REPRINT

PHILOSOPHIA

PHILOSOPHICAL QUARTERLY OF ISRAEL

VOL. 7, NO.2 JUNE 1978

BAYESIAN LEARNING MODELS WITH REVISION OF EVIDENCE

WILLIAM HARPER

BAR-ILAN UNIVERSITY

PHILOSOPHIA

PHILOSOPHICAL QUARTERLY OF ISRAEL

edited by Asa Kasher

BAYESIAN LEARNING MODELS WITH REVISION OF EVIDENCE

WILLIAM HARPER

1. Introduction

This paper is a brief sketch of an extension of the orthodox Bayesian model for learning. On the orthodox Basesian model an ideally rational agent learns from experience by conditionalizing on new evidence relative to the probability function that represents his present belief state. This does not allow for situations where giving up previously accepted evidence is rational for such an agent. Two generalizations of the concept of conditional belief will be combined to produce an extension of the Bayesian learning model that can accommodate two ways in which even an ideally rational agent might be called upon to revise previously accepted evidence.

The version of the orthodox model I shall work with is one where the agent can accept corrigible propositions and where he assigns degree of belief 1 to every proposition he accepts. The motivation for the extension produced here is to provide Bayesian models where one can have acceptance without great restrictions on the kind of learning situations that can be represented. I hope that the flexibility and usefulness of these models will help to support the idea that acceptance can be a key concept for investigating rational belief change.

Our extension will share with the orthodox model the assumption that an ideally rational agent has *coherent* beliefs, where "coherent" has its usual technical sense.[1] This assumption amounts to the claim that the agent has actual degrees of belief which determine a unique probability function P and that his conditional degree of belief in B on A equals the corresponding ratio conditional probability,

$$P(\mathrm{B}/\mathrm{A}) = P(\mathrm{AB})/P(\mathrm{A}),$$

for any A where $P(\mathrm{A}) > 0$. This insures that the agent has conditional beliefs adequate to rationally plan for hypothetical situations where he comes to accept new evidence that does not conflict with any proposition of which he is now certain. These are exactly the situations that would be covered by classical conditionalization. Coherence does not preclude an agent endowed with further conditional beliefs that are not covered by the conditional probabilities determined by *P*. Each of the generalizations of conditional belief will correspond to a generalization of the concept of conditional probability appropriate to guide the hypothetical reasoning an agent would require to rationally plan how he would handle the corresponding kind of revision of what he accepts.

The objects of belief are regarded as propositions and propositions are construed set theoretically. The domain of a belief function will be a *field E* of subsets of a fundamental set T of alternative possible states of the world. The propositional operations AB, A and A∨B are represented by the corresponding set theoretical operations, $A \cap B$, T−A, and $A \cup B$. Entailment is represented by inclusion so that the basic set T of alternatives is a proposition entailed by every proposition in *E* and the empty set ϕ is a proposition that entails every proposition in *E*. These are respectively the necessary proposition and impossible proposition relative to field *E*.

The point of constructing a Bayesian model is, usually, to illuminate some problem of decision or inquiry. This problem determines what ought to count as the fundamental set of alternative possible states and what is to be the appropriate field of propositions. The states are to be specified in whatever detail is required to capture all the distinctions relevant to the problem. Often the states can be restricted to alternative specifications of some relatively specific situation. On this version of the set theoretical treatment, propositions correspond exactly to Savage's events (Savage 1953, pp. 8-12, 87-90).

In addition to coherence we assume the agent is semantically omniscient with respect to the field of propositions in the model.[2] At any time the set ∇ of propositions he accepts is consistent and closed under semantical consequence, so that

(1) $\cap\nabla \neq \Phi$, and

(2) $\cap\nabla \subseteq A$ only if $A \in \nabla$

This has the effect that $\cap\nabla$ acts as a single proposition that corresponds to the total content of all the propositions the agent accepts. It will be convenient to call $\cap\nabla$ the agent's *acceptance context*.

Let K be the acceptance context for a Bayesian agent relative to field of propositions *E*. As such an acceptance context K satisfies

(3) $K = \cap\{A \in E\colon K \subseteq A\}$,

and any proposition A is accepted relative to K just in case $K \subseteq A$. The states in K are exactly the states the agent treats as live options; therefore, the coherence requirement may properly be made relative to K. This has the effect that the agent's degrees of belief determine a probability function *P* on E such that

(4) $P(A)=1$, if $K \subseteq A$

and

(5)

(5) $P(A)=P(A \cap K)$

for all A in E.[3]

From this point of view the conditional probability $P(B/A)$ can be regarded as the degree of belief in B appropriate to an alternative acceptance context K(A) which is the minimal revision of K needed to have a context where A is accepted. Whatever else we may want to say about the idea of minimal revision of an acceptance context, this much is clear: If A is compatible with K, as is required to have $P(A)>O$, then the minimal revision of K needed to accept A is just $K \cap A$. Any context where A is accepted must be included in A. Unless K(A) is included in K, as well, more old information will be given up than is required to revise K to accept A. If $K \cap A$ is not included in K(A) then more new information will be accepted than is required to revise K to accept A. The classic orthodox defenses of the coherence condition that the agent's conditional degree of belief in B given A ought to equal $P(AB)/P(A)$ when $P(A)>O$ are arguments to the effect that this ratio is the proper degree of belief in B to guide decisions where the relevant states are those in $K \cap A$ (Ramsey 1926 pg. 79, de Finetti 1937 pp. 108-109 and Savage 1953 pp. 21-24 and pp. 43, 44). Our observation that $K \cap A$ is the minimal revision of K needed to assume A when $P(A)>O$ suggests that the idea of minimal revision of an acceptance context can be profitably applied to the concept of conditional belief.

2. Assumption Contexts

The first kind of rational revision of previously accepted evidence we want to accommodate in our learning models consists of cases where the agent is simply *confronted with the truth of some proposition that conflicts with what he has previously accepted.* A person who carefully examines the empty money compartment of his wallet does not rationally wonder if perhaps there is still money in it. He is forced to accept that his money is not in it. This could happen even if he had taken careful notice when he put the money in, taken precautions to stow the wallet in an inner pocket, and had no reason to expect that anyone would even try to get at it. Cases like this are *prima facie* ones where it was rational to accept some claim in the first place, and also rational to give it up in the face of a surprising new input from experience. Any theory that allows for acceptance of corrigible claims as evidence must allow for this kind of situation.[4]

I want to extend the Bayesian framework so that rational acceptance followed by equally rational revision can be accommodated. The first step is to endow the agent with the ability to carry out hypothetical reasoning relative to assumptions that conflict with what he accepts. This extension will follow the idea of minimal revision of an acceptance context we have, already, applied to cases covered by orthodox Bayesian conditional probability. Suppose K is the acceptance context for an agent. Hypothetical reasoning relative to assumption A, for this agent, is reasoning relative to an alternative context *K(A)* which he treats as *the minimal revision of K required to assume A.* Let us call this alternative context the agent's "A-assumption context".

We have, already, seen that when A is compatible with K a rational agent ought to treat K∩A as the unique nearest A-context to K. No such specification of what a rational agent must treat as the nearest A-context to K is available for cases where A is incompatible with K. This may lead one to suspect that there may not be any single context which the agent treats as the nearest A-context to K. Might not the agent, for a familiar example, be unable to decide between an I-context where Bizet and Verdi are both Italian, and an F-context where they are both French as minimal revisions needed to assume C – that they are compatriots?[5] If we think of an assumption context in the following way the answer to this question must be: No!

What counts as a C-assumption context for the agent is simply the intersection of all and only those propositions he is sure he would commit himself to by assuming C. His inability to decide between I and F is not evidence that he vacillates between two candidates for K(C); instead it shows that he does not have a sure commitment to either. This has the effect that neither I nor F is accepted relative to his C-assumption context. He may have commitment to I∪F, but his

vacillation indicates that both I-states and F-states are among the possibilities left open by K(C).

The heart of our extension is a set of general constraints on assumption contexts motivated by the idea that K(A) should be a minimal revision of K needed to assume A. These constraints do not determine K(A) uniquely except in the case where A is compatible with K. They are, nevertheless, sufficiently strong to generate interesting consequences for belief functions that are coherent with respect to systems of contexts meeting them.

Our preliminary constraint is required by the idea that K(A) is to be an acceptance context for field *E* of propositions.

(0) $K(A)=\cap\{B\varepsilon E : K(A)\subseteq B\}$,

The next two are obvious from the need to have K(A) be an A-context and the fact that no revision at all is required to accept A if A is already accepted.

(1) $K(A)\subseteq A$

(2) If $K\subseteq A$ then $K(A)=K$

The third constraint is more complex.

(3) If $A\subseteq B$ and $K(B)\cap A\neq\Phi$, then $K(A)=K(B)\cap A$

Its motivation is as follows: Given that $A\subseteq B$ any revision of K sufficient to accept A is also sufficient to accept B. This suggests that $K(A)\subseteq K(B)$, unless minimal revision of K to assume A requires giving up something accepted relative to K(B). That $K(B)\cap A\neq\phi$ rules out this case, for it insures that A is compatible with K(B). Therefore, on the hypothesis that $A\subseteq B$ and $K(B)\cap A\neq\phi$ K(A) ought to be a subset of $K(B)\cap A$. The other inclusion ought to hold because violating it would require K(A) to be a context where more is accepted than is required to revise K to assume A.

An empty K(A), as would be the case with $K(\phi)$, is not a viable acceptance context at all, and cannot be used as a base from which to make further assumptions non-trivially.

(4) If $K(A)=\phi$ then $K(A)(B)=\phi$

Having K(A) empty is for the agent to treat A as though it would commit him to something as impossible as a contradiction.

We want to be able to produce many interesting series of contexts.

$$K, K(A_0), K(A_0)(A_1), \ldots, K(A_0)\ldots(A_n)$$

where multiple revisions of evidence are required and where each transition satisfies our constraints. Such a construction has been carried out using the idea that *K(A) should be the set of all possible A-states the agent treats as conflicting least with the information he accepts relative to K.* Any ordering of the propositions accepted relative to K, that can correspond to the relative importance an agent attaches to preserving his acceptance of them when assuming A, will determine a comparative similarity relation of states to K which generates a K(A) that satisfies the constraints. This procedure can be iterated to produce sequences of revisions as complex as one wants. (The details of this construction will appear elsewhere in a more technical treatment of these matters. A less elegant construction along similar lines is already available in Harper 1976.)

3. Question Openings

The second kind of revision we want to provide for consists in cases where the agent *moves to a weaker acceptance context in order to open up for investigation some proposition that conflicts with what he previously accepted.* It will be convenient to label such moves *"question openings"*. A question opening revision would be appropriate whenever the agent is faced with putative new evidence that conflicts with what he accepts, is compelling enough to shake his confidence in some of what he accepts, but is not so overwhelmingly compelling as to force its own immediate acceptance.

Where K is our agent's present acceptance context and A is a proposition, let *K(?A)* be the context he treats as *the minimal revision of K needed to avoid begging the question against A.* Where A is incompatible with K, opening the question for A requires expanding the context to add some A-states as viable alternatives. Characterization of K(?A) reduces to the problem of selecting which A-states to add. The idea that K(A) is the set of possible A-States the agent treats as least in conflict with the information he accepts relative to K suggests the following unique characterization for K(?A):

$$K(?A)=K\cup K(A).$$

This characterization is designed to provide the right kind of revision for cases where A is incompatible with K. It also performs correctly when A is compatible with K. In these cases no expansion is needed to avoid begging the question against A; therefore, K(?A) ought to be just K. On our definition this follows because the general constraints on K(A) insure that $K(A)=K\cap A$ if $K\cap A\neq\phi$.

We have begun with assumption contexts and defined question opening contexts in terms of them. We could have proceeded the other way around and used the identity

$$K(A)=K(?A)\cap A$$

to define the A-assumption context. The idea of minimal revision to assume A and the idea of minimal revision to avoid begging the question against A are interdependent. Apparatus adequate to handle one of these problems is easily adapted to handle the other as well.

The usefulness of question opening contexts is not confined to modelling learning processes that involve an agent really making pure question opening revisions. For purposes of debate I may want to argue from assumptions that do not beg the question in favor of some proposition A which I accept and am trying to convince you to accept as well. When I plan what to say it may be useful for me to reason hypothetically relative to $K(?\bar{A})$ my question opening context for $\bar{A}$.

Assumption contexts as well as question openings can play this kind of role in debates. This can also happen in other rational activities such as scientific investigations. Many experiments are properly analysed relative to assumptions that the scientist does not really accept. Indeed, I am prepared to argue that this kind of situation would give an important role for acceptance contexts and their revisions in an account of rational scientific practice even if no rational agent ever really accepted any contingent proposition.

4. Learning Models

We want to model learning episodes with sequences of revision which include both question openings and assumption makings. We can use a sequence of propositions to represent a succession of assumption makings. To include question openings, however, we want to represent questions. We make questions distinct from both propositions and states in order to avoid ambiguity. Let ? be a one-one function with domain *E* and range (Rng?) disjoint from $E\cup T$. Sequences drawn from $E\cup$Rng? can represent sequences of revisions that include question openings as well as assumption makings. The occurrence of ?A at the m-th term indicates that the m-th revision in the sequence is a move to open the question for A.

We shall first define a system of sequences of acceptance contexts that can represent various possible sequences of revision. Our full learning models will be characterized by adding degree of belief functions that satisfy appropriate coherence constraints with respect to these systems.

$\langle E T?K\rangle$ *is an IAQ* (System of iterated assumption and question opening contexts) iff E is a field of subsets of T, ? is a one-one function with domain E and range disjoint from $E \cup$T, K is a function mapping $\cup_{n \varepsilon \omega}(E\cup \text{Rng}?)^n$ into PT such that $K\phi \neq \phi$ and for all $n \varepsilon \omega$, $(C)_n \,\varepsilon\, (E\cup\text{Rng}?)^n$, $X \varepsilon E\cup\text{Rng}?$, A, B$\,\varepsilon\,$E.

(0) $K(C)_n = \cap\{A \varepsilon E : K(C)_n \subseteq A\}$
(1) $K(C)_n(A) \subseteq A$
(2) If $K(C)_n \subseteq A$ then $K(C)_n(A) = K(C)_n$
(3) If $B \subset A$ and $K(C)_n A(A) \cap B \neq \phi$ Then
$K(C)_n(B) = K(C)_n(A) \cap B$
(4) If $K(C)_n = \phi$ then $K(C)_n(X) = \phi$
(5) $K(C)_n(?A) = K(C)_n \cup K(C)_n(A)$.

In this definition K is a function mapping sequences drawn from $E\cup$Rng? into subsets of T. $K\phi$, the subset assigned by K to the empty sequence, is the acceptance context that all the sequences of revision begin with. No confusion will arise if we continue to use K to designate this acceptance context as well as the function. Where $(C)_n$ is a sequence in $(E\cup\text{Rng}?)^n$ the sequence $(C)_n(A)$ is $C_o, \ldots, C_{n-1}$, A and $(C)_n(?A)$ is $C_o, \ldots, C_{n-1}$, ?A. This continues the practice of using K(A)(?B) to indicate the result of first revising the initial context K to assume A and then revising K(A) to open the question for B.

Our learning models are the iterated conditional belief systems resulting from adding appropriate degreee of belief functions $P_{(C)n}$ corresponding to the various acceptance contexts $K(C)_n$.

$\langle E T?KP\rangle$ is *ICB* (an iterated conditional belief system) iff $\langle E T?K\rangle$ is IAQ, and P is a function with domain $\cup_{n \varepsilon \omega}(E\cup\text{Rng}?)^n$ such that for each $n \varepsilon \omega$ and $(C)_n \,\varepsilon\, (E\cup\text{Rng}?)^n$

(a) $P_{(C)n}$ is a coherent degree of belief function with respect to $K(C)_n$, if $K(C)_n \neq \phi$
(b) $P_{(C)n}$ maps E into $\{1\}$ if $K(C)_n = \phi$.

The first of these is the obvious coherence constraint corresponding to the idea that $P_{(C)n}$ should be a degree of belief function appropriate to guide hypothetical reasoning relative to $K(C)_n$. The second is a convention to handle cases where $K(C)_n$ is empty. This convention is designed to reflect the idea that having K(A) empty is to treat A as an absurd or impossible assumption. The attitude toward such an assumption captured by the convention is that to assume A would commit one to every proposition even the impossible proposition ϕ.

Our conditions on assumption contexts insure that conditional belief relative to assumptions satisfies Karl Popper's characterization

of a generalization of conditional probability where $P(B/A)$ can be non-trivial even when $P(A)=0$ (see Harper 1975 for a proof and further references to Popper's work). Popper's main transition axiom is equivalent to a transition principle developed by Alfred Rényi for a similar generalization of conditional probability motivated by the idea of using unbounded measures to generate conditional probability functions (Van Fraassen 1976, Rényi 1955). This transition principle corresponds directly to our condition (3) on assumption contexts.

5. A Glimpse Beyond

If an agent treats A as impossible by having $K(\bar{A}) = \phi$ then he, also, treats A as necessary. Let $K^* = \cap\{A \varepsilon E\colon K(A) = \phi\}$ This is the intersection of the set of propositions he treats as necessary in the belief state corresponding to the initial acceptance context K. It acts as an accessibility region around K and around $K(C)_n$ for each sequence $(C)_n$ leading to a non-empty context. Indeed, it *determines* whether a sequence leads to a viable context by controlling the admissible transitions along the way. If $K(C)_n \neq \phi$ then $K(C)_n(A) \neq \phi$ if and only if $A \cap K^{**} \neq \phi$.

It is natural to think of the structure of equivalence classes imposed on the propositions by K* as a representation of the agent's conceptual framework. Our basic ideas of minimal revision for assumption and question opening can be applied, on the next level up, to K*. This will produce further generalizations of conditional belief that correspond to modifications of K* to make new propositions necessary or possible. Such apparatus is suitable for an extension of the learning model that can accommodate conceptual change. These developments are a topic for another paper.

UNIVERSITY OF WESTERN ONTARIO
LONDON, ONTARIO N6A 3K7
CANADA

NOTES

1 Recently much attention in the literature has been devoted to generalizations of the Bayesian model which give up coherence in favor of representing rational belief with indeterminate probabilities. (A few of the many available examples are: Kyburg, H.E., *Probability and the Logic of Rational Belief* Wesleyan University Press, 1961; Dempster, A.P., 'A Generalization of Bayesian Inference' *J. Roy. Statist. Soc. Ser.* B 30, 1968, 205-247; Shafer, G., 'A Theory of Statistical Evidence' in Harper-Hooker,

1976, 365-436; Levi, I. 'On Indeterminate Probabilities' *Journal of Philosophy* 71, 1974, 391-418; I.J. Good and P. Suppes are also prominent among those involved.) The present extension is not in conflict with these developments and can be adapted to such systems in a relatively straightforward manner. Such an adaptation together with a critical examination of the representation of rational belief with indeterminate probabilities will be developed in future work. In this paper restriction to coherent models simplifies the presentation and keeps it as close as possible to the orthodox model presented in the work of Ramsey and Savage.

2 When the set theoretical treatment of propositions is combined with the idea that belief is a functional state determined by its role in guiding decisions, as is the case in Savage's classic presentation of the orthodox model (Savage 1953), one can argue that semantical omniscience is not so much an attribution of super reasoning ability to the agent as it is a criterion for having E be an appropriate field to represent the objects of the agent's beliefs in the problem being modeled. Assent to sentences is only one among many kinds of evidence bearing on what propositions an agent accepts. A situation where the agent assents to a sentence S_0 that would standardly assert A and refuses to assent to some logically equivalent sentence S_1 that would standardly assert A is evidence that the agent does not use both these sentences to express the proposition A. It cannot be evidence that he both accepts A and does not accept A, because such a belief state is not possible. Whatever irrationality the agent is displaying will apply to his failure to understand the sentences in the standard way, not to his attitudes toward propositions. (For an interesting exploration of some of the problems and benefits in this approach to semantical omniscience see Robert Stalnaker 1976a, and 1976b).

3 If we replaced (4) with the stronger condition

$$(4') \; P(A)=1 \text{ iff } K \subseteq A$$

we would be demanding strict coherence relative to K. This would force the identification of acceptance with assigning degree of belief 1. Such an identification is very plausible when one thinks of probability primarily as degree of belief; however, the identification also rules out many standard measure theoretic based probability functions. I am not sure what side of this trade-off is more valuable; but, since a fully adequate development of the (4') version would require finding some way to substitute for all the neat applications of measure theory it rules out, I do not restrict the present extensions of the Bayesian model to versions where (4') holds. (This is an important difference between the present treatment and the related work in Harper 1975 and Harper 1976).

4 Even a strict early positivist who restricts evidence to sense data must allow that later sense data might give me grounds for doubting whether or not I really had certain other sense data in the past.

5 David Lewis (Lewis 1975, pp. 4, 5) uses this example to discredit a distinctive assumption of Stalnaker's theory of conditionals (Stalnaker 1968). The assumption is that non-trivial conditions for

"If it were that C then it would be that I"

depend on there being a unique nearest C-world to the actual world. The important difference that an acceptance context K(C) can leave many propositions undecided while a world W_C cannot is what blocks the force of the example as an objection to having a unique C-assumption context.

6 In 1973 Richard Jeffrey suggested to me in private conversation that it would be useful to provide for question openings. In public addresses at Pittsburgh and Boston, which I attended in 1975, Isaac Levi argued that question openings are an essential part of any rational revision of evidence.

REFERENCES

de Finetti, 1937. "Foresight: Its Logical Laws, Its Subjective Sources" available in Kyburg, H.E. and Smokler, H.E., 1964 pp. 93-158.

Harper, W.L., 1975. "Rational Belief Change, Popper Functions and Counterfactuals" *Synthese* 30 (1975), also reprinted in Harper W.L. and Hooker, C.A., 1976, Vol. I, pp. 73-113.

Harper, W.L., 1976. "Ramsey Test Conditionals and Iterated Belief Change" in Harper W.L. and Hooker, C.A., 1976, Vol. I, pp. 117-136.

Harper, W.L. and Hooker, C.A., (1976). *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, D. Reidel Pub. Co., Dordrecht.

Hockney, D., Harper, W.L., and Freed, B. (1975). *Contemporary Research in Philosophical Logic and Linguistic Semantics*, D. Reidel Pub. Co., Dordrecht.

Kyburg, H.E. and Smokler, H.E. (1964). *Studies in Subjective Probability*, John Wiley and Sons, New York.

Lewis, D.K. "Counterfactuals and Comparative Possibility" in Hockney, D., Harper, W. and Freed, B. (1975), pp. 1-29.

MacKay, A.F. and Merrill, D.D., 1976. *Issues in the Philosophy of Language*, Yale Press, New Haven.

Ramsey, F.P., 1926. "Truth and Probability" available in Kyburg, H.E. and Smokler, H.E., 1964, pp. 63-92.

Rényi, A., 1955. "On a New Axiomatic Theory of Probability", Acta Mathematica 6, pp. 285-335.

Savage, L.J., 1953. The Foundations of Statistics, John Wiley and Sons, New York.

Stalnaker, R., 1968. *"A Theory of Conditionals," Studies In Logical Theory (American Philosophical Quarterly*, Supplementary Monograph Series), Oxford.

Stalnaker, 1976a. "Propositions" in *Issues in the Philosophy of Language*, MacKay, A.F. and Merrill, D.D., 1976, pp. 79-91.

Stalnaker, 1976b. "Assertion" manuscript.

van Fraassen, 1976. "Representation of Conditional Probabilities" forthcoming in *Journal of Philosophical Logic*.

NOTES TO CONTRIBUTORS

Editorial correspondence should be addressed to Professor Asa Kasher, at the Department of Philosophy, Bar-Ilan University, Ramat-Gan, Israel.

All manuscripts submitted for publication should be typed in duplicate and double-spaced. Articles should be accompanied also by abstracts of approximately 100 words.

OFFPRINT order forms are sent out with proofs. Each author is entitled to receive gratis 20 offprints of his article. Authors are requested to consult with the editor before giving consent to the re-publication of their articles.

NOTES TO SUBSCRIBERS

Yearly subscription is $12.00 (US) for individuals and $16.00 (US) for institutions.

Write to:
Subscriptions, **PHILOSOPHIA**, Department of Philosophy, Bar-Ilan University, Ramat-Gan, Israel.

ADVERTISING

All particulars of advertising in the journal may be obtained from **PHILOSOPHIA**, Department of Philosophy, Bar-Ilan University, Ramat-Gan, Israel.

CONTENTS
VOLUME 7, No. 2

PROCEEDINGS OF THE FOURTH ANNUAL MEETING OF THE SOCIETY OF EXACT PHILOSOPHY