if(f) the conjunction of H and B logically implies E. Hence, if E HD-confirms H and H' logically implies H, then E HD-confirms H' relative to B. So HD-confirmation is a loveliness relation. According to IL, E "absolutely confirms" H relative to B iff $\Pr(H \mid E, B) > r$, for some value $r \in [.5, 1]$. Hence, if E absolutely confirms H relative to B and H logically implies H', then E absolutely confirms H' relative to B. So absolute confirmation is a likeliness relation.

More precisely, where '$|\!\sim_{HD}$' stands for confirmation in the sense of Hypothetico-Deductivism, and '$|\!\sim^{r}_{IL}$' stands for qualitative confirmation in the sense of Inductive Logic as explained above:

$$\begin{array}{rcl} E \mid\!\sim_{HD} H, \quad H' \vdash H & \Rightarrow & E \mid\!\sim_{HD} H' \\ E \mid\!\sim^{r}_{IL} H, \quad H \vdash H' & \Rightarrow & E \mid\!\sim^{r}_{IL} H'. \end{array}$$

The epistemic values behind loveliness and likeliness are *informativeness* and *truth*, respectively. Turning back to the question we started from in section 1 – viz.: What is a good theory? – we can say the following: According to HD, a good theory is an informative one, whereas IL says good theories are true.

Putting together the insights of last century's confirmation theory, the answer of the new millenium is this: A *good* theory is *both* informative *and* true. Consequently, we should make these aims explicit in the assessment of theories.

4 Searching Power and Indicating Truth

Given evidence E and background information B, a theory T should be both as informative and as likely as possible. A *strength indicator* s measures how informative T is relative to E and B; a *truth indicator* t measures how likely it is that T is true in view of E and B. Of course, not any function will do.

Definition 2 *Let E be a piece of evidence. A function $f, f : \mathcal{T} \times \mathcal{E} \times \mathcal{B} \rightarrow \Re$,*[6] indicates truth in $Mod(E)$ *iff for any theories T and T' and every body of background information B with $E, B \nvdash \bot$:*

1. $f(T, E, B) \geq 0$
2. $B, E \vdash T \quad \Rightarrow \quad a) \quad f(T, E, B) = 1, \quad [b) \quad f(\neg T, E, B) = 0]$
3. $B, T \vdash T' \quad \Rightarrow \quad f(T, E, B) \leq f(T', E, B)$.

f is a truth indicator *iff f indicates truth in $Mod(E)$, for every piece of evidence E.*

Given the other clauses, the first requirement just stipulates that likeliness is measured on the interval $[0, 1]$.

Part a) of the second requirement states that we cannot demand more – as far as only our aim of arriving at true theories is concerned – than that the evidence our assessment is based on *guarantees* (in the sense of logical implication) that the theory to be assessed is true, given the background assumptions B. Similarly, part b) says that a theory cannot do worse – as far as only our aim at arriving true theories is concerned – than that the conjunction of the data and the background information guarantees that our theory is false.

Finally, the third requirement takes care of the fact that the set of possibilities (possible worlds, models) making a theory T false is a subset of the set of possibilities falsifying any theory that logically implies T, where the set of possibilities is restricted to those that are allowed for by the body of background information B. As a consequence of 3., logically equivalent theories always have the same likeliness.

[6] $\mathcal{T} \subseteq \mathcal{L}$ is the set of all axiomatizations of all theories, and similarly for $\mathcal{E}$ and $\mathcal{B}$. The relativisation to the underlying language L in Mod and elsewhere is suppressed ($\mathcal{L}$ is the set of wffs generated by the first-order language L, and $Mod(X) \subseteq Mod_L$ is the set of possible worlds or models making the wff X true).

Definition 3 *Let E be a piece of evidence. A function $f, f : \mathcal{T} \times \mathcal{E} \times \mathcal{B} \rightarrow \Re$, is an* evidence-based strength indicator for $Mod\,(E)$ *iff for any theories T and T' and every body of background information B with $\neg E, B \nvdash \bot$:*

1. $f\,(T, E, B) \geq 0$
2. $B, T \vdash E \quad \Rightarrow \quad f\,(T, E, B) = 1$
3. $B, T \vdash T' \quad \Rightarrow \quad f\,(T, E, B) \geq f\,(T', E, B)$.

f *is an* evidence-based strength indicator *iff f indicates strength for $Mod\,(E)$, for every piece of evidence E.*

A function $f, f : \mathcal{T} \times \mathcal{B} \rightarrow \Re$, is an evidence-neglecting strength indicator *iff for any theories T and T' and every body of background information B with $B \nvdash \bot$:*

1. $f\,(T, B) \geq 0$
2. $B, T \vdash \bot \quad \Rightarrow \quad f\,(T, B) = 1$
3. $B, T \vdash T' \quad \Rightarrow \quad f\,(T, B) \geq f\,(T', B)$.

A function $f, f : \mathcal{T} \times \mathcal{E} \times \mathcal{B} \rightarrow \Re$, indicates strength with parameter α*, or is an* α-strength indicator*, $\alpha \in [0, 1]$, iff*

$$f\,(T, E, B) = \alpha \cdot f_1\,(T, E, B) + (1 - \alpha) \cdot f_2\,(T, B)\,,$$

where f_1 is an evidence-based, and f_2 an evidence-neglecting, strength indicator.

Given the other clauses, the first requirement in each case again just stipulates that informativeness about the data and informativeness *per se*, as one might call these two notions, are measured on the interval $[0, 1]$.

The second requirement for evidence-based strength indicators says that a theory cannot do better in terms of informing *about the data* than logically implying them. Although this is not questionable, one might object that it is nevertheless inappropriate to ascribe maximal informativeness to any theory logically implying the evidence. The reason is that two theories both logically implying all of the data can still differ in their informativeness: consider, for instance, a complete theory consistent with the data and a theory-like collection of all the data gathered so far (if there is such a collection). This argument is perfectly reasonable. Hence the distinction between evidence-based and evidence-neglecting strength indicators.

As before, the third requirement in each of the two cases simply takes into account that the set of possibilities ruled out or falsified by a theory T is a subset of the set of possibilities ruled out by any theory logically implying T, where the set of possibilities is again restricted to those allowed for by B. So the logically weaker theory T is less informative in that it rules out fewer possibilities than any logically stronger theory. It follows that logically equivalent theories are always equally informative.

The notion of a strength indicator as a weighted mixture of an evidence-based and an evidence-neglecting strength indicator is introduced in order to avoid that one *has to* take sides – though one *can* do so, since the parameter α may be set equal to 0 or to 1. The discussion of how to measure informativeness will be taken up again when the present paradigm-independent theory is spelt out in terms of subjective probabilities.

In all three cases, the third requirement expresses that strength indicators and truth indicators increase and decrease, respectively, with the logical strength of the theory to be assessed. These quantitative requirements correspond to the defining clauses of the qualitative relations of loveliness and likeliness, respectively.

Obviously, an assessment function a should *not* be both a strength indicator and a truth indicator, for there is no such function ($\perp$ would have to be assigned the value 1 in order for a to be a strength indicator, and the value 0 in order to for it be a truth indicator).[7] Let us call this observation the *singularity* of simultaneously indicating strength and truth.

Instead, an assessment function a should *weigh* between these two conflicting aspects: a has to be *sensitive to both informativeness and truth, loveliness and likeliness*.

Definition 4 *Let s be a strength indicator, let t be a truth indicator, and let $\beta \in \Re$. A function $f = f_{s,t} : \mathcal{T} \times \mathcal{E} \times \mathcal{B} \to \Re$ is* sensitive to loveliness and likeliness in the sense of s and t and with demarcation β *– or for short: an s,t-function (with demarcation β) – iff*

0.a Functionality: f is a function of s and t,

0.b Loveliness and Likeliness: f increases with both s and t,

such that it holds for all possible values $s_1, s_2 \in Rg(s) = [0,1]$ of s and all possible values $t_1, t_2 \in Rg(t) = [0,1]$ of t

[7]If condition 2b of indicating truth is not adopted, any strength-indicating truth indicator is a constant function equal to 1, for all T, E, B with $E, B \nvdash \perp$ and $\neg E, B \nvdash \perp$.

1.a *Continuity for Loveliness: Whatever small surplus in loveliness there is, it succeeds if only the difference in likeliness is small enough.*[8]

$$\forall \varepsilon > 0 \quad \exists \delta_\varepsilon > 0 \quad \forall s_1, s_2 \in Rg\,(s) \quad \forall t_1, t_2 \in Rg\,(t) :$$
$$[s_1 > s_2 + \varepsilon \quad \& \quad t_1 > t_2 - \delta_\varepsilon] \quad \Rightarrow \quad f\,(s_1, t_1) > f\,(s_2, t_2)\,.$$

1.b *Continuity for Likeliness: Whatever small surplus in likeliness there is, it succeeds if only the difference in loveliness is small enough.*

$$\forall \varepsilon > 0 \quad \exists \delta_\varepsilon > 0 \quad \forall t_1, t_2 \in Rg\,(t) \quad \forall s_1, s_2 \in Rg\,(s) :$$
$$[t_1 > t_2 + \varepsilon \quad \& \quad s_1 > s_2 - \delta_\varepsilon] \quad \Rightarrow \quad f\,(s_1, t_1) > f\,(s_2, t_2)\,.$$

2. *Demarcation* β

$$\begin{bmatrix} t_1 = 1 & \& & s_1 = 0 \\ & \textit{or} & \\ t_1 = 0 & \& & s_1 = 1 \end{bmatrix} \quad \Rightarrow \quad f\,(s_1, t_1) = \beta.$$

An s, t-function is a function of s and t. Therefore, as the occasion arises, I will sometimes write $f\,(s_1, t_1)$, for some values of s, t, respectively, and other times $f\,(T, E, B)$.

Continuity for Loveliness implies that f increases with s, and Continuity for Likeliness implies that f increases with t. For theorem 2 below Continuity for

[8]This formulation of Continuity is stronger than the following one:

Weak Continuity for Loveliness

$$\forall s_1, s_2 \in Rg\,(s) : s_1 > s_2 \quad \exists \delta_{s_1, s_2} > 0 \quad \forall t_1, t_2 \in Rg\,(t) :$$
$$t_1 > t_2 - \delta_{s_1, s_2} \quad \Rightarrow \quad f\,(s_1, t_1) > f\,(s_2, t_2)\,.$$

The difference is that, in its stronger formulation, Continuity requires δ just to depend on the lower bound ε of the difference between s_1 and s_2, and not on the numbers s_1 and s_2 themselves.

Thus, in case of Weak Continuity, if $s_1 = 0.91$ and $s_2 = 0.8$, and $s'_1 = 0.11$ and $s'_2 = 0$, there may be no *common* upper bound $\delta = \delta_{s_1, s_2} = \delta_{s'_1, s'_2}$ by which t_1 and t'_1 must not be smaller than t_2 and t'_2 in order for $f\,(s_1, t_1) > f\,(s_2, t_2)$ and $f\,(s'_1, t'_1) > f\,(s'_2, t'_2)$, respectively, to hold – the respective upper bounds may be, say, $\delta = .05$ for t_1 and t_2, and $\delta' = .025$ for t'_1 and t'_2.

This is not the case for the version of Continuity assumed to hold of any s, t-function: There the δ is only allowed to depend on the lower bound ε by which s_1 exceeds s_2. Thus, for s_1, s_2 and s'_1, s'_2 there must exist a common δ depending just on the lower bound, say, $\varepsilon = 0.1$ (there are, of course, uncountably many such εs for which there exist (not necessarily distinct) δ_εs).

The difference between Continuity and Weak Continuity plays a crucial role in the proof of theorem 2 below. In a sense, it requires the Bayesian agent with degree of belief function $\Pr$ to be clever enough not to get fooled by the odd convergence behaviour of her degree of belief function.

Likeliness is not needed; nor is Functionality. It suffices that f increases with t. This asymmetry is due to the fact that truth is a qualitative yes-or-no affair: a sentence either is or is not true in some world, whereas informativeness (about some data) is a matter of degree. In case of truth, degrees enter the scence only because we do not know in general, given only the data, whether or not a theory is true in any world the data could be taken from. In case of informativeness, however, degrees are present even if we have a complete and correct assessment of the informational value of the theory under consideration.

The conjunction of Functionality and Loveliness and Likeliness implies the following principles:

3. Likeliness

$$s_1 = s_2 \quad \Rightarrow \quad [f(s_1, t_1) \geq f(s_2, t_2) \quad \Leftrightarrow \quad t_1 \geq t_2]$$

4. Loveliness

$$t_1 = t_2 \quad \Rightarrow \quad [f(s_1, t_1) \geq f(s_2, t_2) \quad \Leftrightarrow \quad s_1 \geq s_2]$$

Continuity implies Loveliness and Likeliness, and

5. Maximality

$$[s = 1 \quad \& \quad t = 1] \quad \Leftrightarrow \quad f(s, t) = \max$$

6. Minimality

$$[s = 0 \quad \& \quad t = 0] \quad \Leftrightarrow \quad f(s, t) = \min.$$

The conjunction of Continuity for Loveliness and Demarcation β implies

7. Limiting Demarcation β

Let $t_1, \ldots, t_n, \ldots$ be a sequence of possible values of t.

$$\left[\begin{array}{c} t_n \to \left\{\begin{array}{ll} 1 & \\ 0 & \text{as } n \to \infty \end{array}\right. \\ s \in (0,1) \end{array}\right] \quad \Rightarrow \quad \exists n \forall m \geq n : f(s, t_m) \begin{array}{c} > \\ < \end{array} \beta$$

However, the conjunction of Continuity and Demarcation β does not imply Symmetry (which, to be clear, is not assumed to hold of an s, t-function):

Symmetry

$$f(s_1, t_1) = f(t_1, s_1).$$

Indeed, it is compatible with Continuity and Demarcation β to consider one aspect, say the likeliness aspect, more important than the other one. This is illustrated by the following s, t-function:

$$f(T, E, B) = s(T, E, B)^x + t(T, E, B), \quad x \in \Re_{>0}.$$

In this respect, the only thing that is ruled out is to totally neglect one of the two aspects.

Furthermore, the conjunction of Continuity and Demarcation β does not imply that for a given value $s_1 \in (0, 1)$ there is a value $t_1 \in (0, 1)$ such that $f(s_1, t_1) = \beta$.[9]

One might want Minimality to run as follows:

Weak Minimality

$$[s = 0 \quad or \quad t = 0] \quad \Leftrightarrow \quad f(s, t) = \min.$$

Weak Minimality and Minimality (and thus Continuity) are, of course, inconsistent. Weak Minimality is also inconsistent with Loveliness and with Likeliness (unless one adds the condition $\neq 0$ to their respective antecedents).

But clearly, a theory which is refuted by the data can still be better than another theory which is also refuted by the data. After all, (almost) every interesting theory from, say, physics, has turned out to be false – and we nevertheless think there has been progress! This, however, is not allowed for by adopting Weak Minimality, for Weak Minimality amounts to saying that in the special case of likeliness being minimal, loveliness does not count anymore; and *vice versa*.

Minimality says that a theory has minimal value only if both loveliness and likeliness are minimal ($= 0$). Similarly, Maximality says that a theory has maximal value only if both loveliness and likeliness are maximal ($= 1$). Finally, and

[9] The functions

$$f_\varepsilon(s, t) = \begin{cases} s + t + \varepsilon, & \text{if} \quad s + t > 1 \\ s + t + \varepsilon/2, & \text{if} \quad s + t = 1 \\ s + t - \varepsilon, & \text{if} \quad s + t < 1 \end{cases}$$

are counterexamples for $\varepsilon > 0$.

before considering some examples, note that any s, t-function is invariant with respect to (or closed under) equivalence transformations, because s and t are so by virtue of the third clauses in definitions 2 and 3.

As an example, consider the following symmetric s, t-function with demarcation $c + 1$:

$$s_c = s(T, E, B) + t(T, E, B) + c,$$

where $c \in \Re$ is some constant. If $c = -1$, then $\beta = 0$, whence

$$d_f = [t(T, E, B) + s(T, E, B) - 1] \cdot f(E, B)$$

satisfies Continuity and Demarcation β. However, unless f is a function of s and t, d_f is not an s, t-function (with demarcation $\beta = 0$) in the sense defined.

5 Loveliness and Likeliness

5.1 The General Theory

What has been seen so far is a general theory of theory assessment which one might call the theory of loveliness and likeliness. In a nutshell, its message is (1) that there are two values a theory should exhibit: informativeness and truth – measured by a loveliness measure or strength indicator s and a likeliness measure or truth indicator t, respectively; (2) that these two values are conflicting in the sense that the former is an increasing, and the latter a decreasing function of the logical strength of the theory to be assessed; and (3) that in assessing a given theory one should weigh between these two conflicting aspects in such a way that any surplus in love(like)liness succeeds, if only the difference in like(love)liness is small enough.

Particular accounts arise by inserting particular strength indicators and truth indicators.

5.2 Loveliness and Likeliness, Bayes Style

The theory can be spelt out in terms of Spohn's ranking theory (cf. Spohn 1988, 1990; see also sections 9-10), and in a logical paradigm that goes back to Hempel (cf. Huber 2002). Here, however, I will focus on the Bayesian version, where I take Bayesianism to be the threefold thesis that (i) scientific reasoning is probabilistic; (ii) probabilities are adequately interpreted as an agent's actual subjective degrees of belief; and (iii) they can be measured by her betting behaviour.

Spelling out the general theory in terms of subjective probabilities simply means that we specify a (set of) probabilistic strength indicators(s) and a (set of) probabilistic truth indicator(s). Everything else is accounted for by the general theory. (Before going on, let me repeat that the present theory of theory assessment is *paradigm-independent* in the sense of not being committed to the credo of any paradigm. This means in particular that I do *not* beg the question by assuming that theory assessment must be probabilistic. What is presented in the following is the Bayesian version of a general theory of theory assessment.)

The nice thing about the Bayesian paradigm is that once one is given theory T, evidence E, and background information B, one is automatically given the relevant numbers $\Pr(T \mid E \wedge B), \ldots$, and the whole problem reduces to the definition of a suitable function of $\Pr$.[10]

In this paradigm it is natural to take

$$t(T, E, B) = \Pr(T \mid E \wedge B) = p$$

as truth indicator, and

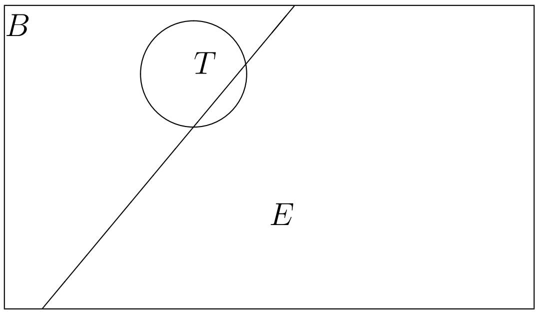
$$s_{e-b}(T, E, B) = \Pr(\neg T \mid \neg E \wedge B) = i, \quad s_{e-n}(T, B) = \Pr(\neg T \mid B) = i'$$

as evidence-based and evidence-neglecting strength indicators, respectively, where $\Pr$ is a *strict* probability.[11]

The choice of p hardly needs any discussion, and for the choice of i consider the following figure with hypothesis or theory T, evidence E, and background

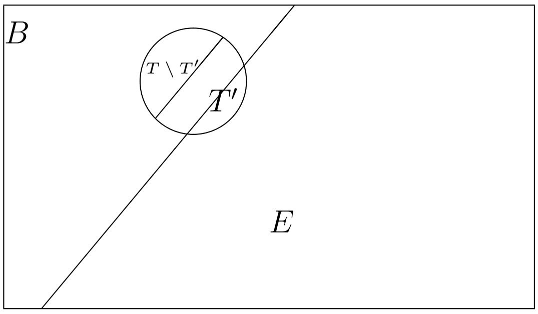
[10] This is not the case in the Hempel paradigm. There the numbers are squeezed out of the logical structure of T, E, and B and nothing else. As a consequence, these values are uniquely determined by T, E, and B and the logical consequence relation. In particular, they are independent of the underlying language (cf. Huber 2002).

[11] The strictness of $\Pr$ is of importance, as will be seen later on. Strictness is often paraphrased as open-mindedness (e.g. Earman 1992), because it demands that no consistent statement is assigned probability 0. Given a subjective interpretation of probability, this sounds like a restriction on what one is allowed to believe (to some degree). Strictness can also be formulated as saying that any statement H_1 which logically implies but is not logically implied by some other statement H_2 – i.e. H_1 strictly implies H_2 – must be assigned a strictly lower degree of belief than H_2. (In case of probabilities conditional on B, logical implication is also conditional on B.) Seen this way, strictness requires degrees of belief which are sufficiently fine-grained. For this reason I prefer to think of strictness not as a restriction on *what* (which propositions) to believe (to some degree), but as a restriction on *how* to believe (propositions), namely, sufficiently fine-grained so that differences so big as to be expressable purely in terms of the logical consequence relation are not swept under the carpet.



information B.

Suppose you are asked to strengthen T by deleting possibilities verifying it, that is, by shrinking the area representing T.[12] Would you not delete possibilities outside E? After all, given E, those are exactly the possibilities known not to be the actual one, whereas those possibilities inside E are still alive options. Indeed, our probabilistic evidence-based strength indicator i increases when T shrinks to



T' as depicted above.

For the probabilistic evidence-neglecting strength indicator i' it does not matter which possibilities one deletes in strengthening T (provided all possibilities have equal weight on the probability measure $\Pr$). i' neglects whether they are inside or outside E.

Finally, the strength indicator i^*_α with parameter $\alpha \in [0, 1]$ is given by

$$i^*_\alpha = \alpha \cdot \Pr(\neg T \mid \neg E, B) + (1 - \alpha) \cdot \Pr(\neg T \mid B) = \alpha \cdot i + (1 - \alpha) \cdot i'.$$

For i^*_α, it depends on α how much it matters whether the deleted possibilities lie inside or outside of E.

Other candidates for measuring informativeness that are (suggested by mea-

[12] I owe this graphical illustration to Luc Bovens.

sures) discussed in the literature[13] are

$$\begin{aligned} i'' &= \Pr(\neg T \mid E \wedge B), \\ \text{cont} &= \Pr(E) \cdot \Pr(\neg T \mid E \wedge B), \\ \inf &= -\log_2 \Pr(T \mid E \wedge B). \end{aligned}$$

These measures, all of which assign minimal informativeness to any theory entailed by the data and the background assumptions, do even worse on this count by requiring the deletion of the possibilities inside E. Another reason why i'', cont, and inf seem to be inapproapriate for measuring informativeness is presented in section 5.4.

Note that the body of background information B plays a role different from that of the piece of evidence E for i, i', and i^*_α, but not for i'', cont, or inf. Clearly, if there is a difference between data on the one hand and background assumptions on the other, then this difference should show up somewhere. Background assumptions, so to speak, determine the set of possibilities in the inquiry, and thus are nothing but restrictions on the set of possible worlds over which inquiry has to succeed (cf. Hendricks 2004). Furthermore, evidence-based strength indicators measure how much a theory informs about the data, but not how much they inform about the background assumptions. However, if one holds there should be no difference between E and B as far as measuring informativeness is concerned, then one can nevertheless adopt the above measures by substituting $E' = E \wedge B$ and $B' = \top$ for E and B, respectively.

5.3 Incremental Confirmation

Let us see how this approach compares to current Bayesian confirmation theory. The following notion is central in this literature (cf. Fitelson 2001): A function $f = f_{\Pr} : \mathcal{L} \times \mathcal{L} \times \mathcal{L} \to \Re$ is a *β-relevance measure based on* $\Pr$ just in case it holds for all $T, E, B \in \mathcal{L}$ with $\Pr(E \wedge B) > 0$:

$$f(T, E, B) \begin{array}{c} > \\ = \\ < \end{array} \beta \quad \Leftrightarrow \quad \Pr(T \mid E \wedge B) \begin{array}{c} > \\ = \\ < \end{array} \Pr(T \mid B).$$

[13] Cf. Carnap/Bar-Hillel (1952), Bar-Hillel/Carnap (1953), and Hintikka/Pietarinen (1966); cf. also Bar-Hillel (1952) and (1955). In Levi (1967), i'' is proposed as, roughly, a measure for the relief from agnosticism afforded by accepting T as strongest relative to total evidence $E \wedge B$.

As

$$\Pr(T \mid E \wedge B) > \Pr(T \mid B) \quad \Leftrightarrow \quad \Pr(\neg T \mid \neg E \wedge B) > \Pr(\neg T \mid B) \quad (1)$$

for $0 < \Pr(E \mid B) < 1$ and $\Pr(B) > 0$ (which is true for every E and B with $\neg E, B \nvdash \bot$ and $E, B \nvdash \bot$, because Pr is assumed to be strict – cf. the definitions of indicating strength and truth), every i, p-function

$$s_c = p + i + c, \quad c \in \Re,$$

is a $c+1$-relevance measure in the Bayesian sense (where p and i depend on $\Pr$).

For $c = -1$, one gets the Joyce-Christensen measure s,

$$s_{\Pr}(T, E, B) = \Pr(T \mid E \wedge B) - \Pr(T \mid \neg E \wedge B)$$

(cf. Joyce 1999 and Christensen 1999). As noted earlier at the end of section 4, for positive f not depending on T, the functions

$$d_f = [i + p - 1] \cdot f(E, B)$$

satisfy Continuity and Demarcation $\beta = 0$ for i and p. For $f = \Pr(\neg E \mid B)$ we get the distance measure d,

$$d_{\Pr}(T, E, B) = \Pr(T \mid E \wedge B) - \Pr(T \mid B),$$

and for $f = \Pr(\neg E \mid B) \cdot \Pr(B) \cdot (E \wedge B)$ we get the Carnap measure c,

$$c_{\Pr}(T, E, B) = \Pr(T \wedge E \wedge B) \cdot \Pr(B) - \Pr(T \wedge B) \cdot \Pr(E \wedge B)$$

(cf. Carnap 1962). Hence the Carnap measure c, the difference measure d, and Joyce-Christensen measure s seem to be three different ways of weighing between the two functions i and p (or between i' and p, for $s = d/\Pr(\neg E \mid B)$ and $c = d \cdot \Pr(B) \cdot \Pr(E \wedge B)$).

Alternatively, the difference between d and s can be seen not as one between the way of weighing, but as one between *what* is weighed – namely two different pairs of functions, viz. i and p for the difference measure d, and i' and p for the Joyce-Christensen measure s. This is clearly seen by rewriting d and s as

$$\begin{aligned} d_{\Pr} &= \Pr(T \mid E \wedge B) + \Pr(\neg T \mid B) - 1, \\ s_{\Pr} &= \Pr(T \mid E \wedge B) + \Pr(\neg T \mid \neg E \wedge B) - 1. \end{aligned}$$

In this sense, the discussion about the right measure of confirmation turns out to be a discussion about the right measure of informativeness of a theory relative to a body of evidence. This view is endorsed by the observation that d and s actually employ the same decision-theoretic considerations:[14]

$$\begin{aligned}
d_{\Pr} &= \Pr(T \mid E \wedge B) - \Pr(T \mid B) \\
&= \Pr(T \mid E \wedge B) - \Pr(T \mid B) \cdot \Pr(T \mid E \wedge B) - \\
&\quad - \Pr(T \mid B) + \Pr(T \mid B) \cdot \Pr(T \mid E \wedge B) \\
&= (1 - \Pr(T \mid B)) \cdot \Pr(T \mid E \wedge B) - \Pr(T \mid B) \cdot (1 - \Pr(T \mid E \wedge B)) \\
&= \Pr(\neg T \mid B) \cdot \Pr(T \mid E \wedge B) - \Pr(T \mid B) \cdot \Pr(\neg T \mid E \wedge B) \\
&= i'(T, B) \cdot \Pr(T \mid E \wedge B) - i'(\neg T, B) \cdot \Pr(\neg T \mid E \wedge B),
\end{aligned}$$

and

$$\begin{aligned}
s_{\Pr} &= \Pr(T \mid E \wedge B) - \Pr(T \mid \neg E \wedge B) \\
&= \Pr(T \mid E \wedge B) - \Pr(T \mid \neg E \wedge B) \cdot \Pr(T \mid E \wedge B) - \\
&\quad - \Pr(T \mid \neg E \wedge B) + \Pr(T \mid \neg E \wedge B) \cdot \Pr(T \mid E \wedge B) \\
&= (1 - \Pr(T \mid \neg E \wedge B)) \cdot \Pr(T \mid E \wedge B) - \\
&\quad - \Pr(T \mid \neg E \wedge B) \cdot (1 - \Pr(T \mid E \wedge B)) \\
&= \Pr(\neg T \mid \neg E \wedge B) \cdot \Pr(T \mid E \wedge B) - \Pr(T \mid \neg E \wedge B) \cdot \Pr(\neg T \mid E \wedge B) \\
&= i(T, E, B) \cdot \Pr(T \mid E \wedge B) - i(\neg T, E, B) \cdot \Pr(\neg T \mid E \wedge B).
\end{aligned}$$

So d and s are exactly alike in the way they combine or weigh between loveliness and likelinesss – which is to form the expected informativeness of the theory (about the data and relative to the background assumptions); their difference lies in the way they measure informativeness.

Given that the theory of loveliness and likeliness has such a nice justification in terms of conflicting epistemic virtues, given that it can be motivated historically[15],

[14]Cf. Hintikka/Pietarinen (1966), and Levi (1961) and (1963), but also Hempel (1960).

[15]As to historical motivation, the ideas behind the strength indicator LO and the truth indicator LI in the Hempel paradigm go back to Hempel's prediction and satisfaction criteria (which is why that paradigm is called Hempel paradigm). It is interesting to see that these two criteria are both present in Hempel's seminal paper on confirmation, who thus seems to have felt the need of both the loveliness and the likeliness concept of confirmation (see also Hempel 1960). This is particularly revealing as his triviality result that every observation report confirms every theory basically rests on the fact that loveliness is an increasing and likeliness a decreasing function of the logical strength of the theory (cf. sections 9-10), and thus amounts to the singularity observation of section 4.

and given that – due its generality – it is free from being committed to the credo of any paradigm, current Bayesian confirmation theory should warmly welcome the connection between i^*_α, p-functions (for $\alpha = 0$ and $\alpha = 1$) and β-relevance measures afforded by the i, p-function s and the i', p-function d. This should be so the more, since in the light of the present approach part of the discussion about the right Bayesian measure of incremental confirmation is one about the right measure of informativeness. Finally, as will be seen in sections 6-8, the present theory provides the only reasonable (and not yet occupied) answer to the question of justification, i.e. why one should stick to well-confirmed theories rather than to any other theories.

5.4 Expected Informativeness

What results do we get from the decision-theoretic way of setting confirmation equal to the expected informativeness for the measures i'', cont, and inf mentioned in section 5.2? "Not many!" one is inclined to answer, where '$i''(T)$' is short for '$i''(T, E, B)$' (similarly for 'cont (T)' and 'inf (T)'):

$$
\begin{aligned}
E\left(i''(T)\right) &= i''(T) \cdot \Pr(T \mid E \wedge B) - i''(\neg T) \cdot \Pr(\neg T \mid E \wedge B) \\
&= \Pr(\neg T \mid E \wedge B) \cdot \Pr(T \mid E \wedge B) - \\
&\quad - \Pr(T \mid E \wedge B) \cdot \Pr(\neg T \mid E \wedge B) \\
&= 0 \\
E\left(\text{cont}(T)\right) &= \text{cont}(T) \cdot \Pr(T \mid E \wedge B) - \text{cont}(\neg T) \cdot \Pr(\neg T \mid E \wedge B) \\
&= \Pr(E) \cdot \Pr(\neg T \mid E \wedge B) \cdot \Pr(T \mid E \wedge B) - \\
&\quad - \Pr(E) \cdot \Pr(T \mid E \wedge B) \cdot \Pr(\neg T \mid E \wedge B) \\
&= 0 \\
E\left(\text{inf}(T)\right) &= \text{inf}(T) \cdot \Pr(T \mid E \wedge B) - \text{inf}(\neg T) \cdot \Pr(\neg T \mid E \wedge B) \\
&= -\log_2 \Pr(\neg T \mid E \wedge B) \cdot \Pr(T \mid E \wedge B) + \\
&\quad + \log_2 \Pr(T \mid E \wedge B) \cdot \Pr(\neg T \mid E \wedge B) \\
&\begin{smallmatrix} > \\ = \\ < \end{smallmatrix}\ 0 \\
&\Leftrightarrow\ \Pr(T \mid E \wedge B) \begin{smallmatrix} > \\ = \\ < \end{smallmatrix} \Pr(\neg T \mid E \wedge B).
\end{aligned}
$$

Hence only inf gives a non-trivial answer, viz. to maximize probability. Maximizing probability is also what the "Acceptance rule based on relative-content measure of utility" from Hempel (1960) requires (I have dropped the body of background information B, because Hempel does not have it, and I took his content measure $m(\cdot)$ to be $1 - \Pr(\cdot)$, which is in accordance with his remarks on p. 76 of Hempel 1965 and with Hempel 1962 and Hempel/Oppenheim 1948). Hempel's "Relative-content measure of purely scientific utility" is this:

$$\begin{aligned}
\mathrm{rc}(T,E) &= i_H(T,E)\cdot\Pr(T\mid E) - i_H(T,E)\cdot\Pr(\neg T\mid E)\\
&= \frac{\Pr(\neg T\wedge E)}{\Pr(\neg E)}\cdot\Pr(T\mid E) - \frac{\Pr(\neg T\wedge E)}{\Pr(\neg E)}\cdot\Pr(\neg T\mid E)\\
&= \frac{\Pr(\neg T\wedge E)}{\Pr(\neg E)}\left(2\cdot\Pr(T\mid E)-1\right).
\end{aligned}$$

However, as noted by Hintikka/Pietarinen (1966), fn. 12, it seems more adequate to consider:

$$\begin{aligned}
E(i_H(T,E)) &= i_H(T,E)\cdot\Pr(T\mid E\wedge B) - i_H(\neg T,E)\cdot\Pr(\neg T\mid E\wedge B)\\
&= \frac{\Pr(\neg T\wedge E)}{\Pr(\neg E)}\cdot\Pr(T\mid E) - \frac{\Pr(T\wedge E)}{\Pr(\neg E)}\cdot\Pr(\neg T\mid E)\\
&= 0.
\end{aligned}$$

Given this result, it is clear why Hintikka/Pietarinen (1966) chose $i' = \Pr(\neg T)$ as measure of information, and thus arrived at the distance measure d as shown above.

In order to finish our digression into decision-theory – the point of which was to show that taking the expected informativeness as assessment value is allowed, but not required by the Bayesian version of the theory of loveliness and likeliness – we give the expected informativeness for the measures i^*_α, $\alpha\in[0,1]$:

$$\begin{aligned}
E(i^*_\alpha(T,E,B)) &= i^*_\alpha(T,E,B)\cdot\Pr(T\mid E\wedge B) - i^*_\alpha(\neg T,E,B)\cdot\Pr(\neg T\mid E\wedge B)\\
&= [\alpha\cdot\Pr(\neg T\mid\neg E\wedge B) + (1-\alpha)\cdot\Pr(\neg T\mid B)]\cdot\Pr(T\mid E\wedge B) -\\
&\quad - [\alpha\cdot\Pr(T\mid\neg E\wedge B) + (1-\alpha)\cdot\Pr(T\mid B)]\cdot\Pr(\neg T\mid E\wedge B)\\
&= \alpha\cdot(\Pr(T\mid E\wedge B) - \Pr(T\mid\neg E\wedge B)) +\\
&\quad + (1-\alpha)\cdot(\Pr(T\mid E\wedge B) - \Pr(T\mid B))\\
&= \alpha\cdot s_{\Pr} + (1-\alpha)\cdot d_{\Pr}.
\end{aligned}$$

5.5 Is It Likely To Be Lovely?

The terms (and only the terms) 'loveliness' and 'likeliness' are taken from Lipton (1993) who suggests[16] the view that a theory which is lovely in his sense (which provides a lot of good explanations) is also likely to be true.

Loveliness, as understood here, is an indicator of the informativenss of a theory, and thus need not have anything to do with explanation. Still, one might ask whether "it is likely to be lovely".

The first way to make this question more precise is to ask whether, given no data at all, a lovely theory is also a likely one. This is, of course, not the case, as is clear from the fact that loveliness and likeliness are conflicting in the sense that the former is an increasing, and the latter a decreasing function of the logical strength of the theory to be assessed.

However, the equivalence in (1) gives rise to another way of putting this question: Given that a piece of evidence E raises the loveliness of T relative to B, does that piece of evidence also raise the likeliness of T relative to B?[17]

Let $E_0, \ldots, E_{n-1}, E_n$ be the evidence seen up to stage $n+1$ of the inquiry. Then the answer is affirmative if, at stage $n+1$, one considers the *total available evidence* $E = E_0 \wedge \ldots \wedge E_{n-1} \wedge E_n$ and asks whether the likeliness of T given the total available evidence E and background information B is greater than the likeliness of T at stage 0 before the first datum came in, i.e. whether

$$\Pr(T \mid E \wedge B) > \Pr(T \mid B).$$

As we have just seen, this holds if and only if the loveliness of T relative to the total available evidence E and background information B, $\Pr(\neg T \mid \neg E \wedge B)$, is greater than T's loveliness at stage 0, when it was equal to $\Pr(\neg T \mid B)$.[18] So on the global scale, lovely theories are likely to be true.

However, the answer is negative on the local scale where one considers just the single datum E_n. At stage n, the loveliness and the likeliness of T relative to

[16] Whether or not he in fact holds this view is another question. It may be doubted, since Lipton applies the term 'lovely' to explanations, and not to theories.

[17] According to the probabilistic evidence-neglecting strength indicator i', the informativeness of a theory is independent of the data, and so it does not make sense to ask whether a piece of evidence E raises the loveliness – in the sense of i' – of some theory T relative to a body of background information B. Therefore only i is considered in the following.

[18] It may justifiedly be argued that the loveliness of T at stage 0 before the first datum came in is *not* $\Pr(\neg T \mid B)$, but rather is not defined. This follows if the "empty datum", i.e. the one before the first datum came in, is represented by $\top$. Stipulating that s_0 is defined and equal to $\Pr(\neg T \mid B)$ should only enable me to make sense of the question whether it is likely to be lovely.

B and the data seen so far are given by

$$s_n = \Pr(\neg T \mid \neg(E_0 \wedge \ldots \wedge E_{n-1}) \wedge B)$$

and

$$t_n = \Pr(T \mid E_0 \wedge \ldots \wedge E_{n-1} \wedge B).$$

Now suppose the next datum E_n at stage $n+1$ raises the loveliness of T relative to B and the data seen so far, i.e.

$$\begin{aligned} s_{n+1} &= \Pr(\neg T \mid \neg(E_0 \wedge \ldots \wedge E_{n-1} \wedge E_n) \wedge B) \\ &> \Pr(\neg T \mid \neg(E_0 \wedge \ldots \wedge E_{n-1}) \wedge B) \\ &= s_n. \end{aligned}$$

Does it follow that

$$\begin{aligned} t_{n+1} &= \Pr(T \mid E_0 \wedge \ldots \wedge E_{n-1} \wedge E_n \wedge B) \\ &> \Pr(T \mid E_0 \wedge \ldots \wedge E_{n-1} \wedge B) \\ &= t_n? \end{aligned}$$

It does not; what holds true is that

$$\begin{aligned} t_{n+1} &> t_n \\ &\Leftrightarrow \\ \Pr(\neg T \mid E_0 \wedge \ldots \wedge E_{n-1} \wedge \neg E_n \wedge B) &> \Pr(\neg T \mid E_0 \wedge \ldots \wedge E_{n-1} \wedge B) = t_{n-1}, \end{aligned}$$

given that the relevant probabilities are non-negative. But t_{n+1} may be smaller than t_n, even if $s_{n+1} > s_n$.[19] Thus, on the global scale a lovely theory is also a likely one, though this does not hold true on the local scale, where single pieces of evidence are considered.

[19] The same holds true on both the local *and the global* scale, if one takes the measure $i'' = \Pr(\neg T \mid E \wedge B)$ instead of $\Pr(\neg T \mid \neg E \wedge B)$. For

$$\Pr(T \mid E \wedge B) + \Pr(\neg T \mid E \wedge B) = 1,$$

for all T, E, B with $\Pr(E \wedge B) > 0$, so

$$\Pr(\neg T \mid E_0 \wedge \ldots \wedge E_{n-1} \wedge E_n \wedge B) < \Pr(\neg T \mid E_0 \wedge \ldots \wedge E_{n-1} \wedge B) \quad \text{and}$$
$$\Pr(\neg T \mid E \wedge B) < \Pr(\neg T \mid B),$$

if

$$\Pr(T \mid E_0 \wedge \ldots \wedge E_{n-1} \wedge E_n \wedge B) > \Pr(\neg T \mid E_0 \wedge \ldots \wedge E_{n-1} \wedge B) \quad \text{and}$$
$$\Pr(T \mid E \wedge B) > \Pr(T \mid B),$$

5.6 Selected Success Stories

This section briefly indicates how the present account of theory assessment is successfully applied to various problems in confirmation theory. Being notoriously short, the discussion of these problems does not do justice to their seriousness.

5.6.1 Problems in Bayesian Confirmation Theory

As noted in Christensen (1999), the following are among the specific problems of (subjective)[20] Bayesian confirmation theory:

1. the symmetry of confirmation: E incrementally confirms T if and only if T incrementally confirms E; and

2. the problem of old evidence: If E is known in the sense of being assigned a degree of belief of 1, then E incrementally confirms no T,

where the background information B has been neglected in both cases.

According to the present approach confirmation is not symmetric. This point is elaborated in sections 9-10. There we require that T be at least as likely as and more informative than (or more likely than and at least as likely as) its negation given E in order for E to confirm T in the combined sense of loveliness and likeliness. Furthermore, confirmation is "antisymmetric" in the Hempel paradigm, because there are no finite sets of statements that formulate both a theory and evidence (cf. Huber 2002).

---

respectively. Though i'' is a decreasing function of the logical strength of T, it is not an evidence-based strength indicator in the sense defined, because $\Pr(\neg T \mid E \wedge B)$ need not equal 1 if $T, B \vdash E$. Moreover, according to the i'', p-function $s''_c = i'' + p + c$, every theory T has the same value $\beta = c + 1$ independently of the given evidence E and background information B.

As I learned recently, Levi (in personal correspondence) now favours $i' = \Pr(\neg T \mid B)$ as a measure of the informativeness of T given B. According to this measure, informativeness is a virtue of a theory T relative to background information B which is independent of the data E. This is not true for $i = \Pr(\neg T \mid \neg E \wedge \neg B)$.

Interestingly, i' violates a condition of adequacy Levi himself (cf. Levi 1986) holds: Any two theories which are logically equivalent given evidence E and background knowledge B should be assigned the same value. This condition does not hold of i, p-functions and has the consequence that any two refuted theories are assigned the same value – which, given the history of science, is, of course, inappropriate for a theory of theory assessment.

[20]Though the importance of interpreting $\Pr$ is often dismissed in Bayesian confirmation theory, some problems – e.g. the problem of old evidence – arise only under particular interpretations of probability. In this case it is the subjective interpretation that takes $\Pr$ to be an agent's actual degree of belief function.

It is clear that the problem of old evidence does not occur in the Hempel paradigm, for there one is not working with subjective probabilities. More importantly, note that, by Jeffrey conditionalisation,

$$i = \Pr(\neg T \mid \neg E \wedge B) \quad \text{and} \quad p = \Pr(T \mid E \wedge B)$$

and thus all functions which are sensitive to loveliness and likeliness in the sense of i and p are invariant with respect to changes in $\Pr(E \mid B)$. This means that no i, p-function faces the more general version of the problem of old evidence. The general version of the problem of old evidence is that T is more confirmed by E relative to B in the sense of $\Pr_2$ than in the sense of $\Pr_1$ just in case $\Pr_2(E \mid B) < \Pr_1(E \mid B)$, where $\Pr_2$ results from $\Pr_1$ by Jeffrey conditioning on E, and E is positively relevant for T given B in the sense of $\Pr_1$.[21] In other words, the problem is that the less reliable the source of information, the higher the degree of confirmation. (The traditional problem of old evidence – i.e. the special case where $\Pr(E \mid B) = 1$ – does not arise, because $\Pr$ is strict, and it is assumed that $\neg E, B \nvdash \perp$ and $E, B \nvdash \perp$.) The more general version of the problem of old evidence is faced by the distance measure d, the log-likelihood ratio l, and the ratio measure r (cf. Huber 2004).

5.6.2 Tacking by Conjunction

If evidence E confirms theory T relative to background information B, then E generally does not confirm (relative to B) the conjunction of T and an arbitrary theory T'. This is in accordance with the present approach, for although adding T' does not decrease the informativeness or loveliness of T relative to E and B, it generally does lead to a decrease in the likeliness of T relative to E and B.

5.6.3 Theory Hostility

It is sometimes claimed that confirmation is inappropriate for the assessment of theories, because confirmation does not take into account the fact that theories should possess several other virtues besides being true or having a high probability. This exclusive focus on truth (or probability) is referred to as theory hostility. An adequate theory of confirmation or theory assessment should yield that theories that are well-confirmed (in its sense) should not only be true or probable; they

[21]In case E is negatively relevant for T given B in the sense of $\Pr_1$, this holds just in case $\Pr_2(E \mid B) > \Pr_1(E \mid B)$. Negative evidence provides more disconfirmation and positive evidence provides more confirmaton, the lower the degree of belief in it.

should also be informative. Obviously this holds of any theory which is confirmed in the combined sense of loveliness and likeliness.

6 What Is the Point?

The crucial question any theory of theory assessment has to face is this: What is the point of having theories that are given high assessment values? That is, why are theories given high assessment values better than any other theories? In terms of confirmation the question is: What is the point of having theories that are well confirmed by the available data relative to some background information? That is, why should we stick to well confirmed theories rather to any other theories?

The traditional answer to this question is that the goal is truth, and that one should stick to well confirmed theories because (in the long run) confirmation takes you to the truth.[22] But as we have seen, truth is only one side of the coin – the other is informativeness. Thus, the answer of the new millenium is that the goal is informative truth, and that one should stick to theories which are given high assessment values because (in the medium run) theory assessment takes you to the most informative among all true theories.

Indeed, if being taken to the most informative among all true theories is not the goal of confirmation theory, it seems there is no point to current (i.e. incremental) Bayesian confirmation theory at all. The traditional approach to confirmation the early Carnap had, i.e. absolute confirmation setting confirmation equal to (logical) probability, has long been abandoned in favour of incremental confirmation setting confirmation equal to increase in probability. Though I do not want to opt for a revival of absolute confirmation, it must be said for reasons of fairness that absolute confirmation at least could be justified by arguing that, in the long run, absolute confirmation takes one to the truth (which is the content of the Gaifman and Snir theorem – cf. the next section).

So if the goal were truth and only truth, there would have been no need for abandoning absolute confirmation. Hence there must be another goal for incremental confirmation. But then, if arriving at (the most) informative (among all) true theories is not the goal of incremental confirmation, what else could it be? Yet, as will be seen in section 8, incremental confirmation does not take one to informative true theories nor to the most informative among all true theories.

[22]This is the line of reasoning the early Carnap (could have) had. He held that confirmation is equal to (logical) probability – absolute confirmation, in contrast to incremental confirmation.

Let us turn back to how theory assessment takes one to the most informative among all true theories. Given a possible world (possibility, model) ω, contingent theory T_1 is to be preferred over contingent theory T_2 in ω if

1. T_1 is true in ω, but T_2 is false in ω; or
2. both T_1 and T_2 are true in ω, but T_1 logically implies T_2, whereas T_2 does not logically imply T_1 (i.e. T_1 is logically stronger than T_2); or
3. both T_1 and T_2 are false in ω, but T_1 is logically stronger T_2.

In case T_1 is logically false, it is worse in ω than every contingent theory T_2 that is true in ω (because they are all true in ω, whereas T_1 is false in ω), but better than every contingent theory T_2 that is false in ω (because T_1 is more informative than each of them). Similarly, if T_1 is logically true, it is worse in ω than every contingent theory T_2 that is true in ω (because they all are more informative than T_1), but better than every contingent theory T_2 that is false in ω (because they all are false in ω, whereas T_1 is true in ω).

Consequently, a function $f, f : \mathcal{T} \times \mathcal{E} \times \mathcal{B} \rightarrow \Re$, is said to *reveal the true assessment* (or *confirmational*) *structure in world* ω if and only if, for some $\beta \in \Re$, for any two contingent theories $T_1, T_2 \in \mathcal{T}$, every logically determined theory $T \in \mathcal{T}$,[23] every background information $B \in \mathcal{B}$ which is true in ω, and any data stream $e_0, \ldots, e_n, \ldots$ ($e_i \in \mathcal{E}$) from ω (i.e. a sequence of sentences expressing distinct propositions all of which are true in ω):

1. If T_1 is true in ω and T_2 is false in ω, then there is a point n such that for all later points $m \geq n$:

$$f(T_1, E_m, B) > \beta > f(T_2, E_m, B);$$

2. if T_1 and T_2 are true in ω, but T_1 is logically stronger than T_2, then there is a point n such that for all later points $m \geq n$:

$$f(T_1, E_m, B) > f(T_2, E_m, B) > \beta;$$

3. if T_1 and T_2 are false in ω, but T_1 is logically stronger than T_2, then there is a point n such that for all later points $m \geq n$:

$$\beta > f(T_1, E_m, B) > f(T_2, E_m, B);$$

[23] It may, of course, be doubted that a scientific theory can be logically determined.

4. if T is logically determined, then it holds for all m:

$$f(T, E_m, B) = \beta;$$

where $E_m = e_0 \wedge \ldots \wedge e_m$.

So f must *stabilize* to the correct answer, i.e. get it right after finitely many steps, and continue to do so forever without necessarily halting (or giving any other sign that it has arrived at the correct answer).[24] The smallest n for which the above holds is called the point of stabilisation; and β is called the demarcation parameter.

The central question is, of course, whether the measures presented do in fact further the goal they are supposed to further, that is, whether they in fact reveal the true assessment structure and thus lead to true and informative theories. Not surprisingly, the answer is: YES! And this is as good a justification for a theory of theory assessment as there can be[25].

More precisely, let $e_0, \ldots, e_n, \ldots$ be a sequence of sentences of L which separates Mod_L, and let e_i^ω be e_i, if $\omega \models e_i$, and $\neg e_i$ otherwise, where $\omega \in Mod_L$. Let $\Pr$ be a strict probability on $\mathcal{L}$, and let $a = a_{\Pr}$ be a function satisfying 0.b Loveliness and Likeliness, 1.a Continuity for Loveliness, and 2. Demarcation β (but is not necessarily an i, p-function) for i and p (e.g. $s_c = i + p + c$, where $\beta = c + 1$). Finally, let $\Pr^*$ be the unique probability measure on the smallest σ-field containing the field $\{Mod(A) : A \in \mathcal{L}\}$, where $Mod(A) = \{\omega \in Mod_L : \omega \models A\}$.

Then there exists $X \subseteq Mod_L$ with $\Pr^*(X) = 1$ such that the following hold for every $\omega \in X$, any two contingent $H_1, H_2 \in \mathcal{L}$, and every $H \in \mathcal{L}$:

[24] Stabilisation to the correct answer is a stronger requirement than convergence to the correct answer (see Kelly 1996). The latter is a bit odd to formulate for revealing the true assessment structure, but in general says that for any $\varepsilon > 0$ (as small as you like) there exists a point n (depending on ε) such that for all later points $m \geq n$ f's conjecture differs form "the truth" only by an amount smaller than ε. The difference between stabilisation and convergence was the reason for appealing to the medium run (stabilisation) as compared to the long run (convergence). Note, however, that the Gaifman-Snir convergence theorem can be used to obtain an almost-sure stabilisation result in the above spirit by assigning 1 to H, if the probability of H is above $.5$, and 0 otherwise (cf. section 8).

[25] ... except for the fact that the result stated below holds only for almost every world and is restricted to data sequences that separate Mod_L. However, this is not due to anything being wrong with the theory of loveliness and likeliness. Instead, this flaw is inherited from the present paradigm. The flaw is serious (Kelly 1996, ch. 13), but not inevitable, because there are other paradigms one might adopt such as ranking theory (Spohn 1988, 1990), where "pointwise reliability" is possible (Kelly 1999). However, the price of pointwise reliability is that the set of possible worlds be countable, and it is fair to say that measure 1 results are not problematic in this case.

1. $\omega \models H_1, \omega \not\models H_2 \quad \Rightarrow \quad \exists n \forall m \geq n : a(H_1, E_m^\omega) > \beta > a(H_2, E_m^\omega)$

2. $\begin{bmatrix} \omega \models H_1, \omega \models H_2 \\ H_1 \vdash H_2, H_2 \not\vdash H_1 \end{bmatrix} \Rightarrow \exists n \forall m \geq n : a(H_1, E_m^\omega) > a(H_2, E_m^\omega) > \beta$

3. $\begin{bmatrix} \omega \not\models H_1, \omega \not\models H_2 \\ H_1 \vdash H_2, H_2 \not\vdash H_1 \end{bmatrix} \Rightarrow \exists n \forall m \geq n : \beta > a(H_1, E_m^\omega) > a(H_2, E_m^\omega)$

4. $\models H$ or $\models \neg H \quad \Rightarrow \quad \forall m : a(H, E_m^\omega) = \beta$,

where $E_m^\omega = \bigwedge_{0 \leq i \leq m} e_i^\omega$.

In other words, every continuously demarcating function reveals the true assessment structure in almost every world when presented with data separating the set of all possible worlds! The relativisation to the body of background information B has been dropped. The above entails that there exists $X \subseteq Mod_L$ with $\Pr^*(X \mid Mod(B)) = 1$, for every $B \in \mathcal{L}$ with $\Pr(B) > 0$, such that 1.-4. hold for every $\omega \in Mod(B)$.

7 Making the Point More Precise

This section makes the claim of the last section more precise. The framework adopted here is from Gaifman and Snir (1982). L_0 is a first order language for arithmetic, which contains all numerals '1', '2', ... as individual constants, and countably many individual variables 'x_1', ... taking values in the set of natural numbers N. Furthermore, L_0 contains the common symbols '$+$', '$\cdot$', and '$=$' for addition, multiplication, and identity, respectively. In addition, there may be finitely many predicates and function symbols denoting certain fixed relations over N. Finally, L_0 contains the quantifiers '$\forall$', '$\exists$', the unary sentential connective '$\neg$', and the binary sentential connectives '$\wedge$', '$\vee$', '$\rightarrow$', and '$\leftrightarrow$'.

The language L is obtained from L_0 by adding finitely many predicates and function symbols. All symbols in L are used autonymously. The set of well-formed formulas of L is denoted by '$\mathcal{L}$'.

A model ω for L consists of an interpretation φ of the empirical symbols which assigns every k-ary predicate 'P' a subset $\varphi(\text{'}P\text{'}) \subseteq N^k$, and every k-ary function symbol 'f' a function $\varphi(\text{'}f\text{'})$ from N^k to N. The interpretation of the symbols in L_0 is the standard one and is kept the same in all models.

Mod_L is the set of all models for L. '$\omega \models A$' says that formula A is true in model $\omega \in Mod_L$. $A[x_1, \ldots, x_k]$ is valid, $\models A[x_1, \ldots, x_k]$, iff $\omega \models A[n_1/x_1, \ldots, n_k/x_k]$

for all $\omega \in Mod_L$ and all numerals $n_1, \ldots, n_k \in L_0$. Here, '$A[n_1/x_1, \ldots, n_k/x_k]$' results from '$A[x_1, \ldots, x_k]$' by uniformously substituting 'n_i' for 'x_i' in 'A', $1 \leq i \leq k$. '$A[x_1, \ldots, x_k]$' indicates that 'x_1', …, 'x_k' are the only variables occurring free in 'A'.

A function $\Pr : \mathcal{L} \to \Re_{\geq 0}$ is a probability on L iff for all $A, B \in \mathcal{L}$:

1. $\models A \leftrightarrow B \quad \Rightarrow \quad \Pr(A) = \Pr(B)$
2. $\models A \quad \Rightarrow \quad \Pr(A) = 1$
3. $\models \neg(A \wedge B) \quad \Rightarrow \quad \Pr(A \vee B) = \Pr(A) + \Pr(B)$
4. $\Pr(\exists x A[x]) = \sup\{\Pr(A[n_1/x] \vee \ldots \vee A[n_k/x]) : n_1, \ldots, n_k \in N, k = 1, 2, \ldots\}$

The conditional probability of A given B, $\Pr(A \mid B)$, is defined as $\Pr(A \wedge B) / \Pr(B)$, provided $\Pr(B) > 0$.

A set of sentences S *separates* a set of models $X \subseteq Mod_L$ just in case for any two distinct $\omega_1, \omega_2 \in X$ there is a sentence $A \in S$ such that $\omega_1 \models A$ and $\omega_2 \not\models A$. Note that the set of all atomic sentences separates Mod_L.

Gaifman and Snir (p. 507) prove the following theorem.

Theorem 1 (Gaifman and Snir) *Let $S = \{A_i : i = 0, 1, \ldots\}$ separate Mod_L, and let $[B](\omega)$ be 1 if $\omega \models B$ and 0 otherwise. Then for every $B \in \mathcal{L}$:*

$$\Pr\left(B \mid \bigwedge_{0 \leq i < n} A_i^{\omega}\right) \to [B](\omega) \text{ almost everywhere as } n \to \infty.$$

Based on the Gaifman and Snir convergence theorem we can now prove

Theorem 2 *Let $e_0, \ldots, e_n, \ldots$ be a sequence of sentences of L which separates Mod_L, and let e_i^{ω} be e_i, if $\omega \models e_i$, and $\neg e_i$ otherwise, where $\omega \in Mod_L$. Let $\Pr$ be a strict probability on L, and let $a = a_{\Pr}$ be a function which satisfies 0.b Loveliness and Likeliness, 1.a Continuity fr Loveliness, and 2. Demarcation β for $i = \Pr(\neg H \mid \neg E \wedge B)$ and $p = \Pr(H \mid E \wedge B)$ with demarcation β. Finally, let $\Pr^*$ be the unique probability measure on the smallest σ-field $\mathcal{A}$ containing the field $\{Mod(A) : A \in \mathcal{L}\}$, where $Mod(A) = \{\omega \in Mod_L : \omega \models A\}$.*

Then there exists $X \subseteq Mod_L$ with $\Pr^*(X) = 1$ such that the following hold for every $\omega \in X$, any two contingent $H_1, H_2 \in \mathcal{L}$, and every $H \in \mathcal{L}$:

1. $\omega \models H_1, \omega \not\models H_2 \quad \Rightarrow \quad \exists n \forall m \geq n : a(H_1, E_m^{\omega}) > \beta > a(H_2, E_m^{\omega})$

2. $\left[\begin{array}{c}\omega\models H_1,\omega\models H_2\\ H_1\vdash H_2,H_2\nvdash H_1\end{array}\right]\Rightarrow\exists n\forall m\geq n:a\left(H_1,E_m^\omega\right)>a\left(H_2,E_m^\omega\right)>\beta$

3. $\left[\begin{array}{c}\omega\not\models H_1,\omega\not\models H_2\\ H_1\vdash H_2,H_2\nvdash H_1\end{array}\right]\Rightarrow\exists n\forall m\geq n:\beta>a\left(H_1,E_m^\omega\right)>a\left(H_2,E_m^\omega\right)$

4. $\models H$ *or* $\models\neg H \quad\Rightarrow\quad \forall m:a\left(H,E_m^\omega\right)=\beta$,

where $E_m^\omega=\bigwedge_{0\leq i\leq m}e_i^\omega$.

PROOF:

1. Assume the conditions stated in theorem 2, and suppose $\omega\models H_1$ and $\omega\not\models H_2$, where $\omega\in X'$ for some $X'\subseteq Mod_L$ with $\Pr^*\left(X'\right)=1$ such that

$$\Pr\left(B\mid E_n^{\omega'}\right)\to[B]\left(\omega'\right)\quad\text{as}\quad n\to\infty$$

for all $B\in\mathcal{L}$ and all $\omega'\in X'$ – such X' exists by the Gaifman and Snir theorem. So

$$\Pr\left(H_1\mid E_n^\omega\right)\to1\quad\text{as}\quad n\to\infty,\quad\text{and}\quad\Pr\left(H_2\mid E_n^\omega\right)\to0\quad\text{as}\quad n\to\infty.$$

First, observe that there exists n_1 such that for all $m\geq n_1$:

$$\Pr\left(\neg H_1\mid\neg E_m^\omega\right)>\Pr\left(\neg H_1\right)>0\quad\text{and}\quad\Pr\left(\neg H_2\mid\neg E_m^\omega\right)<\Pr\left(\neg H_2\right)<1.$$

The reason is that $\Pr$ is strict, the H_i are contingent, and (for $i=1,2$)

$$\Pr\left(\neg H_i\mid\neg E_m^\omega\right)\begin{array}{c}>\\<\end{array}\Pr\left(\neg H_i\right)\quad\Leftrightarrow\quad\Pr\left(H_i\mid E_m^\omega\right)\begin{array}{c}>\\<\end{array}\Pr\left(H_i\right),$$

provided $0<\Pr\left(E_m^\omega\right)<1$.

If $\Pr\left(E_m^\omega\right)=0$, then $\Pr^*\left(Mod\left(E_m^\omega\right)\right)=0$ (Gaifman and Snir 1982, p. 504, Basic Fact 1.3). The union of all such sets $Mod\left(E_m^\omega\right)$ of probability 0 is also of probability 0 (because there are just countably many such sets), i.e.

$$\Pr^*(A)=0,\quad A:=\bigcup\left\{Mod\left(E_m^\omega\right)\in\mathcal{A}:\Pr^*\left(Mod\left(E_m^\omega\right)\right)=0\right\}.$$

Similarly, if $\Pr\left(E_m^\omega\right)=1$, then $\Pr^*\left(Mod\left(\neg E_m^\omega\right)\right)=0$. The union of all such sets $Mod\left(\neg E_m^\omega\right)$ of probability 0 is also of probability 0 (again, there are but countably many), i.e.

$$\Pr^*(B)=0,\quad B:=\bigcup\left\{Mod\left(\neg E_m^\omega\right)\in\mathcal{A}:\Pr^*\left(Mod\left(E_m^\omega\right)\right)=1\right\}.$$

As a consequence,

$$\Pr{}^*(X) = 1, \quad X := X' \setminus (A \cup B).$$

(Note that the strictness of $\Pr$ has not been used.) Assume therefore that $\omega \in X$. As $\Pr(H_1 \mid E_n^\omega) \to_n 1$ and $\Pr(H_2 \mid E_n^\omega) \to_n 0$, there is n_1 such that for all $m \geq n_1$:

$$\Pr(H_1 \mid E_m^\omega) > \Pr(H_1) \quad \text{and} \quad \Pr(H_2 \mid E_m^\omega) < \Pr(H_2),$$

and thus

$$\Pr(\neg H_1 \mid \neg E_m^\omega) > \Pr(\neg H_1) > 0 \quad \text{and} \quad \Pr(\neg H_2 \mid \neg E_m^\omega) < \Pr(\neg H_2) < 1.$$

Hence

$$\Pr(\neg H_1) \leq \inf_{m \geq n_1} \{\Pr(\neg H_1 \mid \neg E_m^\omega)\}, \quad \Pr(\neg H_2) \geq \sup_{m \geq n_1} \{\Pr(\neg H_2 \mid \neg E_m^\omega)\}.$$

As

$$\Pr(H_1 \mid E_n^\omega) \to 1 \quad \text{as} \quad n \to \infty, \quad \Pr(H_2 \mid E_n^\omega) \to 0 \quad \text{as} \quad n \to \infty,$$

and as $\Pr(\neg H_1), \Pr(\neg H_2) \in (0,1)$, it follows from Limiting Demarcation β that there exists n_2 such that for all $m \geq n_2$:

$$a(\Pr(H_1 \mid E_m^\omega), \Pr(\neg H_1)) > \beta > a(\Pr(H_2 \mid E_m^\omega), \Pr(\neg H_2)).$$

As a is an increasing function of $\Pr(\neg H_i \mid \neg E_m^\omega)$, it holds for all $m \geq n := \max\{n_1, n_2\}$:

$$a(\Pr(H_1 \mid E_m^\omega), \Pr(\neg H_1 \mid \neg E_m^\omega)) > \beta > a(\Pr(H_2 \mid E_m^\omega), \Pr(\neg H_2 \mid \neg E_m^\omega)).$$

2. Suppose now that $\omega \models H_1$, $\omega \models H_2$, $H_1 \vdash H_2$, and $H_2 \not\vdash H_1$, where $\omega \in X$ for some $X \subseteq Mod_L$ as before. So

$$\Pr(H_1 \mid E_n^\omega) \to 1 \quad \text{as} \quad n \to \infty, \quad \text{and} \quad \Pr(H_2 \mid E_n^\omega) \to 1 \quad \text{as} \quad n \to \infty,$$

and we can safely assume that $0 < \Pr(E_m^\omega) < 1$ for all m. As before, there exists n_1 such that for all $m \geq n_1$:

$$\Pr(\neg H_2 \mid \neg E_m^\omega) > \Pr(\neg H_2) > 0.$$

Observe that

$$\begin{aligned}
\Pr(\neg H_1 \mid \neg E_m^\omega) - \Pr(\neg H_2 \mid \neg E_m^\omega) &= \frac{1 - \Pr(H_1) - \Pr(E_m^\omega) + \Pr(H_1 \wedge E_m^\omega)}{\Pr(\neg E_m^\omega)} - \\
&\quad - \frac{1 - \Pr(H_2) - \Pr(E_m^\omega) + \Pr(H_2 \wedge E_m^\omega)}{\Pr(\neg E_m^\omega)} \\
&= \frac{\Pr(H_2) - \Pr(H_1)}{\Pr(\neg E_m^\omega)} - \\
&\quad - \frac{[\Pr(H_2 \mid E_m^\omega) - \Pr(H_1 \mid E_m^\omega)] \cdot \Pr(E_m^\omega)}{\Pr(\neg E_m^\omega)} \\
&> \frac{\Pr(H_2) - \Pr(H_1)}{\Pr(\neg E_m^\omega)} - \\
&\quad - \frac{\Pr(H_2 \mid E_m^\omega) - \Pr(H_1 \mid E_m^\omega)}{\Pr(\neg E_m^\omega)}.
\end{aligned}$$

By the above, for any $\varepsilon > 0$ there exists n_ε such that for all $m \geq n_\varepsilon$:

$$\Pr(H_2 \mid E_m^\omega) - \Pr(H_1 \mid E_m^\omega) < \varepsilon.$$

Let $\varepsilon^* = \frac{\Pr(H_2) - \Pr(H_1)}{2}$. Then there exists n_{ε^*} such that for all $m \geq n_{\varepsilon^*}$:

$$\Pr(H_2 \mid E_m^\omega) - \Pr(H_1 \mid E_m^\omega) < \varepsilon^*.$$

Consequently it holds for all $m \geq n_{\varepsilon^*}$:

$$\begin{aligned}
\Pr(\neg H_1 \mid \neg E_m^\omega) - \Pr(\neg H_2 \mid \neg E_m^\omega) &> \frac{\Pr(H_2) - \Pr(H_1)}{\Pr(\neg E_m^\omega)} - \\
&\quad - \frac{\Pr(H_2 \mid E_m^\omega) - \Pr(H_1 \mid E_m^\omega)}{\Pr(\neg E_m^\omega)} \\
&> \frac{2\varepsilon^* - \varepsilon^*}{\Pr(\neg E_m^\omega)} > \varepsilon^*.
\end{aligned}$$

Put differently, for all $m \geq n_{\varepsilon^*}$:

$$\Pr(\neg H_1 \mid \neg E_m^\omega) > \Pr(\neg H_2 \mid \neg E_m^\omega) + \varepsilon^*.$$

By Continuity[26], there is $\delta_{\varepsilon^*} > 0$ such that for all $m \geq n_{\varepsilon^*}$:

$$\Pr(H_1 \mid E_m^\omega) > \Pr(H_2 \mid E_m^\omega) - \delta_{\varepsilon^*} \quad \Rightarrow \quad a(H_1, E_m^\omega) > a(H_2, E_m^\omega).$$

[26]It his here where the assumption enters that δ depends only on ε, which is the lower bound

By the above, for $\delta_{\varepsilon^*} > 0$ there exists $n_{\delta_{\varepsilon^*}}$ such that for all $m \geq n_{\delta_{\varepsilon^*}}$:

$\Pr(H_2 \mid E_m^\omega) - \Pr(H_1 \mid E_m^\omega) < \delta_{\varepsilon^*}$, i.e. $\Pr(H_1 \mid E_m^\omega) > \Pr(H_2 \mid E_m^\omega) - \delta_{\varepsilon^*}$.

So for all $m \geq n_2 := \max\{n_{\varepsilon^*}, n_{\delta_{\varepsilon^*}}\}$:

$$a(H_1, E_m^\omega) > a(H_2, E_m^\omega).$$

By Limiting Demarcation β, and because

$$\Pr(H_2 \mid E_n^\omega) \to_n 1 \quad \text{and} \quad \Pr(\neg H_2 \mid \neg E_m^\omega) > \Pr(H_2) > 0 \quad \forall m \geq n_1,$$

it follows that for all $m \geq n := \max\{n_1, n_2\}$:

$$a(H_1, E_m^\omega) > a(H_2, E_m^\omega) > \alpha.$$

3. Similarly.

4. This follows from Demarcation β. □

Theorem 3 *The same holds true if* $i = \Pr(\neg T \mid \neg E \wedge B)$ *is replaced by* $i^* = \Pr(\neg T \mid B)$*, and hence also if* i *is replaced by* $i^*_\alpha = \alpha \cdot i + (1-\alpha) \cdot i' = \alpha \cdot \Pr(\neg T \mid \neg E \wedge B) + (1-\alpha) \cdot \Pr(\neg T \mid B)$*, for any* $\alpha \in [0,1]$.

Thus it turns out that Bayesianism is a paradigm rich enough to provide more than one pair of functions s, t whose continuous and demarcating weighing reveals the true assessment structure in almost every world when presented with separating data.

---

of the difference between $s_{1_m} = \Pr(\neg H_1 \mid \neg E_m^\omega)$ and $s_{2_m} = \Pr(\neg H_2 \mid \neg E_m^\omega)$ (cf. fn. 9). Otherwise, i.e. when δ depends on s_{1_m} and s_{2_m}, it is possible that for every $m \geq n_{\varepsilon^*}$: $\delta_{s_{1_m}, s_{2_m}} = 1/(m+1)$. In this case

$$\inf_{m \geq n_{\varepsilon^*}} \left\{\delta_{s_{1_m}, s_{2_m}}\right\} = 0,$$

whence there need not be $\delta > 0$ and n_δ such that for all $m \geq n_\delta$:

$$\Pr(H_1 \mid E_m^\omega) > \Pr(H_2 \mid E_m^\omega) - \delta.$$

Therefore, in this case it is not possible to prove that there is n such that for all $m \geq n$:

$$a(H_1, E_m^\omega) > a(H_2, E_m^\omega).$$

The preceding also shows that the present approach is viable even if informativeness and truth are not the only virtues a theory should possess. Whatever these virtues besides truth are, and however they are measured, and even if some or all of these virtues depend on the data (cf. Levi 1961 and 1963), as long as there is a function u such that $u(T, E, B)$ measures the overall value – *without* truth – of T in view of E and B, and as long as for any two theories T_1 and T_2, any separating data sequence $e_0, \ldots, e_n, \ldots$ from any world ω, and any body of background information B true in ω, there is a point j such that for all later points $k \geq j$ in the inquiry: $u(T_1, E_k, B) > u(T_2, E_k, B) + \varepsilon$, for some $\varepsilon > 0$, then the following holds for every f satisfying Continuity and Demarcation β for u (instead of s) and t: There is a point m such that for all later points $n \geq m$: $f(u_{1,m}, t_{1,m}) = f(T_1, E_m, B) > f(u_{2,m}, t_{2,m}) = f(T_2, E_m, B)$, where both of these values are greater than β, if both T_1 and T_2 are true in ω, both of these values are smaller than β if both T_1 and T_2 are false in ω, and β lies between these two values if T_1 is true, but T_2 is false in ω.

8 Relevance Measures and Their Exclusive Focus on Truth

As shown in the preceding section, all one needs to do to reveal the true assessment structure in almost every world when presented separating data is to stick to a function satisfying Loveliness and Likeliness, Continuity for Loveliness, and Demarcation β for i^*_α and p, for some $\alpha \in [0, 1]$. What about the central notion in Bayesian confirmation theory – that of a β-relevance measure? So far, it has not really entered the scene.

The connection to the i, p-function $s_c = i + p + c$ for $c = -1$, and the function d_f for $f = \Pr(\neg E \mid B)$ respectively $f = \Pr(\neg E \mid B) \cdot \Pr(B) \cdot \Pr(E \wedge B)$ has already been pointed out. So for any strict probability $\Pr$, $s_{\Pr}$ and $c_{\Pr}$ and $d_{\Pr}$ reveal the true assessment structure in almost every world when presented with separating data. But there are many other relevance measures, and one would like to know whether they all reach the goal.

If H_1 is true in ω, and H_2 is false in ω, where H_1 and H_2 are contingent, then, after finitely many steps, H_1 has to get a greater value in ω than the demarcation line β which in turn has to be greater than the value of H_2 in ω. Any β-relevance measure r reveals this part of almost any ω's assessment structure. By the Gaifman and Snir convergence theorem,

$$\Pr(H_1 \mid E_n^\omega) \to_n 1 \quad \text{and} \quad \Pr(H_2 \mid E_n^\omega) \to_n 0,$$

whence there exists n such that for all $m \geq n$:

$$\Pr(H_1 \mid E_m^\omega) > \Pr(H_1) \quad \text{and} \quad \Pr(H_2 \mid E_m^\omega) < \Pr(H_2),$$

provided $\Pr$ is strict. Thus, by the definition of a β-relevance measure, it holds for all $m \geq n$:

$$r(H_1, E_m^\omega) > \beta > r(H_2, E_m^\omega).$$

Moreover, the value (in ω) of any contingent hypothesis that is true/false in ω will eventually stay above/below β; the value of any logically determined hypothesis is always equal to β.

So far, so good. But the definition of a β-relevance measure by itself does not imply anything about the relative positions of two hypotheses, if they have the same truth value in some world ω (except that they are both above/below β). This exclusive focus on truth – in contrast to the weighing between the conflicting goals of informativeness and truth of an s, t-function – is what prevents relevance measures from revealing the true assessment structure in general.

As we have seen, β-relevance measures sometimes do weigh between i_α^* and p. Yet, the point is that β-relevance measures are not required to take into account both aspects. They are allowed but not required to weigh between informativeness and truth. In concluding, let us briefly look at the most popular relevance measures (cf. Fitelson 2001) all of which are 0-relevance measures. It is assumed throughout that $\Pr$ is strict.

As already mentioned, the Joyce-Christensen measure s, the distance measure d, and the Carnap measure c get it right in all four cases (in case of Carnap's c, note that the union of all sets $Mod(\pm E_n^\omega)$ with $\Pr(\pm E_n^\omega) = 0$ has probability 0 in the sense of $\Pr^*$, whence $f = \Pr(\neg E_n^\omega \mid B) \cdot \Pr(B) \cdot \Pr(E_n^\omega \wedge B)$ is 0 only for a set of measure 0).

The ratio measure r,

$$r_{\Pr}(H, E, B) = \log\left[\frac{\Pr(H \mid E \wedge B)}{\Pr(H \mid B)}\right],$$

gets it right in case both H_1 and H_2 are true in ω, and $H_1 \vdash H_2$, but $H_2 \nvdash H_1$. In this case

$$r_{\Pr}(H_1, E_n^\omega) \to_n \log[1/\Pr(H_1)] \quad \text{and} \quad r_{\Pr}(H_1, E_n^\omega) \to_n \log[1/\Pr(H_2)],$$

whence there exists n such that for all $m \geq n$:

$$r_{\mathrm{Pr}}\left(H_1, E_m^{\omega}\right) > r_{\mathrm{Pr}}\left(H_2, E_m^{\omega}\right) > 0.$$

However, r does not get it right when both H_1 and H_2 are contingent and false in ω, and such that $H_1 \vdash H_2$, but $H_2 \nvdash H_1$. Here,

$$r_{\mathrm{Pr}}\left(H_1, E_n^{\omega}\right) \to_n \log 0 = -\infty \quad \text{and} \quad r_{\mathrm{Pr}}\left(H_2, E_n^{\omega}\right) \to_n \log 0 = -\infty,$$

and although

$$\frac{1}{\Pr\left(H_1\right)} > \frac{1}{\Pr\left(H_2\right)},$$

it also holds that

$$\Pr\left(H_1 \mid E_m^{\omega}\right) < \Pr\left(H_2 \mid E_m^{\omega}\right),$$

whence there need not be n such that for all $m \geq n$:

$$\log\left[\frac{\Pr\left(H_1 \mid E_n^{\omega}\right)}{\Pr\left(H_1\right)}\right] > \log\left[\frac{\Pr\left(H_2 \mid E_n^{\omega}\right)}{\Pr\left(H_2\right)}\right].$$

For logically determined H, r takes on the value 0, if it is stipulated that $0/0 = 1$.

The situation is even worse for the log-likelihood ratio l,

$$l_{\mathrm{Pr}}\left(H, E, B\right) = \log\left[\frac{\Pr\left(E \mid H \wedge B\right)}{\Pr\left(E \mid \neg H \wedge B\right)}\right].$$

When H_1 and H_2 are contingently true in ω and such that $H_1 \vdash H_2$, but $H_2 \nvdash H_1$, and thus

$$\Pr\left(H_1 \mid E_n^{\omega}\right) \to_n 1, \quad \text{and} \quad \Pr\left(H_1 \mid E_n^{\omega}\right) \to_n 1,$$

it need not be the case that

$$l_{\mathrm{Pr}}\left(H_1, E_n^{\omega}\right) = \log\left[\frac{\Pr\left(\neg H_1 \mid E_n^{\omega}\right) \cdot \Pr\left(\neg H_1\right)}{\Pr\left(\neg H_1 \mid E_n^{\omega}\right) \cdot \Pr\left(H_1\right)}\right]$$

stays strictly above

$$l_{\mathrm{Pr}}\left(H_2, E_n^{\omega}\right) = \log\left[\frac{\Pr\left(\neg H_2 \mid E_n^{\omega}\right) \cdot \Pr\left(\neg H_2\right)}{\Pr\left(\neg H_2 \mid E_n^{\omega}\right) \cdot \Pr\left(H_2\right)}\right].$$

For although

$$\frac{\Pr\left(\neg H_1\right)}{\Pr\left(H_1\right)} > \frac{\Pr\left(\neg H_2\right)}{\Pr\left(H_2\right)},$$

it also holds that
$$\frac{\Pr(H_1 \mid E_n^\omega)}{\Pr(\neg H_1 \mid E_n^\omega)} < \frac{\Pr(H_2 \mid E_n^\omega)}{\Pr(\neg H_2 \mid E_n^\omega)}.$$
In any case,
$$l_{\Pr}(H_1, E_n^\omega) \to_n \infty \quad \text{and} \quad l_{\Pr}(H_2, E_n^\omega) \to_n \infty.$$
The situation is similar when both H_1 and H_2 are contingently false in ω, where
$$l_{\Pr}(H_1, E_n^\omega) \to_n -\infty \quad \text{and} \quad l_{\Pr}(H_2, E_n^\omega) \to_n -\infty.$$
If H is logically determined, one really has to make an effort for l to get it right, for one needs to stipulate that $0 \cdot 1/1 \cdot 0 = 1 \cdot 0/0 \cdot 1 = 1$.

It is interesting to see that the log-likelihood ratio l seems to come out on top when subjectively plausible desiderata are at issue (Fitelson 2001), but to do much more poorly when it comes to the matter-of-fact question whether an assessment function (or measure of confirmation) furthers the goal it is supposed to further – whether it gets it right in the sense that it reveals the true assessment (or confirmational) structure and thus leads to true and informative theories.

Due to their focus on truth, relevance measures – like s, t-functions – separate true from false theories, but due to the exclusiveness of this focus, they do not – in contrast to s, t-functions – distinguish between informative and uninformative true or false theories.

9 Hempel's Triviality Result

In his "Studies in the Logic of Confirmation" (1945) Carl G. Hempel presented the following conditions of adequacy for any relation of confirmation $|\!\sim\ \subseteq \mathcal{L} \times \mathcal{L}$:

1. Entailment Condition: $\quad E \vdash H \quad \Rightarrow \quad E \mid\!\sim H$
2. Consequence Condition: $\quad \{H : E \mid\!\sim H\} \vdash H' \quad \Rightarrow \quad E \mid\!\sim H'$
   - 2.1 Special Consequence Cond.: $\quad E \mid\!\sim H, \quad H \vdash H' \quad \Rightarrow \quad E \mid\!\sim H'$
   - 2.2 Equivalence Condition: $\quad E \mid\!\sim H, \quad H \dashv\vdash H' \quad \Rightarrow \quad E \mid\!\sim H'$
3. Consistency Condition: $\quad \{E\} \cup \{H : E \mid\!\sim H\} \nvdash \bot$
   - 3.1 $\quad E \nvdash \bot, \quad E \mid\!\sim H \quad \Rightarrow \quad E \nvdash \neg H$

3.2 $E \nvdash \bot, \quad E \mid\!\sim H, \quad H \vdash \neg H' \quad \Rightarrow \quad E \not\mid\!\sim H'$

4. Converse Consequence Condition: $E \mid\!\sim H, \quad H' \vdash H \quad \Rightarrow \quad E \mid\!\sim H'$

Hempel then showed that 1.-4. entail that every sentence (observation report) E confirms every sentence (theory) H, i.e. for all $E, H \in \mathcal{L}$: $E \mid\!\sim H$.

Since Hempel's negative result, there has hardly been any progress in constructing a logic of confirmation.[27] One reason seems to be that up to now the predominant view on Hempel's conditions is the analysis Carnap gave in his *Logical Foundations of Probability* (1962), § 87.

9.1 Carnap's Analyis of Hempel's Conditions

In analyzing the consequence condition, Carnap argues that

> ... Hempel has in mind as explicandum the following relation: 'the degree of confirmation of T by E is greater than r', where r is a fixed value, perhaps 0 or $1/2$. (Carnap 1962, p. 475; notation adapted)

In discussing the consistency condition, Carnap mentions that

> Hempel himself shows that a set of physical measurements may confirm several quantitative hypotheses which are incompatible with each other (p. 106). This seems to me a clear refutation of [3.2]. [...] What may be the reasons that have led Hempel to the consistency conditions [3.2] and [3]? He regards it as a great advantage of any explicatum satisfying [3] "that is sets a limit, so to speak, to the strength of the hypotheses which can be confirmed by given evidence" [...] This argument does not seem to have any plausibility for *our* explicandum, (Carnap 1962, pp. 476-7; emphasis in the original)

which is the concept of positive statistical relevance, or "initially confirming evidence", as Carnap says (Carnap 1962, §86), which holds for contingent E and strict $\Pr$ if and only if $\Pr(T \mid E) > \Pr(T)$;

[27]The only two articles I know of are Zwirn/Zwirn (1996) and Milne (2000). Roughly, Zwirn/Zwirn (1996) argue that there is no logic of confirmation (taking into account all of the partly conflicting aspects of confirmation), whereas Milne (2000) argues that there is a logic of confirmation, but that it does not deserve to be called a logic.

> [b]ut it is plausible for the second explicandum mentioned earlier: the degree of confirmation exceeding a fixed value r. Therefore we may perhaps assume that Hempel's acceptance of the consistency condition is due again to an inadvertant shift to the second explicandum. (Carnap 1962, pp. 477-8.)

We can summarize Carnap's analysis as follows: In presenting his first three conditions of adequacy Hempel was mixing up two distinct concepts of confirmation, two distinct explicanda in Carnap's terminology, viz. the concept of incremental confirmation or of positive statistical relevance or of initially confirming evidence, and the concept of absolute confirmation or the likeliness concept of confirmation. Hempel's second and third conditions hold true for the second explicandum, i.e. the likeliness concept of confirmation, but they do not hold true for the first explicandum, i.e. the concept of positive statistical relevance. On the other hand, Hempel's first condition holds true for the first explicandum, but it does so only in a *qualified* form (cf. Carnap 1962, p. 473).

This is indeed the view Carnap seems to hold in § 87 of his (1962). This, however, means that Hempel first had in mind the explicandum of positive statistical relevance for the entailment condition; then he had in mind the second explicandum for the consequence and the consistency conditions; and then, when Hempel presented the converse consequence condition, he got completely confused, so to speak, and had in mind still another explicandum or concept of confirmation. Apart from not being very charitable, Carnap's reading of Hempel also leaves open the question what the third explicandum might have been.

9.2 Hempel Vindicated

Let us turn to our own analysis of Hempel's condition of adequacy by noting first of all that the second explicandum referred to by Carnap also satisfies the entailment condition *without* qualification: If E logically implies T, then $\Pr(T \mid E) = 1 \geq r$, for any value $r \in [0, 1]$.

So the following more charitable reading of Hempel seems plausible: When presenting his first three conditions, Hempel had in mind Carnap's second explicandum, the concept of absolute confirmation or the likeliness concept of confirmation – which, to some extent, is still in accordance with Carnap's analysis. But then, when discussing the converse consequence condition, Hempel also felt the need for a second concept of confirmation: the loveliness concept of confirmation aiming at informative theories.

Given that it was the converse consequence condition which Hempel gave up in his *Studies*, Carnap's analysis cannot explain which concepts Hempel had in mind when he put together conditions of adequacy that lead to the triviality result that every observation report confirms every theory.

The present analysis, however, makes perfect sense of Hempel's argumentation: Though he felt the need for the second concept of confirmation, Hempel also realized that these two concepts were *conflicting*, and so he abandoned the loveliness concept in favour of the likeliness concept.

Let us check this by going through Hempel's conditions:

Hempel's entailment condition is that E confirms T if E logically implies T. As already mentioned, this holds true of the likeliness concept of confirmation without qualification, but it does not hold of the loveliness concept of confirmation. If evidence E logically implies theory T, then – by condition 2a in the definition of indicating truth – the likeliness of T given E is maximal. On the other hand, if E entails T, then T does in general not inform about E.

The consequence condition says that if E confirms a set of theories $\mathcal{T}$, then E confirms every consequence T of $\mathcal{T}$. This condition clearly does not hold of the loveliness concept. It holds without qualification for the likeliness concept only in its restricted form [2.1], the special consequence condition.[28] The latter condition amounts to a qualitative version of condition 3 in the definition of indicating truth, and expresses that likeliness decreases with the logical strength of the theory to be assessed.

However, if E confirms T *because of* the consequence condition, i.e. because T is a logical consequence of the set of all theories confirmed by E, then, by that very condition, E confirms every logical consequence T' of T. On the other hand, E need not confirm any T^* logically implying some or all of the theories confirmed by E.

The third condition of consistency says that every consistent E is compatible with the set $\mathcal{T}$ of all theories T it confirms. As before, this condition does not hold for the loveliness concept of confirmation. It holds for the likeliness concept only with a proviso and only in its restricted form [3.1]: If E is consistent and confirms T, then E confirms no T' which is not consistent with T. In this case, T' logically implies $\neg T$, whence by condition 3 in the definition of indicating truth, the likeliness of T' is not greater than that of $\neg T$. Given the proviso that no consistent E confirms both a theory and its negation, the result follows.[29]

[28] This is also mentioned by Carnap – cf. Carnap (1962), pp. 474-6.

[29] As noted by Carnap on p. 478 of his (1962), the proviso is satisfed if likeliness is measured

However, as before, we have that if E is consistent with the set $\mathcal{T}$ of all theories it confirms, and T' is a logical consequence of any $T \in \mathcal{T}$ or even $\mathcal{T}$ itself, then E is also consistent with $\mathcal{T} \cup \{T'\}$, and so satisfies the consistency condition. On the other hand, E need not be consistent with a theory T^* logically implying some or even all of the theories E confirms.

In particular, if the probability of theory T given evidence E is high ($> 1/2$), then the probability of any theory T' given E – T' being incompatible with T – must be low ($< 1/2$), because the sum of these two probabilities cannot exceed 1.

The culprit, according to Hempel (cf. pp. 103-107, esp. pp. 104-105 of his *Studies*), is the converse consequence condition: If E confirms T and T is logically implied by T^*, then E confirms T^*. Clearly, this is an instance of the loveliness concept of confirmation, but not of the likeliness concept of confirmation. It amounts to a qualitative version of condition 3 in the definitions of evidence-based and evidence-neglecting strength indicators. Furthermore, it coincides with the defining clause of informativeness or loveliness relations by expressing the requirement that loveliness increases with the logical strength of the theory to be assessed.

10 The Logic of Theory Assessment

However, in a sense one can have the cake and eat it: There is a logic of confirmation – or rather: theory assessment – that takes into account both of these two conflicting concepts.

Roughly speaking, HD says that a good theory is informative, whereas IL says that a good theory is true (probable). The driving force behind Hempel's conditions seems to be the insight that *a good theory is both true and informative*. Hence, in assessing a given theory by the available data, one should account for these two conflicting aspects. This is done in the following.

Let $\langle W, \mathcal{A}, \kappa \rangle$ be a ranking space, where W is a non-empty set of possibilities, $\mathcal{A}$ is a field over W, i.e. a set of subsets of W containing $\emptyset$ and closed under complementation and finite intersections, and $\kappa : W \to \mathbf{N} \cup \{\infty\}$ is a ranking function (cf. Spohn 1988, 1990), i.e. a function from W into the set of extended natural numbers $\mathbf{N} \cup \{\infty\}$ such that at least one possibility $\omega \in W$ is assigned rank 0. κ is extended to a function on the field $\mathcal{A}$ by setting $\kappa(\emptyset) = \infty$ and by

by a probability $\Pr$, and $r \geq 1/2$ respectively $r > 1/2$, depending on whether one requires $\Pr(T \mid E) > r$ or $\Pr(T \mid E) \geq r$ in order for E to confirm T in the sense of absolute confirmation.

defining, for each non-empty $A \in \mathcal{A}$,

$$\kappa(A) = \min\{\kappa(\omega) : \omega \in A\}.$$

The conditional rank of $B \in \mathcal{A}$ given $A \in \mathcal{A}$, $\kappa(B \mid A)$, is defined as

$$\kappa(B \mid A) = \begin{cases} \kappa(A \cap B) - \kappa(B) & \text{if} \quad \kappa(A) < \infty \\ 0 & \text{if} \quad \kappa(A) = \infty \end{cases}$$

(Goldszmidt/Pearl 1996, p. 63, define $\kappa(B \mid A) = \infty$, if $\kappa(A) = \infty$.) A ranking function represents an ordering of *dis*belief. So the degree of belief in H given E is given by the degree of disbelief in $\overline{H}$ given E.

For any ranking space $\mathbf{A} = \langle W, \mathcal{A}, \kappa \rangle$ we have

Observation 1

1. $0 = \kappa(W) \leq \kappa(A) \leq \kappa(\emptyset) = \infty$ *for each* $A \in \mathcal{A}$
2. $\kappa(A) = 0$ *and/or* $\kappa(\overline{A}) = 0$ *for each* $A \in \mathcal{A}$
3. $A \subseteq B \quad \Rightarrow \quad \kappa(B) \leq \kappa(A)$*, for any* $A, B \in \mathcal{A}$
4. $\kappa(A \cap B) = \kappa(A) + \kappa(B \mid A)$
5. $\kappa(A \cup B) = \min\{\kappa(A), \kappa(B)\}$ *for all* $A, B \in \mathcal{A}$
6. $\kappa(\bigcup\{A_i : i \in \boldsymbol{N}\}) = \min\{\kappa(A_i) : i \in \boldsymbol{N}\}$ *for each sequence* $(A_i)_{i \in \boldsymbol{N}} \in \mathcal{A}$*, if* $\mathcal{A}$ *is a* σ*-field over* W
7. $\kappa(\bigcup \mathcal{B}) = \min\{\kappa(B) : B \in \mathcal{B}\}$ *for each* $\mathcal{B} \subseteq \mathcal{A}$*, if* $\mathcal{A}$ *is complete*

For $E, H \in \mathcal{A}$, $\kappa(\overline{H} \mid E)$ measures how *likely* H is given E, whereas $\kappa(H \mid \overline{E})$ measures how much H *informs* about E (with the extended natural numbers as range). However, how to measure informativeness and likeliness is not the task of the present section. Here we are interested in the qualitative counterpart of the quantitative notion of degree of confirmation, or rather: the qualitative counterpart of the quantitative assessment value – this being the notion of *an acceptable theory given the data*.

Neglecting the background information B, it is tempting to say that hypothesis H is an acceptable theory for evidence E iff the overall assessment value of H relative to E is greater than that of its complement $\overline{H}$ relative to E. This, however, has the consequence that it depends on the way one combines loveliness and likeliness whether H is an acceptable theory relative to E. One may, for instance, simply take the sum $s + t$, or else one may judge informativeness measured by s more important than likeliness measured by t and stick with $s + t^x$, for some $x > 1$.

The only clear case in which H is acceptable given E is when H is at least as likely given E as is its complement $\overline{H}$ given E, and H informs more about E than does $\overline{H}$; or else, H is more likely given E than is $\overline{H}$ given E, and H informs at least as much about E as does its complement $\overline{H}$. This will indeed be our defining clause below, but before we have to fix a little bit of terminology.

A ranking space $\langle W, \mathcal{A}, \kappa\rangle$ is an *assessment model* for the language $\mathcal{L}$ iff W is the set of all models for $\mathcal{L}$, i.e. $W = Mod_{\mathcal{L}}$, $Mod\,(\alpha) \in \mathcal{A}$ for each $\alpha \in \mathcal{L}$, and $\kappa\,(Mod\,(\alpha)) < \infty$ for each consistent $\alpha \in \mathcal{L}$.

The consequence relation $|\!\sim_\kappa$ over the language $\mathcal{L}$ induced by an assessment model $\langle Mod_{\mathcal{L}}, \mathcal{A}, \kappa\rangle$ is defined as follows:

$$\begin{aligned} \alpha \mathrel{|\!\sim_\kappa} \beta \quad \Leftrightarrow \quad & \kappa\,(Mod\,(\beta) \mid Mod\,(\alpha)) \leq \kappa\,(Mod\,(\neg\beta) \mid Mod\,(\alpha)) \quad \& \\ & \kappa\,(Mod\,(\neg\beta) \mid Mod\,(\neg\alpha)) \leq \kappa\,(Mod\,(\beta) \mid Mod\,(\neg\alpha)), \end{aligned}$$

where at least one of these inequalities is strict. In words: β is an acceptable theory given α (in the sense of κ) iff β is at least as likely as and more informative than its negation $\neg\beta$ given α, or β is at least as informative as and more likely than its negation $\neg\beta$ given α.

On the other (i.e. the syntactical) hand, a relation $|\!\sim \,\subseteq \mathcal{L} \times \mathcal{L}$ is an *assessment* (or *confirmation*) *relation on $\mathcal{L}$* iff $|\!\sim$ satisfies the following principles:

1. $\alpha \mathrel{|\!\sim} \alpha$ — Reflexivity*
2. $\alpha \mathrel{|\!\sim} \beta, \quad \alpha \dashv\vdash \gamma \quad \Rightarrow \quad \gamma \mathrel{|\!\sim} \beta$ — Left Logical Equivalence*
3. $\alpha \mathrel{|\!\sim} \beta, \quad \beta \dashv\vdash \gamma \quad \Rightarrow \quad \alpha \mathrel{|\!\sim} \gamma$ — Right Logical Equivalence*
4. $\alpha \mathrel{|\!\sim} \beta \quad \Rightarrow \quad \alpha \not\mathrel{|\!\sim} \neg\beta$ — Selectivity*
5. $\alpha \mathrel{|\!\sim} \beta \quad \Rightarrow \quad \alpha \mathrel{|\!\sim} \alpha \wedge \beta$ — Weak Composition*
6. $\alpha \mathrel{|\!\sim} \beta \quad \Rightarrow \quad \neg\alpha \mathrel{|\!\sim} \neg\beta$ — Loveliness and Likeliness
7. $\nvdash \alpha \vee \beta \quad \Rightarrow \quad \alpha \vee \beta \mathrel{|\!\sim} \alpha \quad \text{or} \quad \alpha \vee \beta \mathrel{|\!\sim} \beta$ — Either-Or
8. $\alpha \vee \beta \not\mathrel{|\!\sim} \alpha, \quad \nvdash \alpha \vee \beta \quad \Rightarrow \quad \alpha \vee \neg\alpha \mathrel{|\!\sim} \neg\alpha$ — Negation 1

9. $\alpha \vee \neg\alpha \mid\!\sim \alpha, \quad \alpha \vdash \beta \quad \Rightarrow \quad \alpha \vee \neg\alpha \mid\!\sim \beta$ Up

10. $\alpha \wedge \neg\alpha \mid\!\sim \alpha, \quad \alpha \vee \beta \mid\!\sim \alpha \quad \Rightarrow \quad \alpha \wedge \neg\alpha \mid\!\sim \beta$ Down

11. $\alpha \mid\!\sim \alpha \wedge \beta, \quad \alpha \mid\!\sim \alpha \vee \beta \quad \Rightarrow \quad \alpha \not\mid\!\sim \neg\beta$ No Name 1

12. $\alpha \not\mid\!\sim \alpha \wedge \neg\beta, \quad \alpha \mid\!\sim \alpha \vee \beta, \quad \alpha \nvdash \bot, \quad \nvdash \alpha \quad \Rightarrow \quad \alpha \mid\!\sim \beta$ No Name 2

13. $\alpha \vee \beta \mid\!\sim \alpha, \quad \beta \vee \gamma \mid\!\sim \beta, \quad \nvdash \alpha \vee \gamma \quad \Rightarrow \quad \alpha \vee \gamma \mid\!\sim \alpha$ quasi Nr 21

14. $\alpha \vee \beta \mid\!\sim \alpha, \quad \beta \vee \gamma \mid\!\sim \beta, \quad \vdash \alpha \vee \gamma \quad \Rightarrow \quad \alpha \vee \gamma \not\mid\!\sim \neg\alpha$

    supplementary Nr 21

15. $\alpha_i \vee \alpha_{i+1} \mid\!\sim \alpha_{i+1}, \quad \nvdash \alpha_i \vee \alpha_j \quad \Rightarrow \quad \exists n \forall m \geq n : \alpha_m \vee \alpha_{m+1} \mid\!\sim \alpha_m$

    Minimum (no strictly $\prec$-decreasing sequence, where $\alpha \preceq \beta \Leftrightarrow \alpha \vee \beta \mid\!\sim \alpha$ for non-tautological $\alpha \vee \beta$, and $\alpha \prec \neg\alpha \Leftrightarrow \alpha \vee \beta \mid\!\sim \alpha$ for tautol. $\alpha \vee \beta$)

The * starred principles are among the *core principles* in Zwirn/Zwirn (1996). Loveliness and Likeliness says that $\neg\beta$ is at least as likely as and more informative than β given $\neg\alpha$, if β is more likely than and at least as informative as $\neg\beta$ given α (or the other way round). Loveliness and Likeliness is not the same as Negation Symmetry in Milne (2000).

For the remaining principles it is helpful to keep in mind that, for non-tautological $\alpha \vee \beta$, $\alpha \vee \beta \mid\!\sim \alpha$ says that α is not more disbelieved than β, i.e. $\kappa(\alpha) \leq \kappa(\beta)$ (for the ranking function κ constructed from $\mid\!\sim$ in the proof of the second part of theorem 4). For tautological $\alpha \vee \beta$, $\alpha \vee \beta \mid\!\sim \alpha$ says that $\neg\alpha$ is disbelieved, i.e. $\kappa(\alpha) < \kappa(\neg\alpha)$. So Either-Or amounts to the connectedness of the $\leq$-relation between natural numbers: Either the rank of α is not greater than the rank of β, or the rank of β is not greater than that of α. According to Negation 1, if the rank of α is greater than the rank of β, then the rank of α is greater than 0, which, by observation 1.2, holds if and only if the rank of α is greater than the rank of $\neg\alpha$. Up corresponds to observation 1.3: If the rank of $\neg\alpha$ is greater than 0, and if $\neg\beta$ logically implies $\neg\alpha$, then the rank of $\neg\beta$ is also greater than 0. Similarly for Down: If the rank of α is greater than 0 and the rank of β is not smaller than the rank of α, then the rank of β is also greater than 0.

No Name 1 can be put as follows: If $\alpha \wedge \beta$ is at least as likely as its negation given α, and if $\alpha \vee \beta$ informs at least as much about α as does its negation, then β – which is at least as likely as $\alpha \wedge \beta$ and at least as informative as $\alpha \vee \beta$ – is neither less likely nor less informative than its negation given α. As to No

Name 2, the antecedent is a complicated way of saying that β is an acceptable theory for contingent α. quasi Nr 21 is the derived rule (21) of the system **P** in Kraus/Lehmann/Magidor (1990). It amounts to the transitivity of the $\leq$-relation between natural numbers: If the rank of α is not greater than that of β, and if the rank of β is not greater than that of γ, then the rank of α is not greater than that of γ. supplementary Nr 21 combines the transitivity of $\leq$ with observation 1.3: If the rank of α is not greater than that of β, if the rank of β is not greater than that of γ, and if $\neg\alpha$ logically implies γ, then the rank of α is not greater than that of $\neg\alpha$. Finally, Minimum expresses nothing but the fact that there is no strictly $<$-decreasing sequence of natural numbers.

Theorem 4 *The consequence relation $\mid\!\sim_\kappa$ induced by an assessment model $\langle Mod_{\mathcal{L}}, \mathcal{A}, \kappa\rangle$ is an assessment relation on $\mathcal{L}$. Conversely, for each assessment relation $\mid\!\sim$ on $\mathcal{L}$ there is an assessment model $\langle Mod_{\mathcal{L}}, \mathcal{A}, \kappa\rangle$ such that $\mid\!\sim = \mid\!\sim_\kappa$.*

The following principles are admissible (the last follows from Reflexivity):

16.	$\alpha \nvdash \bot \quad \Rightarrow \quad \alpha \not\mid\!\sim \bot$	Consistency*
17.	$\nvdash \alpha \quad \Rightarrow \quad \alpha \not\mid\!\sim \top$	Informativeness
18.	$\alpha \mid\!\sim \beta \rightarrow \gamma, \quad \alpha \mid\!\sim \beta \quad \Rightarrow \quad \alpha \mid\!\sim \gamma$	MPC
19.	$\alpha \mid\!\sim \beta, \quad \alpha \mid\!\sim \gamma \quad \Rightarrow \quad \alpha \mid\!\sim \beta \wedge \gamma$ or $\alpha \mid\!\sim \beta \vee \gamma$	quasi Comp.
20.	$\alpha \wedge \beta \mid\!\sim \gamma, \quad \alpha \wedge \neg\beta \mid\!\sim \gamma \quad \Rightarrow \quad \alpha \mid\!\sim \gamma$	Proof by Cases, D
21.	$\alpha \vee \neg\alpha \mid\!\sim \alpha \quad \Rightarrow \quad \alpha \vee \gamma \mid\!\sim \alpha$	Negation 2
22.	$\alpha \vee \beta \vee \gamma \mid\!\sim \beta \vee \gamma \quad \Rightarrow \quad \alpha \vee \beta \mid\!\sim \beta$ or $\alpha \vee \gamma \mid\!\sim \gamma$	Ranks
23.	$\alpha \vdash \beta \quad \Rightarrow \quad \alpha \vee \beta \mid\!\sim \alpha$	No Name 3

The following principles are not admissible:

i.	$\alpha \vdash \beta \quad \Rightarrow \quad \alpha \mid\!\sim \beta$	Entailment, Supraclassicality*
ii.	$\beta \vdash \alpha \quad \Rightarrow \quad \alpha \mid\!\sim \beta$	Conversion
iii.	$\alpha \mid\!\sim \beta, \quad \gamma \vdash \alpha \quad \Rightarrow \quad \alpha \mid\!\sim \gamma$	Left Monotonicity
iv.	$\alpha \mid\!\sim \beta, \quad \beta \vdash \neg\gamma \quad \Rightarrow \quad \alpha \not\mid\!\sim \gamma$	Strong Selectivity

v. $\alpha \wedge \beta \mid\sim \gamma, \quad \alpha \mid\sim \beta \quad \Rightarrow \quad \alpha \mid\sim \gamma$ Cut

vi. $\alpha \mid\sim \gamma, \quad \alpha \mid\sim \beta \quad \Rightarrow \quad \alpha \wedge \beta \mid\sim \gamma$ Cautious Monotonicity

As to iv., remember the second quotation in section 9.1 from Carnap (1962). In comparing the present approach with standard nonmonotonic logic in the KLM-tradition (cf. Kraus/Lehmann/Magidor 1990), we note two points:

First, the present system is *genuinely nonmonotonic* in the sense that not only Left, but also Right Monotonicity (= Right Weakening) is not admissible:

vii. $\alpha \mid\sim \beta, \quad \beta \vdash \gamma \quad \Rightarrow \quad \alpha \mid\sim \gamma$ Right Monotonicity, Right Weakening

So not only arbitrary strengthening of the premises, but also arbitrary weakening of the conclusion is not allowed. The reason is this: By arbitrary weakening of the conclusion information is lost – and the less informative conclusion need not anymore be worth taking the risk being of led to a false conclusion.

Second, the present approach can explain why everyday reasoning is satisfied with a standard that is weaker than truth-preservation in all possible worlds (e.g. truth-preservation in all normal worlds), and thus runs the risk of being led to a false conclusion: We are willing to take this risk, because we want to arrive at informative conclusions that go beyond the premises.

Finally, one might wonder how the present logic of theory assessment compares to Carnap's analysis of Hempel's conditions and his *dictum* that qualitative confirmation is positive statistical relevance. The following observation gives a first hint:

Observation 2 *For every regular probability* Pr *on some Gaifman-Snir language* $\mathcal{L}$*,*

$$\mid\sim_{\mathrm{Pr}} \quad = \quad \perp^{+}_{\mathrm{Pr}} \cup \{\langle \alpha, \beta \rangle : \alpha \dashv\vdash \beta \dashv\vdash \top\} \cup \{\langle \alpha, \beta \rangle : \alpha \dashv\vdash \beta \dashv\vdash \bot\}$$

is an assessment relation on $\mathcal{L}$*, where* $\perp^{+}_{\mathrm{Pr}}$ *is the relation of positive statistical relevance in the sense of* Pr.

However, confirmation in the combined sense of loveliness and likeliness is not the same as positive statistical relevance. In contrast to positive statistical relevance, Symmetry is not admissible in the present context:

viii. $\alpha \mid\sim \beta \quad \Rightarrow \quad \beta \mid\sim \alpha$ Symmetry

11 Acknowledgements

My research was supported by the Alexander von Humboldt Foundation, the Federal Ministry of Education and Research, and the Program for the Investment in the Future (ZIP) of the German Government through a Sofja Kovalevskaja Award to Luc Bovens.

References

[1] Bar-Hillel, Yehoshua: Semantic Information and Its Measures, in: *Transactions of the Tenth Conference on Cybernetics*, Josiah Macy, Jr. Foundation, New York, 1952, pp. 33-48; reprinted in: Bar-Hillel, Yehoshua: Language and Information. Selected Essays on Their Theory and Application, Reading (Mass.): Addison-Wesley, 1964, pp. 298-310.

[2] Bar-Hillel, Yehoshua: An Examination of Information Theory, in: *Philosophy of Science* 22, 1955, pp. 86-105; reprinted in: Bar-Hillel, Yehoshua: Language and Information. Selected Essays on Their Theory and Application, Reading (Mass.): Addison-Wesley, 1964, pp. 275-297.

[3] Bar-Hillel, Yehoshua/Carnap, Rudolf: Semantic Information, in: *The British Journal for the Philosophy of Science* 4, 1953, pp. 147-157.

[4] Carnap, Rudolf: The Continuum of Inductive Methods, Chicago: University of Chicago Press, 1952.

[5] Carnap, Rudolf: Logical Foundations of Probability, 2^{nd} ed., Chicago: University of Chicago Press, 1962.

[6] Carnap, Rudolf/Bar-Hillel, Yehoshua: An Outline of a Theory of Semantic Information, Technical Report No. 247 of the Research Laboratory of Electronics, Massachusetts Institute of Technology, 1952; reprinted in: Bar-Hillel, Yehoshua: Language and Information. Selected Essays on Their Theory and Application, Reading (Mass.): Addison-Wesley, 1964, pp. 221-274.

[7] Christensen, David: Measuring Confirmation, in: *Journal of Philosophy* 96, 1999, pp. 437-461.

[8] Earman, John: Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory, Cambridge, Mass.: MIT Press, 1992.

[9] Fitelson, Branden: Studies in Bayesian Confirmation Theory, University of Wisconsin-Madison: Doctoral Dissertation, 2001.

[10] Gaifman, Haim/Snir, Marc: Probabilities over Rich Languages, Testing, and Randomness, in: *Journal of Symbolic Logic* 47, 1982, pp. 495-548.

[11] Goldszmidt, Moisés/Pearl, Judea: Qualitative probabilities for default reasoning, belief revision, and causal modeling, in: *Artificial Intelligence* **84**, 1996, pp. 57-112.

[12] Goodman, Nelson: A Query on Confirmation, in: *The Journal of Philosophy* 43, 1946, pp. 383-385.

[13] Goodman, Nelson: On Infirmities of Confirmation-Theory, in: *Philosophy and Phenomenological Research* 8, 1947, pp. l49-151.

[14] Helmer, Olaf/Oppenheim, Paul: A Syntactical Definition of Probability and of Degree of Confirmation, in: *Journal of Symbolic Logic* 10, 1945, pp. 25-60.

[15] Hempel, Carl Gustav: A Purely Syntactical Definition of Confirmation, in: *Journal of Symbolic Logic* 8, 1943, pp. 122-143.

[16] Hempel, Carl Gustav: Studies in the Logic of Confirmation, in: *Mind* 54, 1945, pp. 1-26, pp. 97-121; reprinted in: Hempel, Carl Gustav: Aspects of Scientific Explanation and Other Essays in the Philosophy of Science, New York: The Free Press, 1965 pp. 3-51.

[17] Hempel, Carl Gustav: Inductive Inconsistencies, in: *Synthese* 12, 1960, pp. 439-469; reprinted in: Hempel, Carl Gustav: Aspects of Scientific Explanation and Other Essays in the Philosophy of Science, New York: The Free Press, 1965 pp. 53-79.

[18] Hempel, Carl Gustav: Deductive-Nomological vs. Statistical Explanation, in: Feigl, Herbert/Maxwell, Grover (eds.): Minnesota Studies in the Philosophy of Science, vol. 3, Minneapolis: University of Minnesota Press, 1962, pp. 98-169.

[19] Hempel, Carl Gustav/Oppenheim, Paul: A definition of Degree of confirmation, in: *Philosophy of Science* 12, 1945, pp. 98-115.

[20] Hempel, Carl Gustav/Oppenheim, Paul: Studies in the Logic of Explanation, in: *Philosophy of Science* 15, 1948, pp. 135-175; reprinted in: Hempel, Carl Gustav: Aspects of Scientific Explanation and Other Essays in the Philosophy of Science, New York: The Free Press, 1965 pp. 245-295.

[21] Hendricks, Vincent F.: Forcing Epistemology, Cambridge: Cambridge University Press, 2004.

[22] Hintikka, Jaakko/Pietarinen, Juhani: Semantic Information and Inductive Logic, in: Hintikka, Jaakko/Suppes, Patrick (eds.): Aspects of Inductive Logic, Amsterdam: North-Holland, 1966, pp. 96-112.

[23] Hosiasson-Lindenbaum, Janina: On Confirmation, in: *Journal of Symbolic Logic* 5, 1940, pp. 133-148.

[24] Huber, Franz: Assessing Theories. The Problem of a Quantitative Theory of Confirmation, University of Erfurt: Doctoral Dissertation, 2002.

[25] Huber, Franz: Subjective Probabilities as Basis for Scientific Reasoning?, forthcoming in: *The British Journal for the Philosophy of Science*, 2004.

[26] Joyce, James M.: The Foundations of Causal Decision Theory, Cambridge: Cambridge University Press, 1999.

[27] Kelly, Kevin T.: The Logic of Reliable Inquiry, Oxford: Oxford University Press, 1996.

[28] Kelly, Kevin T.: Iterated Belief Revision, Reliability, and Inductive Amnesia, in: *Erkenntnis* 50, 1999, pp. 11-58.

[29] Kemeny, John G.: A Logical Measure Function, in: *Journal of Symbolic Logic* 18, 1953, pp. 289-308.

[30] Kemeny, John G.: Fair Bets and Inductive Probabilities, in: *Journal of Symbolic Logic 20*, 1955, pp. 263-273.

[31] Kemeny, John G./Oppenheim, Paul: Degree of Factual Support, in: *Philosophy of Science* 19, 1952, pp. 307-324.

[32] Kraus, Sarit/Lehmann, Daniel/Magidor, Menachem: Nonmonotonic Reasonig, Preferential Models and Cumulative Logics, in: *Artificial Intelligence* 40, 1990, pp. 167-207.

[33] Lehman, R. Sherman: On Confirmation and Rational Betting, in: *Journal of Symbolic Logic* 20, pp. 251-262.

[34] Levi, Isaac: Decision Theory and Confirmation, in: *Journal of Philosophy* 58, 1961, pp. 614-625.

[35] Levi, Isaac: Corroboration and Rules of Acceptance, in: *The Britisch Journal for the Philosophy of Science* 13, 1963, pp. 307-313.

[36] Levi, Isaac: Gambling With Truth. An Essay on Induction and the Aims of Science, London: Routledge, 1967.

[37] Levi, Isaac: Probabilistic Pettifoggery, in: *Erkenntnis* 25, 1986, pp. 133-140.

[38] Lipton, Peter: Inference to the Best Explanation, London: Routledge, 1993.

[39] Milne, Peter: Is There a Logic of Confirmation Transfer?, in: *Erkenntnis* **53**, 2000, pp. 309-335.

[40] Shimony, Abner: Coherence and the Axioms of Confirmation, in: *Journal of Symbolic Logic* 20, 1955, pp. 1-28.

[41] Spohn, Wolfgang: Ordinal Conditional Functions: A Dynamic Theory of Epistemic States, in: William L. Harper and Brian Skyrms (eds.): Causation in Decision, Belief Change, and Statistics II, Dordrecht: Kluwer, 1988, pp. 105-134.

[42] Spohn, Wolfgang: A General Non-Probabilistic Theory of Inductive Reasoning, in: R.D. Shachter et. al. (eds.): Uncertainty in Artificial Intelligence 4, Amsterdam: North-Holland, 1990, pp. 149-158.

[43] van Fraassen, Bas C.: Theory Comparison and Relevant Evidence, in: John Earman (ed.): Testing Scientific Theories, Minneapolis: University of Minnesota Press, 1983, pp. 27-42.

[44] Zwirn, Denis/Zwirn, Hervé P.: Metaconfirmation, in: *Theory and Decision* 41, 1996, pp. 195-228.