BAYESIAN RULES FOR THE RATIONAL RECONSTRUCTION OF A SYSTEM OF HYPOTHESES WEAKENED BY ADVERSE OBSERVATIONS

HENRY A. FINCH

> Man kann sich daher nicht genug in acht nehmen, aus Versuchen nicht zu geschwind zu folgern: — Goethe, in *Der Versuch als Vermittler von Objekt und Subjekt.*

Consider the large and important class of scientific theories or belief systems which are logically equivalent to a conjunction of a finite number of hypotheses; i.e. let H symbolize a system of hypotheses such that

$$H \equiv h_1 h_2 \ldots . h_n$$

Let o be a kind of observation whose calculable likelihood given H and prior evidence e is so low that the adverse effect of the unmistakable occurrence of o is seriously to weaken the credibility of H. Reflective as well as spontaneous thought invariably seeks to reconstruct H — especially if H hitherto served well — basically in one or another of four distinguishable ways. A proposition H′ is sought given which the likelihood of o becomes high; ideally Pr(o/H′) is equal to 1. H′ is so related in one of the four fundamental ways to H *that the temptation is strong to believe that H has been saved in whole or part* from the adverse effect of the weakening observation o. I shall show, by Bayes' theorem, that *exact conditions involving prior probabilities and likelihoods are stateable to decide the question whether a proposed H′ with pr(o/H′.e) equal to 1 is or is not a rational reconstruction of H, that is, our conditions will effectively decide whether H is reasonably admissible and not merely and ad hoc hypothesis.* (1)

(1) Since at least the time of Dugald Stewart who had a very good grasp of the issue, the admissibility of a hypothesis H′ when Pr(o/H′.e) = 1 has uniformly been viewed as a question of H′s fruitfulness for future or «novel» observations. Our Bayesian rules are, of source, consistent with the desideratum of theoretical fruitfulness; but they suggest with metrical definiteness that a reference to the prior credibility of H′ and to the relevance of ~H′ is nicely pertinent when deciding to replace *H* by *H′*. To be sure the connections between credibility and theoretical fruitfulness deserve close measure-theoretic and information theoretic analysis. (We may mention

The four methods of reconstruction of H are (a) finding a hypothesis h_{n+1} such that $H.h_{n+1} \equiv H'$ and $Pr(o/H'.e) = 1$; (b) dropping a proper subset of the conjoined hypotheses constituting H, the remaining conjunction being H', with $Pr(o/H') \cong 1$; (c) replacing each h of a proper subset of H by a non-equivalent h; again the likelihood of o on the resulting H' is equal to 1; (d) finding an hypothesis h_{n+1} non-equivalent to H or any conjunctive component of H such that H implies h_{n+1} and also finding an H' to imply h_{n+1} without H' implying H or conversely; as always $Pr(o/H'.e) = 1$.

I take it that the justifiable ambition of method (a) is to assure a lower bound for c(H/e·o) [2] where c is the degree of confirmation or credibility of H on e and o, that is, on the complete history so far of H; c(H/e) being the degree of confirmation of H before it encountered the weakening (undermining) force of o. The ambition of method (a) may also be expressed, rather more obscurely, as the intention to have the entire old system of hypotheses H as «part» of the new system H'. Method (b)'s ambition is apparently to include «part» of H in H' and to assure a value c(H'/e·o) greater than or equal to k, where k is at least c(H/e). Method (c) has the same ambition as (b) but prefers replacing to dropping of components of H. Method (d) which is quite commonly utilized in physics and all highly metricized science has both the intention of raising a «part» of H (really a range

here our exploratory study, «Theoretical Fruitfulness and the Measure of Concepts» pages 165-183 of *Essays in Philosophy*, Penna. State University Press, 1962). See also note 7 below for a heuristic suggestion concerning the possible explication of theoretical fruitfulness in terms of comparative rates of change in credibility. Our rules (a) to (d) and (a') to (d') presuppose that o has a calculable likelihood given He, H'e, ~He, ~H'. I do not consider in the present investigation the reconstruction of H into H' in such ways as to render the likelihood of o given H'.e *calculable* whenever that likelihood given He *was not decidable*. (There are deep waters here not safely navigable since the relationship between deductive and stochastic undecidablity is not too clear to me. By stochastic undecidability I mean the following situation. Let H be a system of hypotheses; let o be an observation, then neither Pr(o/H) nor Pr(~o/H) is effectively calculable.)

(2) This is a reasonable ambition, although I shall not assert that those who are disposed to reconstruct H' by method (a) always restrict themselves so modestly. Method (b) must accept the restriction,

$$Pr(o/H'.e) \leqslant 1 - \frac{c(Ho/e)}{c(H'/e)}.$$

of application of H within which confirmation of H and H′ is exactly or nearly the same) to a minimum level equal to c(H′/e·o), and of attaining $c(H'/e.o) \geqslant c(H/e)$.

It requires no extensive argument to prove that when H′ is constructed in accordance with procedures (a), (b), (c), and (d) the reasonable ambitions of the procedures are not necessarily realized, as even the oldest deductive logic will warn with a caveat against asserting the consequent. To know merely that a procedure will not necessarily succeed is of little positive assistance in discovering the additional conditions under which it will. Our primary aim is to find the required additional conditions from Bayes' theorem.

I shall prove the following decision rules from Bayes' theorem for cases (a), (b), (c), and (d) respectively, H is the hypothesis weakened by o, H′ the reconstructed H. Any hypothesis involved in the transformation of H into H′ or conversely will be called an auxiliary hypothesis. (Bracketed letters after «rule» or «lemma» refer to specified modes of reconstructing H′.) $O \leqslant \Delta < 1$; $\Delta = O$ *except in method (b).*

Rule (a): a sufficient condition that neither c(H/e·o) nor c(H′/e·o) shall fall below the value k is the inequality [3]:

$$\frac{(1-\Delta)\ (1\text{-}k)c(H'/.e)}{k\, c(\sim H'/e)} \geqslant \Pr(o/\sim H'\cdot e) \qquad (\alpha)$$

(There is no loss of generality in taking $k \geqslant c(H'/e)$, and, as required by Bayes' theorem, $c(H'/e) > 0$) where c(H′/e) is the prior confirmation of H′ and the right hand side of the inequality is the likelihood of o (the weakening observation) given ∼H′ and e.).

Lemma (a): Rule (a) is expressible as a necessary and sufficient condition upon $c(h_{n+1}/He)$ in virtue of the relationship $c(H'/e) = c(H/e) \times c(h_{n+1}/H.e)$. It is *therefore clear that our rule not only reccognizes but specifies the degree metrically of «independent» corroboration long demanded by inductive logicians to distinguish permissible auxiliary hypotheses, here h_{n+1}, from recklessly or desperately introduced ad hoc hypotheses.* As is also indicated in lemmas (b), (c), and (d), *inequality* (α) *imposes definite conditions upon all auxiliary hypotheses.*

(3) We prove from (I) below that (α) is a necessary and sufficient condition for $c(H'/e.o) \geqslant k$; but, when, as in case (a), $H' \supset H$, $c(H'/e.o) \leqslant c(H/e.o)$, whence $c(H/e.o) \geqslant k$.

Rule (b): a necessary and sufficient condition that H', a proper subset of the h's conjoined in H, shall have a degree of confirmation on e.o equal to at least c(H/e) is supplied by the inequality (α) upon the substitution therein of c(H/e) for k,

$$\frac{c(H \sim o/e)}{c(H'/e)} \text{ for } \Delta.$$

Lemma (b): Rule (b) is expressible as a necessary and sufficient condition upon the conjunction of dropped h' from H' to form H' in virtue of the following relationship:

$$c(H'/e) \times c\,(\text{conjunction of dropped h's/e:H}) = c(H/e).$$

Rule (c): A necessary and sufficient condition that $H' \equiv H_i H_r$ (where H_i is the conjunction of unmodified h's of H and H_r is the conjunction of the replacements for the h's dropped from H) shall have a degree of confirmation at least equal to c(H/e) is supplied by the inequality (α) upon the substitution therein of c(H/e) for k.

Lemma (c): Rule (c) is expressible as a necessary and sufficient condition upon H_r in virtue of the relationship:

$$c(H'/e) = c(H_i/e) \times c(H_i/H_i e).$$

Rule (d): A necessary and sufficient condition for H' and for h_{n+1} to have a degree of confirmation on e·o equal to at least c(H/e) is supplied by the inequality (α) upon the substitution therein of c(H/e) for k.

Lemma (d): Rule (d) is directly formulated as a necessary and sufficient condition for h_{n+1} as well as for H', because of the deducibility of h_{n+1} from H' assures $c(H'/eo) \leqslant c(h_{n+1}/eo)$ and it is also assured after substitution of c(H/e) for k in (α) that $c(H/e) \leqslant c(H'/eo)$. A necessary condition involving $Pr(o/\sim H'e)$ which $c(h_{n+1}/e)$ must meet is proved in Rule (d') below.

It only remains to demonstrate that inequality (α) is a necessary and sufficient condition for $c(H'/e.o) \geqslant k$, since it is evident that all the other rules are dependent on it. Proof: In Bayes' theorem replace Pr(o/H'.e) by 1-Δ; Bayes' theorem ([4]) is c(H'/e.o) =

([4]) This is Bayes' theorem for known prior probabilities; few modern authors are clearly aware of its latency in Bayes' original memoir; for an exposition of the meaning and use of Bayes' theorem, see H. A. Finch «Con-

$$\frac{c(H'/e) \times Pr(o/H'.e)}{c(H'/e) \times Pr(o/H'\cdot e) + c(\sim H'/e) \times Pr(o/\sim H/.e)};$$

hence

$$c(H'/e\cdot o) = \frac{c(H'/e)\ (1\text{-}\Delta)}{c(H'/e) + c(\sim H'/e) \times Pr(o/\sim H'.e)} \qquad (I).$$

Setting the right hand side of (I) $\geqslant$ k establishes inequality (α) easily as necessary and sufficient for c(H'/e.o) $\geqslant$ k.

We have already noted that the basic inequality (α) satisfies with metrical definiteness the requirement of independent confirmation for all auxiliary hypotheses. We may now explicitly state rules for the conditions which must be satisfied by auxiliary hypotheses involved in transformations from H to H'. The following inequality (β) is demonstrably equivalent to inequality (α):

$$c(H'/e) \geqslant \frac{k\ Pr(o/\sim H'e)}{1 - k\,Pr(\sim o/\sim H'.e) - \Delta(1\text{-}k)}; \qquad (\beta)$$

(β) is, of course, proved by solving inequality (α) for c(H'/e). Hence, for types of reconstruction (a), (b), (c), and (d) respectively, the following conditions upon the auxiliary hypotheses hold:

Rule (a'): A necessary and sufficient condition for c(H'/e.o) $\geqslant$ k is, from inequality (β) and lemma (a),

$$c(h_{n+1}/He) \geqslant \frac{k\ Pr(o/\sim H'e)}{[1 - k\ Pr(\sim o/\sim H'e)]c(H/e)} \quad (a')\ ;\ \text{hence}$$

hence inequality (a') is sufficient for c(H/e.o) $\geqslant$ k.

Rule(b'): A necessary and sufficient condition for c(H'e.o) $\geqslant$ c(H/e) is, from inequality (β) and lemma (b):

firming Power of Observations Metricized,» *Philosophy of Science,* July, Oct. 1960. See also the remarks on corroboration,prior and posterior probability by T. J. Good on page 92 of *Foundations of Statistical Inference* by L. J. Savage and others, London, 1962.

$$c(\text{conjunction of dropped h's}/eH) \leqslant \frac{[1 - c(H/e)\ Pr(\sim o/\sim H'\ e)] - \dfrac{c(H \sim o/e)\ c(\sim H/e)}{c(H'/e)}}{Pr\,(o/\sim H'.\,e)} \qquad (b')$$

Rule (c'): A necessary and sufficient condition for c(H'/e.o) ⩾ c(H/e) is, from inequality (β) and lemma (c)

$$c(H_r/H_i e) \geqslant \frac{c(H/e)\ Pr(o/\sim H'e)}{c(H_i/e)\ [1 - c(H/e)\ Pr(\sim o/\sim H'.e)]} \qquad (c')$$

Rule (d'): A necessary and sufficient condition for c(H'/e.o) ⩾ c(H/e) is, from inequality (β),

$$c(H'/e) \geqslant \frac{c(H/e)\ Pr(o/\sim H'e)}{1 - c(H/e)\ Pr(\sim o/\sim H'e)},$$

but, ex hypothesi, h_{n+1} is such that $H' \supset h_{n+1}$; therefore $c(H'/e) \leqslant c(h_{n+1}/e)$ which, together with the immediately preceding inequality, yields as a necessary condition for $c(h_{n+1}/e)$, Rule (d'):

$$c(h_{n+1}/e) \geqslant \frac{c(H/e)\ Pr(o/\sim H'e)}{1 - c(H/e)\ Pr(\sim o/\sim H'e)} \qquad (d')$$

Manifestly, Rules (a) through (d') must be satisfied before a reconstruction of H into H' is successful even when Pr(o/H'e) = 1. (The general strategy in deriving rules (a) to (d) and (a') to (d') is easily applicable to combinations of methods (a), (b), (c), (d).

A further remarkable characteristic of the strategic inequality of Rule (a) testifies once again to the extraordinary heuristic fruitfulness of Bayes' theorem. ([5]) Popper has rightly insisted that the addition of auxiliary or ad hoc hypotheses in the reconstruction of scientific theories ought not to weaken the testability of reconstructed theories. The fundamental Bayesian inequality (α) is logically a necessary and sufficient condition upon Pr(o/∼H'·e); *this likelihood is undetermined by Pr(o/H'.e)* ([6]) and hence can not be arbitrarily influenced by find-

([5]) For a Bayesian measure of the stability of a system of hypotheses as it is tested by observation, see Finch op. cit., Part I, p. 302.

([6]) Since c(H'o/e) + c(∼H'o/e) = Pr(o/e) and since Pr(o/H'e) = 1-Δ,

$$Pr(o/\sim H'.e = \frac{Pr(o/e) - c(H'/e)(1-\Delta)}{1 - c(H'/e)}.$$ Therefore, as my mathematical

ing such an H′ that likelihood of o given H′ and e is equal to 1. Testability can not be weakened when the strategic parameter for deciding admissibility of a theory's revision is logically and metrically dependent upon the *contradictory* of the new system of hypotheses; and it is surely advisable while maintaining testability not to neglect the sequential history of hypotheses as they rise or fall in credibility with expanding evidence. [7]

Henry A. FINCH

Pennsylvania State University
University Park, Pennsylvania
July 10, 1964

colleague Professor Nathan Fine who graciously read the text pointed out to me, the right hand side of inequalities (a′) (b′) (c′) (d′) can be expressed as functions of c(H/e), Pr(o/e) and C(H′e). Elimination of c(H′/e) is possible in virtue of the connections noted in lemmas (a) to (d) so that inequalities (a′) to (d′) will have on their right side only expressions computable from c(H.o/e) and e(∼H′.o/e); which indeed may be convenient since Pr(o/∼He) may be easier to calculate than Pr(o/∼H′.e).

(7) The following heuristic suggestion merits, I think, further investigation. Let c(h/e) $\leqslant$ c(h_2/e) then if c(h_1/e.o) $>$ c(h_2/e.o), I find it plausible to consider h_1 more fruitful over the range of application determined by e and o than h_2. When I reflection the need for the condition c(h_1/e) $\leqslant$ c(h_2/e), it appears to me that the basis for the plausibility of the judgment of h_1's superiority in fruitfulness is ultimately the greater rate of growth of c(h_1) as the domain of application is extended from e to e and o. Such superior rates of growth when assessed for their integral effects are crucial for growing credibility over progressively wider and varied domains of application. Fruitfulness is such relevance as maximizes information over expanding domains of application. We may refer to pages 304 of our (1960) essay on confirming power cited in note 4. for a conceptualization of hypothesis testing as a procedure which diminishes uncertainty of choice between hypotheses by extracting information from observations; the importance of *rates of change* in degrees of credibility or confirmation is noticed ibid. page 294. See also I. J. Good's masterly independent 1960 study, 'Weight of Evidence, Corroboration, Explanatory Power, Information and the Utility of experiments», *Journal of the Royal Statistical Society* B, 22, pages 319-331. Our effort to state a formal rationale for the evolution of scientific theories is encouraged by a comment of S. Toulmin who, despite an avowed skepticism concerning evidential calculi, agrees that the scientist's task typically is «to accomodate some new discovery to his inherited ideas without needlessly jeoparding the intellectual gains of his predecessors» (p. 112, *«Foresight and Understanding»*, N.Y., 1963).