Wann ist dann ein Satz wahr? Tarski verweist auf eine anaphorische Struktur und simplifiziert auf folgende Art:

(10) »*Schnee ist weiß*« ist wahr, dann und nur dann, wenn Schnee weiß ist.

Der Satz in Anführungszeichen bezieht sich auf einen nachkommenden Satz. Beide Sätze stehen in Beziehung zueinander, obwohl sie gemäß dem Trennungsprinzip auf verschiedenen Ebenen verwendet werden. Anscheinend stehen beide Sätze anaphorisch zueinander in Beziehung, wenn die Spezifizierung des gebrauchten Satzes von der Struktur des ersteren in irgendeiner Form abhängt.

4. Schluss

Mit der Wahrheit ist es ähnlich wie mit der Selbstbenennung des Odysseus als »Niemand«. Es treten ähnliche Strukturen wie in der *Odyssee* auf, in der viele plausible Sätze nicht wahr und viele phantastische Sätze wahr sind. Wenn wir uns mit Szenerien wie in der *Odyssee* konfrontiert sehen, in denen wir nicht zwischen plausiblen Lügen und phantastischen Wahrheiten unterscheiden können, so muss es eine Reihe von Sätzen geben, die wahr sind. Unsere Kenntnis der Botanik erlaubt uns, zu behaupten, dass *Gras grün ist*. Wir können uns aber genauso gut auf Sätze beziehen, die phantastische Situationen beschreiben, indem wir behaupten, dass *Odysseus aus Ithaca stammte*. Sollten wir mit jemandem über ein spezielles Gras diskutieren, das nicht ganz grün erscheint, so hätten wir es mit Abstraktionsproblemen zu tun. Sollte sich jemand nicht auf Homers Gestalt beziehen, sondern nur den Ulysses von James Joyce kennen, dann würde die Aussage gemäß dem Trennungsprinzip selbstverständlich nicht stimmen.

LITERATUR

Davidson, D. (1990), The Structure and Content of Truth. *Journal of Philosophy* 87, 279–328.

Etchemendy, J. (1988), Tarski on Truth and Logical Consequence. *Journal of Symbolic Logic* 53, 51–79.

Frege, G. (1892), Über Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100, 25–50.

Tarski, A. (1935), Der Wahrheitsbegriff in den Formalisierten Sprachen. *Studia Philosophica* 1, 261–405.

O. Neumaier/C. Sedmak/M. Zichy (Hg.): *Philosophische Perspektiven. Beiträge zum VII. Internationalen Kongress der ÖGP.* Frankfurt/M.–Lancaster 2005: Ontos, 309–315.

THE LOGIC OF CONFIRMATION

Franz Huber (Konstanz)*

The paper presents a new analysis of Hempel's (1945) conditions of adequacy, differing from the one in Carnap (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 were conflicting, and so he gave up the concept of confirmation aiming at informative theories. It is then shown that one can have the cake and eat it: There is a logic of confirmation that accounts for both of these two conflicting aspects.

1. Introduction

In his »Studies in the Logic of Confirmation« (1945) Carl G. Hempel presented the following conditions of adequacy for any relation of confirmation $\partial \subseteq L \times L$ on a language (set of sentences closed under negation and conjunction) L: For all E, H in L,

(1) Entailment: $E \vdash H \Rightarrow E \partial H$
(2) Consequence: $\{H: E \partial H\} \vdash H' \Rightarrow E \partial H'$
(2.1) Special Consequence: $E \partial H,\ H \vdash H' \Rightarrow E \partial H'$
(2.2) Equivalence: $E \partial H,\ H \spadesuit\!\vdash H' \Rightarrow E \partial H'$
(3) Consistency: $\{E\} \cup \{H: E \partial H\} \heartsuit \neg E$
(3.1) $E \heartsuit \bot,\ E \partial H \Rightarrow E \heartsuit \neg H$
(3.2) $E \heartsuit \bot,\ E \partial H,\ H \vdash \neg H' \Rightarrow E \heartsuit H'$
(4) Converse Consequence: $E \partial H,\ H' \vdash H \Rightarrow E \partial H'$

* *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.

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 L: $E \partial H$ (obviously (1) and (4) are sufficient for this to hold).

Since Hempel's negative result, there has hardly been any progress in constructing a logic of confirmation. 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 (viz. the logic of positive statistical relevance), but that it does not deserve to be called a *logic*. One reason for this seems to be that up to now the predominant view on Hempel's conditions is the analysis Carnap (1962, § 87) gave in his *Logical Foundations of Probability*.

Carnap's analysis can be summarized as follows: In presenting his first three conditions Hempel was mixing up two distinct concepts of confirmation, two distinct *explicanda* in Carnap's terminology, viz.

(i) the concept of *incremental confirmation* (positive statistical relevance, initially confirming evidence in Carnap's terminology) according to which E confirms H iff E increases the probability of H, $\Pr(H \mid E) > \Pr(H)$, and

(ii) the concept of *absolute confirmation* according to which E confirms H iff the probability of H given E is greater than some value r, $\Pr(H \mid E) > r$.

Hempel's second and third conditions hold true for the second *explicandum*, but they do not hold true for the first *explicandum*. On the other hand, Hempel's first condition holds true for the first *explicandum*, but it does so only in a *qualified* form (Carnap 1962, 473) – namely only if H is not already assigned probability 1. This, however, means that Hempel first had in mind the *explicandum* of incremental confirmation for the entailment condition; then he had in mind the *explicandum* of absolute confirmation 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.

2. Conflicting Concepts of Confirmation

We present another analysis based on the following two notions. A relation $\partial \subseteq L \times L$ is a *likeliness* relation on the language L iff

$$E \partial H,\ H \vdash H' \Rightarrow E \partial H';$$

∂ is a *loveliness* relation on L iff

$$E \partial H,\ H' \vdash H \Rightarrow E \partial H'.$$

These two concepts underlie the two main approaches to confirmation that have been put forth in the last century: qualitative *Hypothetico-Deductivism* HD, and quantitative probabilistic *Inductive Logic* IL.

According to HD, E HD-confirms H iff H logically implies E. Hence, if E HD-confirms H and H' logically implies H, then E HD-confirms H'. So HD-confirmation is a loveliness relation.

According to IL, the degree of *absolute* confirmation of H by E equals the probability of H given E, $\Pr(H \mid E)$. The natural *qualitative* counterpart of this quantitative notion is that E »absolutely confirms« H iff $\Pr(H \mid E) > r$, for some value r in $[.5, 1)$ (Carnap's second *explicandum*). If E absolutely confirms H and H logically implies H', then E absolutely confirms H'. So absolute confirmation is a likeliness relation.

This is not the way Carnap (1962, ch. VII) defined qualitative IL-confirmation. He rather required that E *raises* the probability of H in order for E to qualitatively IL-confirm H, i.e. $\Pr(H \mid E) > \Pr(H)$. Nevertheless, the above seems to be the natural qualitative counterpart of the degree of absolute confirmation. The reason is that later on, the difference between $\Pr(H \mid E)$ and $\Pr(H)$ was taken as degree of *incremental confirmation*, and Carnap's proposal is the natural qualitative counterpart of this notion of incremental confirmation. Let us say that E *incrementally confirms* H iff $\Pr(H \mid E) > \Pr(H)$.

The loveliness concept underlying HD aims at *informative* theories, whereas the likeliness concept underlying IL aims at *true* (probable) theories. These two concepts are conflicting in the sense that the first increases, whereas the second decreases with the logical strength of the theory to be assessed.

3. Hempel Vindicated

Turning back to Hempel's conditions, note first that Carnap's second *explicandum* satisfies the entailment condition *without* qualification: If E logically implies H, then $\Pr(H \mid E) = 1 > r$, for any value r in $[.5, 1)$. 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. 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«, the present analysis makes perfect sense of his 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.

4. The Logic of Theory Assessment

However, in a sense one can have the cake and eat it: There is a logic of confirmation 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 $< W, \mathbf{A}, \kappa>$ be a ranking space, where W is a non-empty set of possibilities, $\mathbf{A}$ is a field over W, i. e. a set of subsets of W containing the empty set $\emptyset$ and closed under complementation and finite intersections, and κ: $W \rightarrow \mathbf{N} \cup \{\infty\}$ is a *ranking function* (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 w in W is assigned rank 0. κ is extended to a function on $\mathbf{A}$ by setting $\kappa(\emptyset) = \infty$ and defining, for each non-empty A in $\mathbf{A}$,

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

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

$$\kappa(B \mid A) = \kappa(A \cap B) - \kappa(A) \text{ if } \kappa(A) < \infty,$$
$$= 0 \qquad \text{if } \kappa(A) = \infty.$$

A ranking function represents an ordering of *dis*belief. For A, B in $\mathbf{A}$, $\kappa(B^C \mid A) - \kappa(B \mid A)$ measures how *likely* B is given A, whereas $\kappa(B \mid A^C) - \kappa(B^C \mid A^C)$

measures how much B *informs about* A, where A^C is the set-theoretical complement of A w. r. t. W.

A ranking space $< W, \mathbf{A}, \kappa>$ is an *assessment model* for the language L iff W is the set Mod_L of all models for L, $Mod(\alpha) \in \mathbf{A}$ for each α in L, and $\kappa(Mod(\alpha)) < \infty$ for each consistent α in L. The consequence relation ∂_κ on the language L induced by an assessment model $<Mod_L, \mathbf{A}, \kappa>$ is defined as follows: For all H, E in L,

$$E \partial_\kappa H \Leftrightarrow \kappa(Mod(H) \mid Mod(E)) \leq \kappa(Mod(\neg H) \mid Mod(E)) \;\&$$
$$\kappa(Mod(\neg H \mid Mod(\neg E)) \leq \kappa(Mod(H) \mid Mod(\neg E))$$

where at least one of these inequalities is strict. In words: H is an acceptable theory given E (according to κ) iff H is at least as likely as and more informative than $\neg H$ given E, or H is more likely than and at least as informative as $\neg H$ given E.

On the other (the syntactical) hand, a relation $\partial \subseteq L \times L$ is an *assessment relation* on L iff ∂ satisfies the following principles, for all E, H in L:

(A1) $E \partial E$	Reflexivity*
(A2) $E \partial H, E \spadesuit\!\vdash E', H \spadesuit\!\vdash H' \Rightarrow E' \partial H'$	L-Equivalence*
(A3) $E \partial H \Rightarrow E \partial E \wedge H$	Weak Composition*
(A4) $E \partial H \Rightarrow \neg E \partial \neg H$	Loveliness and Likeliness
(A5) $\heartsuit E \vee H \Rightarrow E \vee H \partial E$ or $E \vee H \partial H$	Either-Or
(A6) $E \vee H \{ E, \heartsuit E \vee H \Rightarrow E \vee \neg E \partial \neg E$	Negation
(A7) $E \vee F \partial E, F \vee H \partial F, \heartsuit E \vee H \Rightarrow E \vee H \partial E$	quasi Nr 21
(A8) $E \vee F \partial E, F \vee H \partial F, \vdash E \vee H \Rightarrow E \vee H \{ \neg E$	suppl. Nr 21
(A9) $E_i \vee E_{i+1} \partial E_{i+1}, \heartsuit E_i \vee E_j \Rightarrow \exists n \forall m \geq n: E_m \vee E_{m+1} \partial E_m$	Minimum
(A10) $E \partial E \wedge H, E \partial E \vee H \Rightarrow E \{ \neg H$	
(A11) $E \{ E \wedge \neg H, E \partial E \vee H, \heartsuit E \Rightarrow E \partial H$	
(A12) $E \wedge \neg E \partial E, E \vee H \partial E \Rightarrow E \wedge \neg E \partial H$	

The * starred principles are among the *core principles* in Zwirn/Zwirn (1996). Quasi Nr 21 without the restriction $\heartsuit E \vee H$ is the derived rule (21) of the system **P** in Kraus/Lehmann/Magidor (1990).

Theorem: The consequence relation ∂_κ induced by an assessment model $<Mod_L, \mathbf{A}, \kappa>$ for L is an assessment relation on L. For each assessment relation ∂ on L there is an assessment model $<Mod_L, A, \kappa>$ for L such that $\partial = \partial_\kappa$.

The following principles are admissible:

(A15) $E \partial F \rightarrow H, E \vdash F \Rightarrow E \partial H$ — MPC
(A16) $E \partial F, E \vdash H \Rightarrow E \partial F \wedge H$ or $E \partial F \vee H$ — quasi Composition
(A17) $E \not\partial E \wedge \neg E \Rightarrow E \not\partial E \wedge \neg E$ — Consistency*
(A18) $E \not\partial E \vee \neg E \Rightarrow E \not\partial E \vee \neg E$ — Informativeness
(A19) $E \wedge F \partial H, E \wedge \neg F \partial H \Rightarrow E \partial H$ — Proof by Cases

The following principles are not admissible:

(N1) $E \vdash H \Rightarrow E \partial H$ — Entailment (Supraclassicality)*
(N2) $H \vdash E \Rightarrow E \partial H$ — Conversion
(N3) $F \partial H, E \vdash F \Rightarrow E \partial H$ — Left Monotonicity
(N4) $E \partial F, F \vdash \neg H \Rightarrow E \not\partial H$ — Strong Selectivity
(N5) $E \wedge F \partial H, E \partial F \Rightarrow E \partial H$ — Cut
(N6) $E \partial H, E \partial F \Rightarrow E \wedge F \partial H$ — Cautious Monotonicity

In comparing the present approach with standard nonmonotonic logic in the KLM-tradition (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 is not admissible:

(N7) $E \partial F, F \vdash H \Rightarrow E \partial H$ — Right Monotonicity (Right Weakening)

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 be worth taking the risk of being 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): We are willing to take the risk of being led to a false conclusion, because we want to arrive at informative conclusions.

Finally, one might wonder how the present logic of theory assessment compares to Carnap's *dictum* that qualitative confirmation is positive statistical relevance. A first answer is given by

Observation: For every regular probability Pr on a language L,
$\partial_{Pr} = \perp^{+}_{Pr} \cup \{<E, H>: E \dashv\vdash H \dashv\vdash E \vee \neg E\} \cup \{<E, H>: E \dashv\vdash H \dashv\vdash E \wedge \neg E\}$

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

However, theory assessment is not the same as positive statistical relevance, for Symmetry is not admissible:

(N8) $E \partial H \Rightarrow H \partial E$ — Symmetry

REFERENCES

Carnap, R. (²1962), *Logical Foundations of Probability*, Chicago: University of Chicago Press.

Hempel, C.G. (1945), Studies in the Logic of Confirmation, in: *Mind* 54, 1–26, 97–121.

Kraus, S./Lehmann, D./Magidor, M. (1990), Nonmonotonic Reasoning, Preferential Models, and Cumulative Logics, in: *Artificial Intelligence* 40, 167–207.

Milne, P. (2000), Is There a Logic of Confirmation Transfer?, in: *Erkenntnis* 53, 309–335.

Spohn, W. (1988), Ordinal Conditional Functions: A Dynamic Theory of Epistemic States, in: Harper, W.L./Skyrms, B., eds. (1988), *Causation in Decision, Belief Change, and Statistics II*, Dordrecht: Kluwer, 105–134.

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

Zwirn, D./Zwirn, H.P. (1996), Metaconfirmation, in: *Theory and Decision* 41, 195–228.