GLENN SHAFER

A SUBJECTIVE INTERPRETATION OF CONDITIONAL PROBABILITY*

This paper shows how the rule of conditioning for subjective probabilities can itself be given a purely subjective interpretation. This interpretation is based on the assumption that a person has subjective probabilities for how his information and probabilities may change over time. This means we are concerned not with how the person should or will change his beliefs, but rather with what he believes about how these beliefs will change. We assume that the person expects to have additive expectations that are consistent over time, and that each state of nature he considers possible specifies how his knowledge will progress and his expectations will change over time. And we deduce that he expects to have additive subjective probabilities that will change over time in conformity with the rule of conditioning.

The first step of the argument is developed in Section 1 below. There I show that if each state of nature specfies what the person will know at time t, then these specifications form an "information partition". In other words, they specify a disjoint partition of the set of states of nature and specify that the person's knowledge at time t will amount to knowledge of which element of the partition contains the true state of nature. We are accustomed to assuming the existence of information partitions. We implicitly do so, for example, when we assume that a person's information will come from the observation of a random quantity. We are less accustomed to the idea that the existence of information partitions can be deduced from simpler assumptions. So far as I know, no argument similar to the one given in Section 1 has been published previously.

The argument for the rule of conditioning is given in Section 3. This argument can be seen as a version of the argument that Thomas Bayes gave for the third proposition in his famous essay on probability. In Shafer (1982) I gave a different version of Bayes's argument. There I used rooted trees to specify the possible ways a person's knowledge may progress. And I followed Bayes in limiting attention to a simple kind of "expectation": a contract that awards a fixed prize contingent on the happening of a given

Journal of Philosophical Logic 12 (1983) 453–466. 0022–3611/83/0124–0453$01.40.

event. The present paper uses an explicit time scale in place of rooted trees and uses the modern language of expectation for random quantities. When we use this modern language to express Bayes's assumption that the value of a contract is unchanged if the prize it awards is replaced by a prize that has the same value if and when the event on which the contract depends happens (Postulate IV in Section 4.1 of Shafer (1982)), we see that this assumption is equivalent to a form of the rule of iterated expectation (see Axioms III and III′ in Section 3 below). The implications of an essentially identical form of the rule of iterated expectation have been explored by Goldstein (1981).

I conclude the paper by showing how the framework developed here provides a natural setting for the theorems of Aumann (1976) and Geanakoplos and Polemarchakis (1982), which assert that two people who begin with a common probability distribution but then receive different information will regain their agreement on the probability of any particular event if they repeatedly tell each other their current probabilities for that event.

The argument of the present paper does not address directly the question of whether a person should change his beliefs in conformity with the rule of conditioning when he acquires new information. But it does enable us to say that if he knows in advance that he will acquire the information or its denial, then he will be violating his own expectations if he fails to change his beliefs in this way. For him to have a policy of doing otherwise would be "incoherent" in an obvious sense. This point has been made in the vocabulary of betting by Teller (1973). The present paper strengthens the point by showing how naturally conditioning emerges once foresight about the possibilities for new information is brought into the subjective probability model.

When we see how naturally conditioning emerges from the assumption of foresight about new information, it is natural to ask whether conditioning is justified without this assumption. A number of examples have been advanced to suggest that it is not. These include puzzles where the choice of a conditional probability depends on the protocol for the acquisition of information (see Bar-Hillel and Falk (1982) and Faber (1976)) and statistical problems where assumptions must be made concerning why given data is present or missing (Dawid and Dickey (1977)).

1. THE EXISTENCE OF INFORMATION PARTITIONS

Let Ω be the set of states of nature that Gracchus, at time 0, considers possible. The elements of Ω are descriptions of the world, in some degree of detail, and Gracchus believes that exactly one of these descriptions is true. Let us call subsets of Ω *events*.

Let us say that an event A is *Gracchus's knowledge* at time t (relative to Ω) if the following two conditions are satisfied at time t: (i) Gracchus is certain that the true state of nature is in A, and (ii) there is no proper subset B of A such that Gracchus is certain that the true state of nature is in B. The words "certain" and "knowledge" are used here in a purely subjective sense. When we say that Gracchus knows something, we mean only that he fully believes it; when we say that he is certain, we mean he is subjectively certain. He may be wrong. Let us, however, impose some order on Gracchus's "knowledge" by assuming that it is consistent and closed under logical implication: if Gracchus knows certain things at time t, then he immediately deduces any logical consequences and hence also knows these logical consequences at time t.

A set $\mathscr{G}$ of subsets of Ω is called a *partition* of Ω if the elements of $\mathscr{G}$ are non-empty and disjoint and have Ω as their union. A partition $\mathscr{G}$ of Ω is said to be *Gracchus's information partition* for time t, where $t \geqslant 0$, if each element ω of Ω specifies that Gracchus's knowledge at time t will be the element of $\mathscr{G}$ that contains ω. In general, an information partition for Gracchus at time t need not exist, for the elements of Ω may fail to say anything about Gracchus's knowledge at time t. But, as the following proposition tells us, if each element of Ω does specify what Gracchus's knowledge at time t will be, then our assumptions imply that these specifications determine an information partition.

PROPOSITION 1. Suppose, for a particular $t \geqslant 0$, that each element ω of Ω specifies a subset $G_t(\omega)$ of Ω containing ω and specifies that $G_t(\omega)$ will be Gracchus's knowledge at time t. Then the $G_t(\omega)$ form a partition of Ω.

Proof. That the $G_t(\omega)$ are non-empty and have Ω as their union follows from the fact that $\omega \in G_t(\omega)$. So in order to show that they form a partition we need only show that they are disjoint – i.e., if $G_t(\omega_1) \cap G_t(\omega_2) \neq \emptyset$, then $G_t(\omega_1) = G_t(\omega_2)$.

Denote by $\mathbf{G}_t$ the mapping that maps ω to $G_t(\omega)$. (Throughout this paper bold-face type will indicate implicit dependence on ω.) This mapping, since it is determined by Ω, is known to Gracchus from time 0 on. So if his knowledge at time t turns out to be A, then he can immediately deduce and hence will know at time t that the true state of nature must be some ω such that $G_t(\omega) = A$ – i.e., must be in the set $\{\omega | G_t(\omega) = A\} = \mathbf{G}_t^{-1}(A)$. In order for this not to contradict the statement that his knowledge is only A, we must have $A \subset \mathbf{G}_t^{-1}(A)$. So $G_t(\omega) \subset \mathbf{G}_t^{-1}(G_t(\omega))$ for all $\omega \in \Omega$. This means that if $\omega' \in G_t(\omega)$, then $G_t(\omega') = G_t(\omega)$.

Now consider a pair ω_1 and ω_2 such that $G_t(\omega_1) \cap G_t(\omega_2) \neq 0$. Choose an element ω' of $G_t(\omega_1) \cap G_t(\omega_2)$. Since $\omega' \in G_t(\omega_1)$, $G_t(\omega') = G_t(\omega_1)$. And since $\omega' \in G_t(\omega_2)$, $G_t(\omega') = G_t(\omega_2)$. So $G_t(\omega_1) = G_t(\omega_2)$. □

Throughout the remainder of this paper, let us assume that Gracchus has an information partition for every $t \geqslant 0$. Denote this partition by $\mathscr{G}_t$, and denote by $G_t(\omega)$ the element of $\mathscr{G}_t$ that ω says will be Gracchus's knowledge at time t. Following Shafer (1982), let us call an event A *exact* if $A \in \mathscr{G}_t$ for some t. Not all events are exact; if Ω is large most are not exact.

The argument for Proposition 1 establishes, of course, only the subjective existence of an information partition for Gracchus at time t; the conclusion of the argument is that Gracchus *believes* his knowledge at time t will be an element of the partition, and this belief may be false. It is not clear, however, what would be meant by the objective existence of an information partition. From an objective viewpoint, whether Gracchus is right in believing his knowledge at time t will be an element of the partition formed by the $G_t(\omega)$ comes down to whether his knowledge at time t is $G_t(\omega_0)$, where ω_0 is the true state of nature; it is difficult to say whether Gracchus is right or wrong to believe, in the case of an ω that is not the true state of nature, that his knowledge will be $G_t(\omega)$ if ω is the true state of nature, and so it is difficult to say whether he is right or wrong to believe that the possibilities for what his knowledge will be form a partition.

2. PARTITIONS AND RANDOM QUANTITIES

Here I establish a vocabulary and notation for discussing partitions and random quantities.

Suppose $\mathscr{A}_1$ and $\mathscr{A}_2$ are partitions of Ω, and for each $\omega \in \Omega$, let $A_i(\omega)$

denote the unique element of $\mathscr{A}_i$ that contains ω. If $A_1(\omega) \subset A_2(\omega)$ for all ω, let us say that $\mathscr{A}_1$ is a *refinement* of $\mathscr{A}_2$ and $\mathscr{A}_2$ is a *coarsening* of $\mathscr{A}_1$. (We write $A \subset B$ when A is a subset of B, proper or improper. Thus this definition allows us to call $\mathscr{A}_1$ a refinement of $\mathscr{A}_2$ when $\mathscr{A}_1 = \mathscr{A}_2$.) The *meet* of $\mathscr{A}_1$ and $\mathscr{A}_2$, denoted $\mathscr{A}_1 \wedge \mathscr{A}_2$, is their finest common coarsening, the partition whose element containing ω is

$$\{\omega' \mid \text{There exists a sequence } A_1, A_2, \ldots, A_k \text{ of elements of } \mathscr{A}_1 \cup \mathscr{A}_2 \text{ such that } \omega \in A_1, \omega' \in A_k, \text{ and } A_i \cap A_{i+1} \neq \emptyset \text{ for } i = 1, \ldots, k-1\}.$$

Their *join*, denoted $\mathscr{A}_1 \vee \mathscr{A}_2$, is their coarsest common refinement, the partition whose element containing ω is $A_1(\omega) \cap A_2(\omega)$.

If $0 < t' < t$, then $\mathscr{G}_t$ is a refinement of $\mathscr{G}_{t'}$. (We assume that Gracchus does not expect to forget anything). If $\mathscr{G}_t = \mathscr{G}_{t'}$, then we say that Gracchus expects no new information from time t' through time t. If $\mathscr{G}_t = \bigvee_{t'<t} \mathscr{G}_t$ – i.e., $G_t(\omega) = \bigcap_{t'<t} G_{t'}(\omega)$ – then we say that Gracchus expects no new information at time t.

Let us call a bounded real-valued function of Ω a *random quantity*. In accordance with the convention that bold-face type indicates an implicit dependence on Ω, let us use bold-face for random quantities but not for their values; the value of the random quantity $\mathbf{X}$ at ω will be denoted $X(\omega)$. Given a subset A of Ω, denote by $\mathbf{1}_A$ the random quantity whose values are

$$1_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{if } \omega \notin A. \end{cases}$$

We may abbreviate $\mathbf{1}_\Omega$ to $\mathbf{1}$, and we may similarly write $\mathbf{0}$ for $\mathbf{1}_\emptyset$.

The *partition of Ω determined by the random quantity* $\mathbf{X}$ is the partition consisting of all non-empty subsets of Ω of the form $\{\omega \mid X(\omega) = s\} = \mathbf{X}^{-1}(\{s\})$, where s is a real number. If this partition is coarser than a given partition $\mathscr{A}$, let us call $\mathbf{X}$ *measurable* with respect to $\mathscr{A}$. Notice that $\mathbf{1}_A$ is measurable with respect to $\mathscr{A}$ if and only if A is a union of elements of $\mathscr{A}$. If $\mathbf{X}$ is measurable with respect to $\mathscr{G}_t$ – i.e., if for every $G \in \mathscr{G}_t$, X is constant on G, then we say that Gracchus expects (at time 0) to know the value of $\mathbf{X}$ at time t.

3. EXPECTATION AND PROBABILITY

I now formulate, within the framework of the preceding section, the ideas of expectation and probability. The approach is to adopt axioms governing Gracchus's expectations and to deduce the properties of his probabilities from these axioms.

Let us continue to assume that Gracchus has an information partition $\mathscr{G}_t$ for every $t \geqslant 0$. And assume that $\mathscr{G}_0 = \{\Omega\}$ and that there exists a time $T > 0$ such that $\mathscr{G}_T$ consists of singletons. At time 0, Gracchus cannot identify a proper subset of Ω that contains the true state of nature, but by time T he expects to know exactly which element of Ω it is.

Let us assume that each element ω of Ω specifies, for each random quantity $\mathbf{X}$ and each $t \in [0, T]$, a real number $E_{t,\omega}(\mathbf{X})$, called *Gracchus's expectation* for $\mathbf{X}$ at time t. Let us denote by $\mathbf{E}_t(\mathbf{X})$ the random quantity whose value at ω is $E_{t,\omega}(\mathbf{X})$. And let us assume that Gracchus expects (at time 0) to know the value of $\mathbf{E}_t(\mathbf{X})$ at time t. This means that if $G \in \mathscr{G}_t$ and $\omega, \omega' \in G$, then $E_{t,\omega}(\mathbf{X}) = E_{t,\omega'}(\mathbf{X})$.

If $G \in \mathscr{G}_t$, let us write $E_{t,G}(\mathbf{X})$ for the common value of $E_{t,\omega}(\mathbf{X})$ for $\omega \in G$. And let us write $E_0(\mathbf{X})$ for $E_{0,\Omega}(\mathbf{X})$.

The assumption that Gracchus expects to know the value of $\mathbf{E}_t(\mathbf{X})$ at time t is based on the intuitive idea that this value is an aspect of his opinion at time t; it is the value he assigns at time t to a contract that will pay him $\mathbf{X}$ at time T – i.e., a contract requiring that he be paid $\mathbf{X}$ at time T if $\mathbf{X} > 0$ but requiring that he pay $-\mathbf{X}$ at time T if $\mathbf{X} < 0$. This intuitive interpretation also leads us to assume that Gracchus's expectations obey the following axioms for all $t \in [0, T]$ and all random quantities $\mathbf{X}$ and $\mathbf{Y}$.

I. $\quad E_{t,\omega}(\mathbf{X}+\mathbf{Y}) = E_{t,\omega}(\mathbf{X}) + E_{t,\omega}(\mathbf{Y}) \quad \text{for all } \omega \in \Omega.$

II. $\quad \inf_{\omega' \in G_t(\omega)} X(\omega') \leqslant E_{t,\omega}(X) \leqslant \sup_{\omega' \in G_t(\omega)} X(\omega') \quad \text{for all } \omega \in \Omega.$

III. If $E_{t,\omega}(\mathbf{X}) = E_{t,\omega}(\mathbf{Y})$ for all $\omega \in \Omega$, and if $0 \leqslant t' \leqslant t$, then $E_{t',\omega}(X) = E_{t',\omega}(\mathbf{Y})$ for all $\omega \in \Omega$.

Axioms I and II are the usual axioms for expectation; see, for example, p. 74 of de Finetti (1974). The following proposition lists some of the consequences of these two axioms.

PROPOSITION 2. Suppose $t \in [0, T]$, $G \in \mathscr{G}_t$, a is a real number, and **X** and **Y** are random quantities. Then

(i) $\mathbf{E}_t(\mathbf{X}+\mathbf{Y}) = \mathbf{E}_t(\mathbf{X}) + \mathbf{E}_t(\mathbf{Y})$;

(ii) $\mathbf{E}_t(\mathbf{0}) = \mathbf{0}$;

(iii) $\mathbf{E}_t(\mathbf{1}) = \mathbf{1}$;

(iv) $\mathbf{E}_t(\mathbf{1}_G) = \mathbf{1}_G$;

(v) $\mathbf{E}_t(\mathbf{1}_G\mathbf{X}) = \mathbf{1}_G E_{t,G}(\mathbf{X})$;

(vi) $\mathbf{E}_t(a\mathbf{X}) = a\mathbf{E}_t(\mathbf{X})$.

Proof. Statement (i) is merely another way of writing Axiom I. Statements (ii), (iii), and (iv) follow immediately from Axiom II. To prove (v), notice first that by Axiom I,

$$E_{t,\omega}(\mathbf{X}) = E_{t,\omega}(\mathbf{1}_G\mathbf{X}) + E_{t,\omega}(\mathbf{1}_{\bar{G}}\mathbf{X}),$$

where $\bar{G} = \Omega - G$. By Axiom II, $E_{t,\omega}(\mathbf{1}_{\bar{G}}\mathbf{X}) = 0$ if $\omega \in G$ and $E_{t,\omega}(\mathbf{1}_G\mathbf{X}) = 0$ if $\omega \notin G$. So

$$E_{t,\omega}(\mathbf{1}_G\mathbf{X}) = \begin{Bmatrix} E_{t,\omega}(\mathbf{X}) = E_{t,G}(\mathbf{X}) & \text{if } \omega \in G \\ 0 & \text{if } \omega \notin G \end{Bmatrix} = \mathbf{1}_G(\omega)E_{t,G}(\mathbf{X}).$$

For a proof of (vi), see p. 75 of de Finetti (1974). □

Axiom III says, in effect, that Gracchus's expectations at a given time are consistent with his knowledge about his later expectations; if he knows that he will value **X** and **Y** equally at a future time, then he will value them equally now. This axiom is not as familiar as Axioms I and II; in fact, I have not been able to find anything similar to it in the literature. There is a sense, however, in which this axiom goes back to Thomas Bayes, for something very similar to it is implicit in Bayes's argument for the third proposition in his famous essay on probability. (Compare the axiom to Postulate IV in 4.1 of Shafer (1982).)

Given Axiom II, Axiom III is equivalent to the following statement.

III′ If $0 \leqslant t' \leqslant t$, then $\mathbf{E}_{t'}(\mathbf{X}) = \mathbf{E}_{t'}(\mathbf{E}_t(\mathbf{X}))$.

This statement is more familiar than III; we recognize it as a version of the rule of iterated expectation. If we set $t' = 0$, then III′ reduces to the simpler expression $E_0(\mathbf{X}) = E_0(\mathbf{E}_t(\mathbf{X}))$, which can be compared with Goldstein's (1981) formula $P(X) \Rightarrow P(P_D(X))$. (Goldstein uses P for "prevision" instead of E for "expectation" and his D, for "data", represents a point in time marked not by a clock but rather by the completion of a data-gathering effort.)

Let us now define probability: the quantity $E_{t,\omega}(1_A)$ is called *Gracchus's probability* at time t. When we are thinking about probability, let us write $P_{t,\omega}(A)$ for $E_{t,\omega}(1_A)$, $\mathbf{P}_t(A)$ for $\mathbf{E}_t(1_A)$, $P_{t,G}(A)$ for $E_{t,G}(1_A)$, and $P_0(A)$ for $E_0(1_A)$.

PROPOSITION 3. Suppose $0 \leqslant t' \leqslant t \leqslant T$, $G \in \mathscr{G}_t$, and A and B are events. Then

(i) if $A \cap B = \emptyset$, then $\mathbf{P}_t(A \cup B) = \mathbf{P}_t(A) + \mathbf{P}_t(B)$;

(ii) $\mathbf{P}_t(\emptyset) = 0$;

(iii) $\mathbf{P}_t(\Omega) = 1$;

(iv) $\mathbf{P}_t(G) = 1_G$;

(v) $\mathbf{P}_t(G \cap A) = 1_G P_{t,G}(A)$;

(vi) $\mathbf{P}_{t'}(A) = \mathbf{E}_{t'}(\mathbf{P}_t(A))$;

(vii) $\mathbf{P}_{t'}(G \cap A) = \mathbf{P}_{t'}(G) P_{t,G}(A)$;

(viii) if $G \in \mathscr{G}_{t'}$, then $P_{t',G}(A) = P_{t,G}(A)$.

Proof. Statements (i)–(v) follow immediately from the corresponding statements in Proposition 2. Statement (vi) follows from III′. To prove (vii), use (v) and (vi), together with (vi) of Proposition 2:

$$\begin{aligned} \mathbf{P}_{t'}(G \cap A) &= \mathbf{E}_{t'}(\mathbf{P}_t(G \cap A)) = \mathbf{E}_{t'}(1_G P_{t,G}(A)) \\ &= \mathbf{E}_{t'}(1_G) P_{t,G}(A) = \mathbf{P}_{t'}(G) P_{t,G}(A). \end{aligned}$$

To prove (viii), notice that (iv) and (v), applied to t', yield $\mathbf{P}_{t'}(G) = 1_G$ and $\mathbf{P}_{t'}(G \cap A) = 1_G P_{t',G}(A)$. Substituting these expression for $\mathbf{P}_{t'}(G)$ and $\mathbf{P}_{t'}(G \cap A)$ in (vii), we obtain $1_G P_{t',G}(A) = 1_G P_{t,G}(A)$. Since $G \neq \emptyset$, it follows that $P_{t',G}(A) = P_{t,G}(A)$.

Statement (viii) of Proposition 3 tells us that if Gracchus knows he will recieve no new information from time t' through time t, then $\mathbf{P}_t(A) = \mathbf{P}_{t'}(A)$. It also tells us that if G is an exact event, then the probability

$P_{t,G}(A)$ will be the same for any t such that $G \in \mathcal{G}_t$. It is legitimate, therefore, to call $P_{t,G}(A)$ the "conditional probability of A given G", omitting reference to the time t.

Statement (vii) of Proposition 3 is a version of the third proposition in Thomas Bayes's famous essay on probability; see Shafer (1982). If we set $t' = 0$, the statement becomes

$$P_0(G \cap A) = P_0(G)P_{t,G}(A),$$

and if $P_0(G) > 0$, this can be written

$$P_{t,G}(A) = \frac{P_0(G \cap A)}{P_0(G)}, \qquad (*)$$

in agreement with the usual "definition" of conditional probability.

Notice, however, that $(*)$ justifies the usual formula $P(A|B) = P(A \cap B)/P(B)$ only in the case where B is an "exact event". In order to justify this formula, we have built into our probability model assumptions about what the possibilities for new information are. This approach gives meaning and justification to "conditioning on B" only when B has been designated as one of these possibilities, and not every subset B of Ω can be so designated. □

4. DISCRETE PARTITIONS

Let us call a partition $\mathscr{A}$ *discrete* if $P_0(A) > 0$ for all $A \in \mathscr{A}$ and $\Sigma\{P_0(A)|A \in \mathscr{A}\} = 1$. A discrete partition can have only countably many elements. Moreover, probabilities are "countably addivite" and "continuous" when we are dealing with discrete partitions; if $\mathscr{A}$ is discrete and B is an event, then $P_0(B) = \Sigma\{P_0(A \cap B|A \in \mathscr{A}\}$ and $P_0(B \cap (\cap_i A_i)) = \lim_{i\to\infty} P_0(B \cap A_i)$ whenever $A_1 \supset A_2 \supset \ldots$ is a decreasing sequence of events each of which is measurable with respect to $\mathscr{A}$. A coarsening of a discrete partition is also discrete.

PROPOSITION 4. Suppose $\mathcal{G}_t$ is discrete. Then

(i) $$\mathbf{P}_t(A) = \frac{P_0(\mathbf{G}_t \cap A)}{P_0(\mathbf{G}_t)} \quad \text{for all } A \subset \Omega;$$

(ii) if Gracchus expects no new information at time t, then

$$\mathbf{P}_t(A) = \lim_{t' \uparrow t} \mathbf{P}_{t'}(A) \text{ for all } A \subset \Omega.$$

Proof. Since $\mathscr{G}_t$ is discrete, the random quantity $P_0(\mathbf{G}_t)$ is always positive. Hence (i) follows from (*). If $0 < t' < t$, then $\mathscr{G}_{t'}$, being a coarsening of $\mathscr{G}_t$, is also discrete. If Gracchus expects no new information at time t, then $G_t(\omega) = \bigcap_{t'<t} G_{t'}(\omega)$. It follows from the discreteness of $\mathscr{G}_t$, moreover, that $G_{t'}(\omega)$, which decreases as t' increases, changes only a countable number of times as t' approaches t from below. Hence the intersection $\bigcap_{t'<t} G_{t'}(\omega)$ can be thought of as the intersection of a countable sequence of $G_{t'}(\omega)$, with t' approaching t from below. So

$$P_{t,\omega}(A) = \frac{P_0(G_t(\omega) \cap A)}{P_0(G_t(\omega))} = \frac{\lim_{t' \uparrow t} P_0(G_{t'}(\omega) \cap A)}{\lim_{t' \uparrow t} P_0(G_{t'}(\omega))}$$

$$= \lim_{t' \uparrow t} P_{t',\omega}(A).$$

The statements in Proposition 4 are not, of course, assertions about what Gracchus's probability for A will be at time t. They are merely assertions about what Gracchus thinks, at time 0, about his probability for A at time t. Our theory is founded on assumptions about the states of nature Gracchus considers possible at time 0, and hence is limited to conclusions about what he thinks at time 0. On the other hand, there is a sense in which it is not possible for Gracchus to adopt a probability for A at time t that does not obey statement (i) of Proposition 4. For the correctness of this statement is specified by each state of nature in Ω. In order to violate the statement Gracchus must reformulate his set Ω of possible states of nature, and in doing so he will destroy the identity of the event A. □

5. RENEWAL OF AGREEMENT THROUGH EXCHANGE OF NEW OPINIONS

Here I use the framework developed in the preceding sections to study a problem first raised by Aumann (1976).

Suppose Gracchus and Maevius agree, at time 0, on a space Ω of possible states of nature. Suppose information partitions exist for both Gracchus and Maevius at all times $t \geqslant 0$, and denote these partitions by $\mathscr{G}_t$ and $\mathscr{M}_t$. Suppose also that each element ω of Ω specifices expectations for both Gracchus and Maevius that satisfy the assumptions of Section 2, and that at time 0 Gracchus and Maevius have the same expectations and hence the same probabilities. Since their information partitions differ, they do not necessarily expect to have the same expections and probabilities at later times. We will use the lower case letters g and m as superscripts to distinguish Gracchus's expectations and probabilities at time t from Maevius's.

PROPOSITION 5 (Aumann, 1976). Suppose the partition $\mathscr{G}_t \vee \mathscr{M}_t$ is discrete. And suppose Gracchus and Maevius expect (at time 0) to know each other's probabilities for A at time t. Then $\mathbf{P}_t^g(A) = \mathbf{P}_t^m(A)$.

Proof. To say that Gracchus expects to know $\mathbf{P}_t^m(A)$ at time t is to say that $\mathbf{P}_t^m(A)$ is measurable with respect to $\mathscr{G}_t$. Since it is also measurable with respect to $\mathscr{M}_t$, this implies it is measurable with respect to $\mathscr{G}_t \wedge \mathscr{M}_t$. Our assumption that Maevius expects to know $\mathbf{P}_t^g(A)$ at time t similarly implies that $\mathbf{P}_t^g(A)$ is measurable with respect to $\mathscr{G}_t \wedge \mathscr{M}_t$.

Fix an element ω of Ω and let R denote the element of $\mathscr{G}_t \wedge \mathscr{M}_t$ that contains ω. Since $\mathbf{P}_t^g(A)$ is measurable with respect to $\mathscr{G}_t \wedge \mathscr{M}_t$, $P_{t,G}^g(A) = P_{t,\omega}^g(A)$ for all elements G of $\mathscr{G}_t$ that are contained in R. It follows, by (vii) of Proposition 3, that

$$P_0(G \cap A) = P_0(G)P_{t,\omega}^g(A)$$

for all G in $\mathscr{G}_t$ that are contained in R. Summing this expression over all these G, and using the fact that $\mathscr{G}_t$ is discrete, we obtain

$$P_0(R \cap A) = P_0(R)P_{t,\omega}^g(A).$$

Reasoning similarly concerning $\mathbf{P}_t^m(A)$, we obtain

$$P_0(R \cap A) = P_0(R)P_{t,\omega}^m(A).$$

Since $P_0(R) > 0$, we may conclude from these two expressions that $P_{t,\omega}^g(A) = P_{t,\omega}^m(A)$. □

Aumann's formulation of the preceding proposition involves the notion of "common knowledge." According to Aumann's definition, an event A is *common knowledge at* ω (at time t) if $A \supset G_t(\omega) \cap M_t(\omega)$. Thus the value of a random quantity $\mathbf{X}$ is common knowledge at ω (at time t) if $\mathbf{X}$ is constant on $G_t(\omega) \cap M_t(\omega)$. Aumann shows that if the values of $\mathbf{P}_t^g$ and $\mathbf{P}_t^m$ are common knowledge at ω (at time t), then $\mathbf{P}_t^g(\omega) = \mathbf{P}_t^m(\omega)$. It is a corollary that if the values of $\mathbf{P}_t^g$ and $\mathbf{P}_t^m$ are common knowledge at all

ω (at time t), then $\mathbf{P}_t^g = \mathbf{P}_t^m$, and this is our Proposition 5. For to say that $\mathbf{P}_t^g$ and $\mathbf{P}_t^m$ are common knowledge at all ω (at time t) is to say that they are measurable with respect to $\mathscr{G}_t \wedge \mathscr{M}_t$ – i.e. (cf. the first paragraph of the above proof), that Gracchus and Maevius both expect to know both $\mathbf{P}_t^g$ and $\mathbf{P}_t^m$ at time t. The only shortcoming of Aumann's approach is the difficulty in understanding the intuitive justification for his definition of "common knowledge." (See Milgrom (1981).) Our approach, since it takes the evolution of further knowledge to be excplicitly specified in the probability model, allows us to use a much more transparent vocabulary.

Aumann's result was supplemented by Geanakoplos and Polemarchakis (1982), who showed how two people's probabilities for an event could become common knowledge as a result of their repeatedly telling each other their probabilities for that event. In order to formulate Geanakoplos and Polemarchakis's result in our framework, assume that Ω specifies a subset A of Ω and specifies that between time 1 and time 2 Gracchus and Maevius repeatedly tell each other their current probabilities for A. Suppose, for concreteness, that Ω specifies that information is to be communicated at each of an infinite sequence of times $t_1 < t_2 < \ldots$ such that $1 < t_1$ and $\lim_{i\to\infty} t_i = 2$. At each time t_i either Gracchus or Maevius is to tell the other his current probability for A, or both are to do so simultaneously. (When we say Gracchus tells Maevius his "current" probability for A, we mean, of course, that he tells him the value of $\mathbf{P}_t^g(A)$, where $t_{i-1} < t < t_i$.) The schedule of speakers is rigidly specified by Ω, and both Gracchus and Maevius appear on it an infinite number of times. It is also specified that Gracchus and Maevius will receive no other information from time 1 through time 2. This means, in particular, that both Gracchus and Maevius expect to receive no new information at time 2.

Let us assume that $\mathscr{G}_1 \vee \mathscr{M}_1$ is discrete.

PROPOSITION 6 (Geanakoplos and Polemarchakis, 1982). Under the preceding assumptions, $\mathbf{P}_2^g(A) = \mathbf{P}_2^m(A)$.

Proof. Given $\omega \in \Omega$, let $C(\omega)$ denote the set of all ω' that agree with ω as to what numbers will be announced by Gracchus and Maevius through time 2. The sets $C(\omega)$ form a partition of Ω, which we will denote by $\mathscr{C}$. (This partition can be thought of as the information partition at time 2 for a third person, say Caja, who begins at time 0 with the same set of states of nature as Gracchus and Maevius, but whose information through time 2 is limited to the announcements they make.)

Notice that $\mathscr{C}$ is a coarsening of $\mathscr{G}_1 \vee \mathscr{M}_1$. (Someone who had both Gracchus's and Maevius's knowledge at time 1 could predict their whole sequence of announcements.) And $\mathscr{G}_2 = \mathscr{G}_1 \vee \mathscr{C}$ and $\mathscr{M}_2 = \mathscr{M}_1 \vee \mathscr{C}$. (Gracchus, Maevius, and Caja all hear the same anouncements.)

Set $I = \{i \mid \text{Gracchus is scheduled to speak at time } t_{i+1}\}$. Since I is infinite, there are values of t_i arbitrarily close to 2 for $i \in I$. By statement (ii) of Proposition 4, $\mathrm{P}^g_{t_i}(A)$ approaches $\mathrm{P}^g_2(A)$ as these t_i are chosen closer to 2. But each of the $\mathrm{P}^g_t(A)$ is measurable with respect to $\mathscr{C}$. (Caja hears the value of $\mathrm{P}^g_{t_i}(A)$ at time t_{i+1} and so knows it at time 2.) So their limit, $\mathrm{P}^g_2(A)$, is measurable with respect to $\mathscr{C}$. (So Caja knows $\mathrm{P}^g_2(A)$ at time 2.) Since $\mathscr{C}$ is a coarsening of $\mathscr{M}_2 = \mathscr{M}_1 \vee \mathscr{C}$, $\mathrm{P}^g_2(A)$ is also measurable with respect to $\mathscr{M}_2$. So we conclude that Maevius knows $\mathrm{P}^g_2(A)$ at time 2 – or, more precisely, that he expects at time 0 to know it.

By the same argument applied to Maevius's announcements, Gracchus expects to know $\mathrm{P}^m_2(A)$ at time 2. So it follows from Proposition 5 that $\mathrm{P}^g_2(A) = \mathrm{P}^m_2(A)$. □

NOTE

* Research supported by NSF grant MCS 800213.

REFERENCES

Aumann, Robert J., 'Agreeing to disagree', *The Annals of Statistics* 6 (1976), 1236–1239.

Bar-Hillel, Maya, and Falk, Ruma, 'Some teasers concerning conditional probability', *Cognition* 11 (1982), 109–122.

Dawid, A. P. and Dickey, J. M., 'Likelihood and Bayesian inference from selectively reported data', *Journal of the American Statistical Association* 72 (1977), 845–850.

Faber, Roger J., 'Discussion: re-encountering a counter-intuitive probability', *Philosophy of Science* 43 (1976), 283–285.

de Finetti, Bruno, *Theory of Probability*, Vol. 1, Wiley, New York, 1974.

Geanakoplos, John D. and Polemarchakis, Heraklis M., 'We can't disagree forever', *Journal of Economic Theory* 28 (1982), 192–200.

Goldstein, Michael, 'Revising previsions: a geometric interpretation', *Journal of the Royal Statistical Society, Series B* 43 (1981) 105–130.

Milgrom, Paul, 'An axiomatic characterization of common knowledge', *Econometrica* 49 (1981) 219–222.

Shafer, Glenn, 'Bayes's two arguments for the rule of conditioning', *Annals of Statistics* 10 (1982) 1075–1089.
Teller, Paul, 'Conditionalization and observation', *Synthese* 26 (1973), 218–258.

Department of Mathematics,
University of Kansas,
Lawrence,
KS 66045, U.S.A.