CARL G. WAGNER

# GENERALIZED PROBABILITY KINEMATICS*

ABSTRACT. Jeffrey conditionalization is generalized to the case in which new evidence bounds the possible revisions of a prior below by a Dempsterian lower probability. Classical probability kinematics arises within this generalization as the special case in which the evidentiary focal elements of the bounding lower probability are pairwise disjoint.

## 1. NEW EVIDENCE AND PARTIAL PROBABILITY

Let $p$ be a probability measure on the possibility set $\Theta$,[1] and suppose that new evidence suggests the desirability of revising $p$. Suppose that the total evidence determines a family $\mathscr{E}$ of nonempty, pairwise disjoint subsets of $\Theta$ and a collection $\{\mu_E : E \in \mathscr{E}\}$ of positive real numbers summing to 1, restricting the possible revisions of $p$ to those probability measures $q$ satisfying

(1) $\quad \forall E \in \mathscr{E}[q(E) = \mu_E].$

Unless each $E \in \mathscr{E}$ is a singleton, there are infinitely many probability measures $q$ satisfying (1). Suppose, however, that there is reason to judge that any acceptable revision $q$ should also satisfy

(2) $\quad \forall A \subseteq \Theta,\ \forall E \in \mathscr{E}\ [q(A/E) = p(A/E)].$[2]

As is easily proved, (1) and (2) jointly determine a uniquely acceptable revision $q$, defined for all $A \subseteq \Theta$ by

(3) $\quad q(A) = \sum_{E \in \mathscr{E}} \mu_E p(A/E),$

a generalization of ordinary conditionalization first explored by Jeffrey (1965).[3]

It is worth asking how, in the initial step of revising $p$ according to the above scheme, one might come to fix the values of a possible revision $q$ on members of the family $\mathscr{E}$. In an interesting class of cases, to be explored in depth in this paper, it appears that one arrives at restrictions of this type indirectly, first using the total evidence to construct a (fully specified) probability measure $u$ on a related possibil-

*Erkenntnis* **36**: 245–257, 1992.

ity set $\Omega$, and then inferring certain restrictions on $q$ from an entailment relation between *outcomes* in $\Omega$ and *events* in $\Theta$.

Jeffrey's famous mudrunner[4] example is a case in point. In that example, $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, with outcomes corresponding to a certain horse's winning on a muddy track ($\theta_1$), losing on a muddy track ($\theta_2$), winning on a dry track ($\theta_3$), and losing on a dry track ($\theta_4$). A gambler's prior $p$ is defined on all subsets of $\Theta$, including the (for him) important event $E = \{\theta_1, \theta_2\}$ that the track is muddy. New evidence now arrives in the form of a "fresh weather forecast (which) leads him to change his degree of belief"[5] in members of the family $\mathscr{E} = \{E, \bar{E}\}$ to the new values $\mu_E = 0.6$ and $\mu_{\bar{E}} = 0.4$.

Since weather forecasts assign probabilities to meteorological, not racetrack, conditions, this reassessment of the probabilities of $E$ and $\bar{E}$ clearly takes place indirectly, in something like the following way: The forecast assigns probability 0.6 to rain ($\omega_1$) and 0.4 to clear weather ($\omega_2$), and the gambler adopts these probabilities in the form of the probability measure $u$ on $\Omega = \{\omega_1, \omega_2\}$, with $u(\omega_1) = 0.6$ and $u(\omega_2) = 0.4$. Assuming that rain entails a muddy track and clear weather entails a dry track implies for any reasonable revision $q$ of $p$ that $q(E) \geqslant u(\omega_1)$ and $q(\bar{E}) \geqslant u(\omega_2)$, and the fact that $u(\omega_1) + u(\omega_2) = 1 = q(E) + q(\bar{E})$ then implies that $q(E) = u(\omega_1)$ and $q(\bar{E}) = u(\omega_2)$. Note that this analysis leaves the values of $q$ on other nonempty proper subsets of $\Theta$ undetermined.[6]

In the foregoing example the fact that $\mathscr{E}$ is a disjoint family makes the inference from $u$ on $\Omega$ to $q$ on $\mathscr{E}$ so trivial as to be almost unconscious. In the following section, however, we shall examine a generalization of this sort of evidentiary situation in which the events in a possibility set $\Theta$ entailed by the individual outcomes in a related possibility set $\Omega$ may be arbitrary nonempty subsets of $\Theta$. As we shall see, evidence of this sort circumscribes the possible revisions of a prior by placing a *lower bound* on the revised probability of each event $E \subseteq \Theta$.

## 2. NEW EVIDENCE AND LOWER PROBABILITY

Suppose that the total evidence relevant to revising some prior probability on $\Theta$ allows us to assess on a related possibility set $\Omega$ a probability measure $u$ which is positive on all nonempty subsets of $\Omega$. Suppose that our understanding of the relationship between outcomes in $\Omega$ and those in $\Theta$ is expressed by a function $\Gamma : \Omega \rightarrow 2^\Theta - \{\emptyset\}$, with $\Gamma(\omega)$ denoting the narrowest event in $\Theta$ presently understood to be entailed by the outcome $\omega \in \Omega$.[7] How do $u$ and $\Gamma$ constrain the possible revisions of the prior in question?

Following Dempster (1967), let us consider the mappings $m$, $b$, and $a$, defined for all $E \subseteq \Theta$ by

(4) $$m(E) = u\{\omega \in \Omega : \Gamma(\omega) = E\},$$

(5) $$b(E) = \sum_{H \subseteq E} m(H) = u\{\omega \in \Omega : \Gamma(\omega) \subseteq E\},$$

and

(6) $$a(E) = u\{\omega \in \Omega : \Gamma(\omega) \cap E \neq \emptyset\}.$$

Clearly, $m(\emptyset) = b(\emptyset) = a(\emptyset) = 0$, and

(7) $$a(\Theta) = b(\Theta) = \sum_{E \subseteq \Theta} m(E) = 1.$$

Since $\{\omega \in \Omega : \Gamma(\omega) \subseteq \bar{E}\}$ and $\{\omega \in \Omega : \Gamma(\omega) \cap E \neq \emptyset\}$ partition $\Omega$, it follows from (5) and (6) that

(8) $$\forall E \subseteq \Theta[a(E) = 1 - b(\bar{E})].$$

We shall call members of the family

(9) $$\mathscr{E} = \{E \subseteq \Theta : m(E) > 0\}$$

*evidentiary focal elements*. As one might expect, the mapping $b$ is a probability measure (and thus, by (8), identical with $a$) just when every evidentiary focal element is a singleton subset of $\Theta$ (Shafer, 1976, Theorem 2.8).

Since $b(E)$ (respectively, $a(E)$)is the sum of the probabilities of all those outcomes in $\Omega$ which entail (respectively, do not preclude) the event $E$, it is clear that $u$ and $\Gamma$ restrict the possible revisions of the prior in question to those probability measures $q$ satisfying

(10) $$\forall E \subseteq \Theta\,[b(E) \leq q(E) \leq a(E)].$$

It follows from (8) that it is redundant to postulate both the lower bound $b$ and the upper bound $a$ on $q$, and thus (10) may be replaced with the simpler, equivalent restriction

(11) $$\forall E \subseteq \Theta\,[b(E) \leq q(E)].$$

It should be noted that the set of probability measures $q$ satisfying

(11) is always nonempty. As indicated by the folllowing lemma, such measures arise by allocating each of the quantities $m(E)$ among elements of $E$ in arbitrary fashion.

LEMMA 1. *Given m and b, as defined by* (4) *and* (5), *a probability measure q satisfies* (11) *if and only if there exists a mapping* $\lambda: \Theta \times 2^{\Theta} \to [0, 1]$ *such that*

(i) $E \subseteq \Theta$ *and* $\theta \notin E \Rightarrow \lambda(\theta, E) = 0$,

(ii) $\sum_{\theta \in \Theta} \lambda(\theta, E) = 1$ *for all nonempty* $E \subseteq \Theta$, *and*

(iii) $q(\theta) = \sum_{E \subseteq \Theta} \lambda(\theta, E) m(E)$ *for every* $\theta \in \Theta$.

*Proof*: See Dempster (1967, pp. 327–330). ◇

Lemma 1 yields yet another construction of those $q$ satisfying (11), a construction which further supports the intuition that (11) is the appropriate embodiment of the evidence manifested in $u$ and $\Gamma$. Imagine the possibility of going beyond the specification of $u$ and $\Gamma$ and assessing a joint probability measure $Q$ on $\Theta \times \Omega$. Such a $Q$ would of course be *compatible with u and* $\Gamma$ in the sense that

(12) $\forall \theta \in \Theta, \forall \omega \in \Omega\, [\theta \notin \Gamma(\omega) \to Q(\theta, \omega) = 0]$

and

(13) $Q_{\Omega} = u$,

where $Q_{\Omega}$ denotes the marginalization of $Q$ to $\Omega$.[8] Were we possessed of such a $Q$ the solution to our revision problem would be at hand. We would simply update the prior in question to $Q_{\Theta}$, the marginalization of $Q$ to $\Theta$.[9] The following theorem is therefore of considerable interest.

THEOREM 1. *A probability measure q on* $\Theta$ *satisfies* (11), *with b defined by* (5), *if and only if there exists a probability measure Q on* $\Theta \times \Omega$ *satisfying* (12) *and* (13) *such that* $q = Q_{\Theta}$, *the marginalization of Q to* $\Theta$.

*Proof*: If $q = Q_{\Theta}$, where $Q$ satisfies (12) and (13), then for all $E \subseteq \Theta$,

$$q(E) = Q_{\Theta}(E) = \sum_{\omega \in \Omega} \sum_{\theta \in E} Q(\theta, \omega)$$

$$\geq \sum_{\substack{\omega\in\Omega\\ \Gamma(\omega)\subseteq E}} \sum_{\theta\in E} Q(\theta,\omega) = \sum_{\substack{\omega\in\Omega:\\ \Gamma(\omega)\subseteq E}} \sum_{\theta\in\Theta} Q(\theta,\omega)$$

$$= \sum_{\substack{\omega\in\Omega:\\ \Gamma(\omega)\subseteq E}} Q_\Omega(\omega) = \sum_{\substack{\omega\in\Omega:\\ \Gamma(\omega)\subseteq E}} u(\omega) = b(E).$$

Suppose, conversely, that $q$ satisfies (11). For all $\theta \in \Theta$ and $\omega \in \Omega$, let $Q(\theta, \omega) = \lambda(\theta, \Gamma(\omega))u(\omega)$, where $\lambda$ satisfies (i), (ii), and (iii) of Lemma 1. It is easy to show that $Q$ satisfies (12) and (13) and that $Q_\Theta = q$. ◇

Theorem 1 establishes that, from the standpoint of $u$ and $\Gamma$, the possible revisions of some prior may be construed either as probability measures on $\Theta$ bounded below by $b$, or as marginalizations to $\Theta$ of probability measures on $\Theta \times \Omega$ compatible with $u$ and $\Gamma$. We shall find the latter characterization particularly useful in the next section.

We conclude this section by identifying the special case of our $(u, \Gamma)$–evidentiary model previously treated by Jeffrey.

THEOREM 2. *If* $\mathscr{E} = \{E \subseteq \Theta : m(E) > 0\}$ *is a pairwise disjoint family and* $q$ *is a probability measure on* $\Theta$, *then* $q(E) \geq b(E)$ *for all* $E \subseteq \Theta$ *if and only if* $q(E) = m(E)$ *for all* $E \in \mathscr{E}$.

*Proof*: If $q(E) \geq b(E)$ for all $E \subseteq \Theta$, then, since $b(E) \geq m(E)$ by (5), it follows from (7), the pairwise disjointness of $\mathscr{E}$, and the additivity of $q$, that $1 = \Sigma_{E\in\mathscr{E}}\, m(E) \leq \Sigma_{E\in\mathscr{E}}\, q(E) \leq 1$, and hence that $q(E) = m(E)$ for all $E \in \mathscr{E}$. (Note that if $E \in \mathscr{E}$, no proper subset of $E$ has positive $m$-measure, and so, for such $E$, $m(E) = b(E)$.)

Suppose, on the other hand, that $q(E) = m(E)$ for all $E \in \mathscr{E}$. Given an arbitrary $E \subseteq \Theta$, let $\mathscr{H} = \{H \in \mathscr{E} : H \subseteq E\}$. Clearly, $q(E) \geq \Sigma_{H\in\mathscr{H}}\, q(H) = \Sigma_{H\in\mathscr{H}}\, m(H) = \Sigma_{H\subseteq E}\, m(H) = b(E)$. ◇

Theorem 2 establishes that when the family $\mathscr{E}$ of evidentiary focal elements is pairwise disjoint (equivalently, when for all $\omega_1, \omega_2 \in \Omega$, either $\Gamma(\omega_1) = \Gamma(\omega_2)$ or $\Gamma(\omega_1) \cap \Gamma(\omega_2) = \emptyset$) the "lower probability" restriction (11) is equivalent to Jeffrey's "partial probability" restriction (1), with $\mu_E = m(E)$. In the next section we proffer generalizations of Jeffrey's condition (2) and formula (3).

## 3. A NATURAL GENERALIZATION OF JEFFREY CONDITIONALIZATION

With respect to the task of identifying a uniquely acceptable revision of a prior $p$ on $\Theta$, the restrictions delineated in the previous section, arising solely from consideration of the probability measure $u$ on $\Omega$ and the entailment mapping $\Gamma \rightarrow 2^{\Theta} - \{\emptyset\}$, are rarely decisive. To arrive at such a revision we require an additional condition, generalizing Jeffrey's (2), specifying the extent to which the uncertainty assessments incorporated in the heretofore ignored prior $p$ are still judged to be operative.

It is useful at this point to introduce some additional notation. For all $E \subseteq \Theta$, let $E_{\star} = \{\omega \in \Omega : \Gamma(\omega) = E\}$, so that, for example, $m(E) = u(E_{\star})$. For all $A \subseteq \Theta$ and all $B \subseteq \Omega$, let "$A$" $= A \times \Omega$ and "$B$" $= \Theta \times B$, so that, for example, $Q(\text{“}A\text{”}) = Q_{\Theta}(A)$ for all $A \subseteq \Theta$ and $Q(\text{“}B\text{”}) = Q_{\Omega}(B)$ for all $B \subseteq \Omega$. As in the previous section, $\mathscr{E} = \{E \subseteq \Theta : m(E) > 0\}$, the family of evidentiary focal elements.

Suppose now that we avail ourselves of Theorem 1, construing the possible revisions of $p$ as marginalizations to $\Theta$ of probability measures $Q$ on $\Theta \times \Omega$ satisfying (12) and (13). And suppose that, upon reflection, we judge that only those $Q$ satisfying the additional condition

(14) $\quad \forall A \subseteq \Theta, \ \forall E \in \mathscr{E}[Q(\text{“}A\text{”})/\text{“}E_{\star}\text{”}) = p(A/E)]$[10]

would be reasonable candidates for marginalization. To adopt (14) is to judge that the total impact of the occurrence of the event $E_{\star}$ is to preclude the occurrence of any outcome $\theta \notin E$, and that, within $E$, $p$ remains operative in the assessment of relative uncertainties.

Conditions (12), (13), and (14) may well admit of an infinite number[11] of probability measures on $\Theta \times \Omega$, but, happily for our enterprise, their marginalizations to $\Theta$ are identical.

THEOREM 3. *If $Q$ is any probability measure on $\Theta \times \Omega$ satisfying* (12), (13) *and* (14) *and $Q_{\Theta}$ is the marginalization of $Q$ to $\Theta$, then $Q_{\Theta} = q$, where*

(15) $$q(A) = \sum_{E \in \mathscr{E}} m(E)p(A/E)$$

*for all $A \subseteq \Theta$.*

*Proof*: The family $\{\text{“}E_{\star}\text{”} : E \in \mathscr{E}\}$ is a partition of $\Theta \times \Omega$, and $Q(\text{“}E_{\star}\text{”}) = Q_{\Omega}(E_{\star}) = u(E_{\star}) = m(E)$ for all $E \in \mathscr{E}$. Hence, for all

$A \subseteq \Theta$, $\quad Q_\Theta(A) = Q(\text{"}A\text{"}) = \Sigma_{E \in \mathscr{E}} Q(\text{"}E_\star\text{"} \cap \text{"}A\text{"}) = \Sigma_{E \in \mathscr{E}} Q(\text{"}E_\star\text{"}) Q(\text{"}A\text{"}/\text{"}E_\star\text{"}) = \Sigma_{E \in \mathscr{E}} m(E) p(A/E)$. $\diamond$

It is to be emphasized that probability measures on $\Theta \times \Omega$ enter the preceding discussion as formal, conceptual tools, there being no presupposition (and no need, given Theorem 3) that a fully defined probability measure on $\Theta \times \Omega$ be attainable. Indeed, formula (15) involves only uncertainty measures on $\Theta$, the prior $p$, and the measure $m$ induced by $u$ and $\Gamma$ according to (4).

Along with Theorem 1, Theorem 3 establishes that restricting the possible revisions of a prior $p$ to those probability measures satisfying (11), and thus formally realizable as the marginalization to $\Theta$ of some $Q$ satisfying (12) and (13), determines, under the additional restriction (14), a unique revision $q$ defined by (15). The striking resemblance of (15) to Jeffrey's updating formula (3), along with Theorem 2, establishing the equivalence of (11) with Jeffrey's (1) when $\mathscr{E}$ is a disjoint family, indicates the naturalness of our proffered generalization of Jeffrey conditionalization. The argument for this naturalness is completed by the following elaboration of Theorem 2, which establishes the equivalence of (14) with Jeffrey's (2), when evidentiary focal elements are pairwise disjoint.

THEOREM 4. *If* $\mathscr{E}$, *as defined by* (9), *is a pairwise disjoint family, then stipulating of a probability measure* $q$ *on* $\Theta$ *that* $q = Q_\Theta$ *for some probability measure* $Q$ *on* $\Theta \times \Omega$ *satisfying* (12), (13) *and* (14) *is equivalent to stipulating that* $q(E) = m(E)$ *for all* $E \in \mathscr{E}$ *and that* $q(A/E) = p(A/E)$ *for all* $A \subseteq \Theta$ *and all* $E \in \mathscr{E}$.

*Proof*: We note first that the pairwise disjointness of $\mathscr{E}$ implies for any $Q$ satisfying (12), for all $A \subseteq \Theta$, and for all $E \in \mathscr{E}$ that

$$(16) \qquad \begin{aligned} Q(\text{"}A\text{"} \cap \text{"}E_\star\text{"}) &= Q(A \times E_\star) = Q((A \cap E) \times E_\star) \\ &= Q((A \cap E) \times \Omega) = Q_\Theta(A \cap E) \end{aligned}$$

In particular, setting $A = E$ in (16) yields

$$(17) \qquad Q_\Theta(E) = Q(E \times E_\star) = Q(\Theta \times E_\star) = Q(\text{"}E_\star\text{"}).$$

From (16) and (17), it follows for all $Q$ satisfying (12), for all $A \subseteq \Theta$, and for all $E \in \mathscr{E}$, that

$$(18) \qquad Q(\text{"}A\text{"}/\text{"}E_\star\text{"}) = Q_\Theta(A/E).$$

Now if $q = Q_\Theta$, where $Q$ satisfies (12), (13), and (14), (18) and (14) imply that $q(A/E) = p(A/E)$ for all $A \in \Theta$ and all $E \in \mathscr{E}$. By Theorem 1, (12) and (13) imply (11), which, along with Theorem 2, implies $q(E) = m(E)$ for all $E \in \mathscr{E}$.

Suppose, on the other hand, that $q(E) = m(E)$ for all $E \in \mathscr{E}$ and that $q(A/E) = p(A/E)$ for all $A \subseteq \Theta$ and all $E \in \mathscr{E}$. The former condition, along with Theorem 2, implies (11), which, with Theorem 1, implies the existence of some $Q$ satisfying (12) and (13) such that $Q_\Theta = q$. That $Q$ satisfies (14) follows from (18) and the equality $q(A/E) = p(A/E)$. ◇

## 4. EXAMPLES

*The Linguist.* You encounter a native of a certain foreign country and wonder whether he is a Catholic northerner ($\theta_1$), a Catholic southerner ($\theta_2$), a Protestant northerner ($\theta_3$), or a Protestant southerner ($\theta_4$). Your prior probability $p$ over these possibilities (based, say, on population statistics and the judgment that it is reasonable to regard this individual as a random representative of his country) is given by $p(\theta_1) = 0.2$, $p(\theta_2) = 0.3$, $p(\theta_3) = 0.4$, and $p(\theta_4) = 0.1$. The individual now utters a phrase in his native tongue which, due to the aural similarity of the phrases in question, might be a traditional Catholic piety ($\omega_1$), an epithet uncomplimentary to Protestants ($\omega_2$), an innocuous southern regionalism ($\omega_3$), or a slang expression used throughout the country in question ($\omega_4$). After reflecting on the matter you assign subjective probabilities $u(\omega_1)=0.4$, $u(\omega_2)=0.3$, $u(\omega_3)=0.2$, and $u(\omega_4)=0.1$ to these alternatives. In the light of this new evidence how should you revise $p$?

Under the assumption that no Protestant would utter $\omega_1$ or $\omega_2$ and no northerner $\omega_3$, the entailment mapping $\Gamma$ from members of $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4\}$, to subsets of $\Theta = \{\theta_1, \theta_2, \theta_3, \theta_4\}$, is given by $\Gamma(\omega_1) = \Gamma(\omega_2) = \{\theta_1, \theta_2\}$, $\Gamma(\omega_3) = \{\theta_2, \theta_4\}$, and $\Gamma(\omega_4) = \Theta$. From (4) it follows that the nonzero values of $m$ are $m\{\theta_1, \theta_2\} = 0.7$, $m\{\theta_2, \theta_4\} = 0.2$, and $m(\Theta) = 0.1$, and from (5) it follows that the nonzero values of $b$ are $b\{\theta_1, \theta_2\} = b\{\theta_1, \theta_2, \theta_3\} = 0.7$, $b\{\theta_1, \theta_2, \theta_4\} = 0.9$, $b\{\theta_2, \theta_4\} = b\{\theta_2, \theta_3, \theta_4\} = 0.2$ and of course $b(\Theta) = 1$. These, along with the zero values of $b$, establish a lower bound on the possible revisions of $p$.

To decide to further circumscribe such revisions by (14) amounts in this case to judging that the class of utterances $\{\omega_1, \omega_2\}$, taken as a whole, provides no information (differing from assessments incorporated in $p$) with respect to the speaker's geographical origins, that $\omega_3$ is similarly uninformative with respect to the speaker's religion, and that $\omega_3$ is similarly uninformative with respect to any nonempty proper subset of $\Theta$. If you so judge,[12] Theorem 3 will then warrant your updating $p$ to $q$, as defined by (15). As an illustration, $p\{\theta_1, \theta_3\} = 0.6$, the prior probability that the speaker is a northerner, will be revised to $q\{\theta_1, \theta_3\} = (0.7)(0.4) + (0.2)(0) + (0.1)(0.6) = 0.34$.

*The Toxic Waste Dump*. A small dump is found on the premises of an abandoned chemical factory. This factory is known to have produced toxic wastes of type $\theta_1$, $\theta_2$, and $\theta_3$, with $p(\theta_1) = 0.4$, $p(\theta_2) = 0.3$, and $p(\theta_3) = 0.3$ being the proportions of the total mass of wastes attributable to the three types. Without aditional evidence one might adopt the quantities $p(\theta_i)$ as estimates of the proportions of these types of wastes in the dump.

Suppose that we now learn that the dump resulted from shipments $\omega_1$, $\omega_2$, $\omega_3$, and $\omega_4$, with the proportions of the total mass of the dump attributable to these shipments given by $u(\omega_1) = 0.1$, $u(\omega_2) = 0.2$, $u(\omega_3) = 0.5$, and $u(\omega_4) = 0.2$. The specific composition of these shipments is unknown, but fragmentary records indicate that shipments $\omega_1$ and $\omega_2$ contained no wastes of type $\theta_1$, and shipment $\omega_3$ contained no wastes of type $\theta_3$.

Along with $\Gamma(\omega_1) = \Gamma(\omega_2) = \{\theta_2, \theta_3\}$, $\Gamma(\omega_3) = \{\theta_1, \theta_2\}$, and $\Gamma(\omega_4) = \Theta = \{\theta_1, \theta_2, \theta_3\}$, $u$ yields by (4) the function $m$ on $2^\Theta$ with nonzero values $m(\{\theta_2, \theta_3\}) = 0.3$, $m(\{\theta_1, \theta_2\}) = 0.5$, and $m(\Theta) = 0.2$. The associated set function $b$, defined by (5), takes the nonzero values $b(\{\theta_2, \theta_3\}) = 0.3$, $b(\{\theta_1, \theta_2\}) = 0.5$, and $b(\Theta) = 1$. Note that for every $E \subseteq \Theta$, $b(E)$ is a lower bound on the proportion of the total mass of the dump comprised of wastes in the class $E$.

To judge in this case that (14) holds is to judge that *within the aggregate of shipments* $\omega_1$ *and* $\omega_2$ the wastes of type $\theta_2$ and $\theta_3$ may reasonably be assumed to be represented in proportion to the quantities $p(\theta_2)$ and $p(\theta_3)$[13] and, similarly, that within shipment $\omega_3$ wastes of type $\theta_1$ and $\theta_2$ are represented in proportion to the quantities $p(\theta_1)$ and $p(\theta_2)$, and within shipment $\omega_4$ wastes of type $\theta_i$ are represented in proportion to the quantities $p(\theta_i)$ for $i = 1, \ldots, 3$. So judging, one is warranted in employing (15) to construct an updated estimate $q(\theta_i)$ of the fraction of the total mass of the dump comprised of wastes of type $\theta_i$. For example $q(\theta_3) = (0.3)(0.5) + (0.5)(0) + (0.2)(0.3) = 0.21$.

## 5. DISCUSSION

Theorems 1 and 3, elaborated by Theorems 2 and 4, constitute a natural generalization of Jeffrey conditionalization to a class of evidentiary situations in which evidentiary focal elements need not be, as in Jeffrey's model, disjoint. The problem of updating a prior probability in the light of such evidence is of practical, as well as theoretical, importance. (We encountered it in the process of designing a diagnostic expert system.) Are there feasible alternatives to our approach to this problem?

*Dempster's Rule*. Students of the theory of belief functions (of which the mapping $b$, defined by (5), is an example) may wonder whether one could not simply update $p$ to $p \oplus b$, the result of combining $p$ and $b$ by Dempster's rule (see Shafer, 1976, Chapter 3). As is easily seen, however, $p \oplus b$, while always a probability measure on $\Theta$, may well fail to be bounded below by $b$, in violation of the basic restrictive condition (11). Moreover, $p$ and $b$ are not reasonably combined by Dempster's rule, since they do not satisfy Shafer's criterion of being based on "entirely distinct bodies of evidence,"[14] $p$ based on the old evidence, and $b$ on the total evidence, old as well as new.

It is perhaps of mathematical interest to note, however, that one can always *formally* reconstruct formula (15) by means of Dempster's rule.

THEOREM 5. *Suppose that $p$ is a probability measure on $\Theta$, that $m: 2^\Theta \to [0,1]$ such that $m(\phi) = 0$ and $\Sigma_{E \subseteq \Theta} m(E) = 1$, and that $p(E) > 0$ for all $E \in \mathscr{E} = \{E \subseteq \Theta : m(E) > 0\}$. If we define $M(E) = 0$ for all $E \notin \mathscr{E}$ and*

$$(19) \qquad M(E) = \frac{m(E)/p(E)}{\sum_{F \in \mathscr{E}} m(F)/p(F)}, \quad \text{for all } E \in \mathscr{E},$$

*and let $B(A) = \Sigma_{E \subseteq A} M(E)$ be the belief function induced by $M$, then, for all $A \subseteq \Theta$,*

$$(20) \qquad q(A) \underset{\text{def}}{=} \sum_{E \in \mathscr{E}} m(E) p(A \mid E) = p \oplus B(A),$$

*where* $p \oplus B$ *denotes the result of combining* $p$ *and* $B$ *by Dempster's rule.*

*Proof*: Let $a \in A$. By Dempster's rule (see Shafer 1976, pp. 59–60),

$$p \oplus B(a) = p(a) \sum_{\substack{E \in \mathscr{E} \\ a \in E}} M(E) \bigg/ \sum_{\theta \in \Theta} p(\theta) \sum_{\substack{E \in \mathscr{E} \\ \theta \in E}} M(E)$$

$$= p(a) \sum_{\substack{E \in \mathscr{E} \\ a \in E}} \frac{m(E)}{p(E)} \bigg/ \sum_{\theta \in \Theta} p(\theta) \sum_{\substack{E \in \mathscr{E} \\ \theta \in E}} \frac{m(E)}{p(E)}$$

$$= q(a) \bigg/ \sum_{\theta \in \Theta} q(\theta) = q(a).$$

Since $q$ and $p \oplus B$ are additive, (20) follows for all $A \subseteq \Theta$. ◇

Note that when $\mathscr{E}$ is a pairwise disjoint family the above method furnishes a particularly simple formal reconstruction of Jeffrey's rule by means of Dempster's rule (cf., Shafer 1981).

*Maxent.* Students of maximum entropy approaches to probability revision may recall that the probability measure defined by Jeffrey's formula (3) minimizes the Kullback–Leibler information number $I(q, p) = \Sigma_\theta \, q(\theta) \log(q(\theta)/p(\theta))$ over all probability measures $q$ satisfying (1), and wonder if the probability measure defined by our formula (15) similarly minimizes $I(q, p)$ over all probability measures $q$ bounded below by $b$. The answer is negative, as shown by just about any case in which it happens that the prior $p$ is itself bounded below by $b$. Convinced by Skyrms (1987), among others, that MAXENT is not a tenable updating rule, we are undisturbed by this fact. Indeed we take it as additional evidence against MAXENT that (15), firmly grounded on evidence establishing the lower bound $b$ and a considered judgement that (14) holds, might violate MAXENT.[15] From this point of view, the fact that Jeffrey's rule coincides with MAXENT is simply a misleading fluke, put in its proper perspective by the natural generalization of Jeffrey conditionalization described in this paper.

That what Diaconis and Zabell (1982) call "mechanical updating" schemes fail to ground the revision formula (15) is fortunate from another perspective as well. For without the temptations of mechanical updating, we are left with the hard work of judging whether the crucial condition (14), which generalizes Jeffrey's (2), is reasonably assumed. And it is worth being reminded that conditionalization needs to be grounded both on new evidence *and* on a judgement about the continued relevance of our prior assessments of uncertainty.

## NOTES

* Research supported by grants from the University of Tennessee and the National Science Foundation (DIR – 8921269). Portions of this work were completed while the author was a Visiting Fellow in the Department of Philosophy at Princeton University.

[1] In this paper all possibility sets are assumed to be finite. A probability measure on $\Theta$ is a mapping $p: 2^{\Theta} \to [0, 1]$, where $2^{\Theta}$ is the set of all subsets of $\Theta$, such that $p(\phi) = 0$, $p(\Theta) = 1$, and $\forall A, B \subseteq \Theta\ [A \cap B = \phi \to p(A \cup B) = p(A) + p(B)]$. When $\theta \in \Theta$, we shall abuse notation and write $p(\theta)$ rather than the correct, but cumbersome, $p(\{\theta\})$. As usual, elements of $\Theta$ are called *outcomes* and subsets of $\Theta$ are called *events*. An event is said to occur if one of its constituent outcomes occurs.

[2] It is implicit in (2) that, in the words of Van Fraassen (1980), "zeros are not raised" on members of $\mathscr{E}$, i.e., that $p(E) > 0$ for all $E \in \mathscr{E}$.

[3] Strictly speaking (3) represents a mild generalization of Jeffrey's original rule, which requires $\mathscr{E}$ to be a partition of $\Theta$. This stipulation has the slightly odd consequence that ordinary conditionalization is not a special case, but only a limiting case, of Jeffrey conditionalization. Formula (3), on the other hand, is a genuine generalization of ordinary conditionalization, reducing to the latter when $\mathscr{E}$ has just one member.

[4] Jeffrey, 1965, pp. 158–59.

[5] Jeffrey, 1965, p. 158.

[6] However, lower bounds on the values of $q$ on subsets other than $E$ and $\bar{E}$ are derivable, it being one aim of this paper to elaborate on just this observation. See, especially, §2.

[7] To assert that the event $E = \{\theta_1, \ldots, \theta_k\} \subseteq \Theta$ is entailed by the outcome $\omega \in \Omega$ is to assert that "$\omega$ occurs" entails "$\theta_1$ occurs $\vee \cdots \vee$ $\theta_k$ occurs." If $\omega$ entails $E$ and $E \subseteq F$, then $\omega$ also entails $F$. The event $\Gamma(\omega)$ may be equivalently described as the set of those $\theta \in \Theta$ which are not inconsistent with $\omega$, it being assumed that no $\omega \in \Omega$ is inconsistent with every $\theta \in \Theta$, i.e., that every $\omega \in \Omega$ entails $\Theta$.

[8] For all $E \subseteq \Omega$, $Q_{\Omega}(E) =_{\text{def}} Q(\Theta \times E)$. The set of $Q$ satisfying (12) and (13) is clearly nonempty.

[9] For all $E \subseteq \Theta$, $Q_{\Theta}(E) =_{\text{def}} Q(E \times \Omega)$.

[10] It is implicit in (14) that "zeros are not raised" on $\mathscr{E}$, i.e., that $p(E) > 0$ for every evidentiary focal element $E$.

[11] Of course, (12), (13), and (14) are always satisfied by at least one probability measure $Q$ on $\Theta \times \Omega$, namely, the measure $Q$ defined by $Q(\theta, \omega) = u(\omega)p(\theta|\Gamma(\omega))$. If, for all $\omega_1$, $\omega_2 \in \Omega$, $\omega_1 \neq \omega_2$ implies that $\Gamma(\omega_1) \neq \Gamma(\omega_2)$, then this is the only probability measure $Q$ satisfying (12), (13), and (14). For in this case, (14) asserts that for all $A \subseteq \Theta$ and all $\omega \in \Omega$, $Q(\text{“}A\text{”}/\text{“}\{\omega\}\text{”}) = p(A/\Gamma(\omega))$. Along with (12) and (13), this implies that $Q(\theta, \omega) = u(\omega)p(\theta|\Gamma(\omega))$.

[12] You may, of course, not so judge. You may, for example, have evidence that utterers of $\omega_3$ are overwhelmingly Protestant.

[13] This condition will of course be met if in each of the separate shipments $\omega_1$ and $\omega_2$ wastes of type $\theta_2$ and $\theta_3$ are represented in proportion to the quantities $p(\theta_2)$ and $p(\theta_3)$. But an assumption of this strength is not necessary in order to support (14).

[14] Shafer, 1976, p. 57.

[15] Of course the probability measure defined by (15) does minimize $I(q, p)$ over all those $q$ which are marginalizations to $\Theta$ of some $Q$ on $\Theta \times \Omega$ that satisfies (12) and (13) (this being equivalent to $b \leq q$) *and in addition satisfies* (14). For by Theorem 3 there is only one such $q$. But the counterpart of getting Jeffrey's (3) from (1) and the minimization of $I(q, p)$ alone, on which MAXENT enthusiasts pride themselves, would clearly be to get (15) from (11) (equivalently, (12) and (13)) and the minimization of $I(q, p)$ alone, ignoring (14) just as they ignore Jeffrey's (2). And this they cannot do.

## REFERENCES


Dempster, A. P.: 1967, 'Upper and Lower Probabilities Induced by a Multivalued Mapping', *Annals of Mathematical Statistics* **38**, 325–339.

Diaconis, P. and Zabell, S.: 1982, 'Updating Subjective Probability', *Journal of the American Statistical Association* **77**, 822–830.

Jeffrey, R.: 1965, *The Logic of Decision*, McGraw-Hill, New York.

Shafer, G.: 1976, *A Mathematical Theory of Evidence*, Princeton University Press, Princeton.

Shafer, G.: 1981, 'Jeffrey's Rule of Conditioning', *Philosophy of Science* **48**, 337–362.

Skyrms, B. 1981, 'Updating, Supposing, and Maxent', *Theory and Decision* **22**, 225–246.

Van Fraassen, B. C.: 1980, 'Rational Belief and Probability Kinematics', *Philosophy of Science* **47**, 165–187



Manuscript submitted date November 15, 1989
Final version received date October 10, 1990



Mathematics Dept.
University of Tennessee
Knoxville, TN 37996
U.S.A.