Order $A \times A$ by the ordering of the lengths of chords joining pairs of points in A. Which axioms of an absolute-difference structure (Definition 8) are satisfied (proof or counterexample)? (4.10)

19. Show that the axioms of an absolute-difference structure (Definition 8) are all necessary for the representation of Theorem 6, except for the solvability axiom. (4.10)

20. Suppose that dissimilarities of the 15 pairs of six stimuli, a, b, c, d, e, f, are ranked (from 15, highest dissimilarity, to 1, lowest) as in the following matrix.

	a	b	c	d	e	f
a		4	7	8	14	15
b			1	5	11	13
c				$2\frac{1}{2}$	$9\frac{1}{2}$	12
d					6	$9\frac{1}{2}$
e						$2\frac{1}{2}$

Show that no absolute-difference representation is possible. Show that with only one inversion of order, a solution is possible. In the resulting representation, the cf distance must be more than three times as large as the cd distance. (4.10)

21. Suppose that Re^2 is ordered as follows:

$$(x, y) \succsim (x', y') \quad \text{iff} \quad |x - y| \geqslant |x' - y'|.$$

Does every pair of points in Re have a midpoint (Definition 11)? If so, what is it? Answer the same questions with respect to the lexicographic ordering of the plane (Exercise 6). (4.12)

22. Suppose that four subjects, s, t, u, v, exhibit the following preference orderings over objects a, b, c, d, where S is subject and R is rank.

S \ R	1	2	3	4
s	a	b	c	d
t	d	c	b	a
u	c	b	a	d
v	b	c	d	a

Show that no one-dimensional unfolding representation is possible. (4.12)

Chapter 5 Probability Representations

□ 5.1 INTRODUCTION

The debate about what probability is and about how probabilities shall be calculated has been prolonged and is, in many respects, unresolved; nonetheless few disagreements exist about the mathematical properties of numerical probability. It is accepted that probabilities are numbers between 0 and 1 and that they are attached to entities called events. Moreover, it is widely agreed that events in a given context are subsets—although not necessarily all subsets—of a set known as a sample space. Sample spaces are intended to represent all possible observations that one might make in particular situations. The classical assumptions about events are embodied in the following definition:

DEFINITION 1. *Suppose that X is a nonempty set (sample space) and that $\mathscr{E}$ is a nonempty family of subsets of X. Then $\mathscr{E}$ is an* algebra of sets *on X iff, for every $A, B \in \mathscr{E}$:*

1. $-A \in \mathscr{E}$.
2. $A \cup B \in \mathscr{E}$.

Furthermore, if $\mathscr{E}$ is closed under countable unions, i.e., whenever $A_i \in \mathscr{E}$, $i = 1, 2, \ldots$, it follows that $\bigcup_{i=1}^{\infty} A_i \in \mathscr{E}$, then $\mathscr{E}$ is called a σ-algebra *on X. The elements of $\mathscr{E}$ are called* events.

In words, to be called an algebra of sets, $\mathscr{E}$ must have the two properties of being closed under complementation and unions. As is shown in Section 5.3.1 (Lemma 1), it follows that $\mathscr{E}$ is also closed under intersections and differences.

Observe that $\{\varnothing, X\}$ is an algebra of sets and that it is a subalgebra of every algebra on X since if A is in $\mathscr{E}$, then $-A$ is in $\mathscr{E}$ and so $X = A \cup -A$ is in $\mathscr{E}$ and $\varnothing = -X$ is in $\mathscr{E}$. Only more complex algebras are usually of interest.

Later, in Section 5.4.1, we shall have occasion to consider a somewhat weaker definition in which we are assured that $A \cup B$ is in $\mathscr{E}$ only when $A \cap B = \varnothing$ and A, B are in $\mathscr{E}$.

The following axiomatic definition of numerical probability is the basis of most current work in probability theory. It was first stated explicitly by Kolmogorov (1933).

DEFINITION 2. *Suppose that X is a nonempty set, that $\mathscr{E}$ is an algebra of sets on X, and that P is a function from $\mathscr{E}$ into the real numbers. The triple $\langle X, \mathscr{E}, P\rangle$ is a* (finitely additive) probability space *iff, for every $A, B \in \mathscr{E}$:*

1. $P(A) \geqslant 0$.
2. $P(X) = 1$.
3. *If* $A \cap B = \varnothing$, *then* $P(A \cup B) = P(A) + P(B)$.

It is a countably additive probability space *if in addition:*

4. $\mathscr{E}$ *is a σ-algebra on* X.
5. *If* $A_i \in \mathscr{E}$ *and* $A_i \cap A_j = \varnothing$, $i \neq j$, $i, j = 1, 2, \ldots$, *then*

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

We do not study the surprising consequences of Definition 2; the interested reader should consult a good book on probability theory, e.g., Feller (1957, 1966). Our plan, instead, is to treat Definition 2 as (part of) a representation theorem; specifically, we inquire into conditions under which an ordering $\succsim$ of $\mathscr{E}$ has an order-preserving function P that satisfies Definition 2. Obviously, the ordering is to be interpreted empirically as meaning "qualitatively at least as probable as." Put another way, we shall attempt to treat the assignment of probabilities to events as a measurement problem of the same fundamental character as the measurement of, e.g., mass or momentum. From this point of view, the debates about the meaning of probability are, in reality, about acceptable empirical methods to determine $\succsim$. It is not evident why the measurement of probability should have been the focus of

more philosophic controversy than the measurement of mass, of length, or of any other scientifically significant attribute; but it has been. We are not suggesting that the controversies over probability have been unjustified, but merely that other controversial issues in the theory of measurement may have been neglected to a degree.

To those familiar with the debate about probability (see, for example, Carnap, 1950; de Finetti, 1937; Keynes, 1921; Nagel, 1939; Savage, 1954, 1961) these last remarks may seem slightly strange. Theories about the representation of a qualitative ordering of events have generally been classed as subjective (de Finetti, 1937), intuitive (Koopman, 1940a,b; 1941), or personal (Savage, 1954), with the intent of emphasizing that the ordering relation $\succsim$ *may* be peculiar to an individual and that he *may* determine it by any means at his disposal, including his personal judgment.[1] But these "mays" in no way preclude orderings that are determined by well-defined, public, and scientifically agreed upon procedures, such as counting relative frequencies under well-specified conditions. Even these objective procedures often contain elements of personal judgment; for example, counting the relative frequency of heads depends on our judgment that the events "heads on trial 1" and "heads on trial 2" are equiprobable. This equiprobability judgment is part of a partly "objective," partly "subjective" ordering of events.

Presumably, as science progresses, objective procedures will come to be developed in domains for which we now have little alternative but to accept the considered judgments of informed and experienced individuals.

Ellis pointed out that the development of a probability ordering is

> ... analogous to that of finding a thermometric property, which... was the first step towards devising a temperature scale.
>
> The comparison between probability and temperature may be illuminating in other ways. The first thermometers were useful mainly for comparing atmospheric temperatures. The air thermometers of the seventeenth century, for example, were not adaptable for comparing or measuring the temperatures of small solid objects. Consequently, in the early history of thermometry, there were many things which possessed temperature which could not be fitted into an objective temperature order. Similarly, then, we should not necessarily expect to find any single objective procedure capable of ordering all propositions in respect of probability, even if we assume that all propositions possess probability. Rather, we should expect there to be certain kinds of propositions that are much easier to fit into an objective probability order than others... (Ellis, 1966, p. 172).

Since virtually all representation theorems yield a unique probability measure, agreement about the probability ordering of an algebra of events

[1] Many of the relevant papers are collected together in Kyburg and Smokler (1964).

or, as is more usual, agreement about a method to determine that ordering is sufficient to define a unique numerical probability—an objective probability. Various authors have studied conditions under which two related probability measures on the same algebra of events, one thought of as objective and the other subjective, must agree. Perhaps the most interesting results are those included in Edwards (1962); also see Section 8.4.2.

In any case, the difficulties in measuring probability do not appear to be inherently different from those that arise when we apply extensive and other measurement methods elsewhere. In measuring lengths, for example, rods are practical only for relatively short distances; certainly other methods must be used in astronomical research, and considerable disagreement exists among astronomers about which of the alternatives is appropriate (for a general discussion, see Chapter 15 of North, 1965).

Besides the methods of ordering events by relative frequencies or by direct human judgments, there are indirect methods of inferring the ordering. Savage's (1954) book, *The Foundations of Statistics*, introduced an ordering of "personal probability" inferred from decisions among acts with uncertain outcomes. Briefly, if a, b are outcomes and A is an event, we denote by $a_A \cup b_{-A}$ an act whose outcome is a if A occurs and b if A does not occur. (The reason for the union notation will become clear in Chapter 8.) If a is preferred to b, and if the act $a_A \cup b_{-A}$ is preferred to the act $a_B \cup b_{-B}$, then we may infer that A is more probable than B (for the decision maker whose preferences are studied). This idea is discussed further in Section 5.2.4, where more complete references are given. Usually, the idea is introduced in conjunction with another: the measurement of both probability and utility such that acts are ordered by their expected utilities. This was Savage's approach, and we devote a separate chapter, 8, to a new version of it. A fine survey of work on this topic is available in the book by Fishburn (1970).

Our concern in this chapter is with conditions under which an ordering of events can be represented by a probability measure. Whether or not these conditions are satisfied by event orderings established by one or another experimental procedure is not discussed here; see Luce and Suppes (1965, pp. 321–327).

5.2 A REPRESENTATION BY UNCONDITIONAL PROBABILITY

5.2.1 Necessary Conditions: Qualitative Probability

Since we know the representation theorem that we want to prove, we may proceed as in Chapters 3 and 4 to derive necessary conditions on the assumption that the representation exists.

As always, if P is to be order preserving, $\succsim$ must be a weak ordering of $\mathscr{E}$. Since this property is familiar and its difficulties as an empirical law have already been discussed earlier (Section 1.3.1), albeit in other contexts, we need not repeat ourselves; slight rewording suits those comments to this context.

We next show that the certain event X is strictly more probable than the impossible event $\varnothing$, the empty set, and that any event A is at least as probable as $\varnothing$. Since, by Axiom 2 of Definition 2, $P(X) = 1$ and since $X \cup \varnothing = X$ and $X \cap \varnothing = \varnothing$, it follows from Axiom 3 that

$$1 = P(X) = P(X \cup \varnothing) = P(X) + P(\varnothing) = 1 + P(\varnothing),$$

and so $P(\varnothing) = 0$. Since P is order preserving and $P(X) = 1 > 0 = P(\varnothing)$, $X \succ \varnothing$. Moreover, by Axiom 1, $P(A) \geqslant 0 = P(\varnothing)$, so $A \succsim \varnothing$. Observe that $A \sim \varnothing$ does *not* imply $A = \varnothing$. That is to say, nonempty events may be equivalent in probability to the impossible event, and these *null events*, as they are called, correspond to sets with zero probability. We can also derive that $X \succsim A$ and a number of other properties, but it is unnecessary to state them separately because they turn out to be direct consequences of the properties that we list (see Section 5.3.1).

Next, we wish to consider how the order relation and the operation of union relate. Since Axiom 3 of Definition 2 is our main tool and since it is conditional on the two events being disjoint, we may anticipate disjointness playing a role in the corresponding qualitative property. Consider three events A, B, C, with A disjoint from B and C. If $B \succsim C$, then by the order-preserving property and Axiom 3,

$$P(A \cup B) = P(A) + P(B) \geqslant P(A) + P(C) = P(A \cup C),$$

whence $A \cup B \succsim A \cup C$. Conversely, if $A \cup B \succsim A \cup C$,

$$P(A) + P(B) = P(A \cup B) \geqslant P(A \cup C) = P(A) + P(C),$$

and so subtracting $P(A)$, $P(B) \geqslant P(C)$, whence $B \succsim C$. Summarizing, if $A \cap B = A \cap C = \varnothing$, then $B \succsim C$ if and only if $A \cup B \succsim A \cup C$. Intuitively, this condition seems plausible. If B is at least as probable as C, then adjoining to each the same disjoint event should not alter the ordering, and conversely, deleting such an event also should not alter it.

The primary relation between complementation and ordering is: If $A \succsim B$, then $-B \succsim -A$. This is not stated as a separate axiom because it follows from the others (Lemma 4, Section 5.3.1).

The final necessary condition that we derive in this section is an Archimedean property. We first derive a special case of it. Suppose that $A_1, \ldots, A_i, \ldots$

are mutually disjoint events (i.e., for $i \neq j$, $A_i \cap A_j = \varnothing$), each of which is equivalent in probability to some fixed event A, i.e., for $i = 1, 2, \ldots$, $A_i \sim A$. By a finite induction on Axiom 2 of Definition 1, $\bigcup_{i=1}^{n} A_i$ is an event. If a probability measure exists, then by a finite induction on Axiom 3 of Definition 2,

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) = nP(A).$$

If $P(A) > 0$, i.e., if $A > \varnothing$, then since $P \leqslant 1$ it follows that n cannot exceed $1/P(A)$. Thus, there can be at most finitely many mutually disjoint events each of which is equivalent in probability to a nonnull event. Actually, we need a slightly stronger form of this property which is formulated as follows:

DEFINITION 3. *Suppose that $\mathscr{E}$ is an algebra of sets and that $\sim$ is an equivalence relation on $\mathscr{E}$. A sequence $A_1, \ldots, A_i, \ldots$, where $A_i \in \mathscr{E}$, is a* standard sequence relative to $A \in \mathscr{E}$ *iff for $i = 1, 2, \ldots$, there exist B_i, $C_i \in \mathscr{E}$ such that:*

(i) $A_1 = B_1$ *and* $B_1 \sim A$;
(ii) $B_i \cap C_i = \varnothing$;
(iii) $B_i \sim A_i$;
(iv) $C_i \sim A$;
(v) $A_{i+1} = B_i \cup C_i$.

We see that if event A_i is assumed to correspond to i disjoint copies of A, then it follows that A_{i+1}—which equals the disjoint union of B_i, which in turn is equivalent to A_i, and of C_i, which is equivalent to A—corresponds to $i + 1$ disjoint copies of A. So this is a sensible inductive definition which generalizes our original notion. Since by induction $P(A_i) = iP(A)$, we must assume the finiteness of any standard sequence relative to an $A > \varnothing$. Note that our usual formulation of the Archimedean property, namely that every strictly bounded standard sequence be finite, is somewhat simplified here since every standard sequence is bounded by X.

We summarize these properties as:

DEFINITION 4. *Suppose that X is a nonempty set, that $\mathscr{E}$ is an algebra of sets on X, and that $\succsim$ is a relation on $\mathscr{E}$. The triple $\langle X, \mathscr{E}, \succsim\rangle$ is a* structure of qualitative probability *iff for every $A, B, C \in \mathscr{E}$:*

1. *$\langle \mathscr{E}, \succsim\rangle$ is a weak order.*
2. *$X > \varnothing$ and $A \succsim \varnothing$.*
3. *Suppose that $A \cap B = A \cap C = \varnothing$. Then $B \succsim C$ iff $A \cup B \succsim A \cup C$.*

Furthermore, the structure is Archimedean *iff*

4. *For every $A > \varnothing$, any standard sequence relative to A is finite.*

5.2.2 The Nonsufficiency of Qualitative Probability

The question we now ask and answer in the negative is: Does every (finite) structure of qualitative probability have a finitely additive, order-preserving probability representation? The following ingenious counterexample is due to Kraft, Pratt, and Seidenberg (1959). Let $X = \{a, b, c, d, e\}$ and let $\mathscr{E}$ be all subsets of X. Consider any order for which

$$\{a\} > \{b, c\}, \qquad \{c, d\} > \{a, b\} \qquad \text{and} \qquad \{b, e\} > \{a, c\}. \tag{1}$$

If an order-preserving representation P exists, then

$$P(a) > P(b) + P(c), \qquad P(c) + P(d) > P(a) + P(b),$$
$$P(b) + P(e) > P(a) + P(c).$$

Adding these three inequalities and canceling $P(a) + P(b) + P(c)$, we obtain $P(d) + P(e) > P(a) + P(b) + P(c)$, whence $\{d, e\} > \{a, b, c\}$. Therefore, if we can construct an order that simultaneously satisfies inequality (1), $\{a, b, c\} > \{d, e\}$, and the axioms of qualitative probability, then we know that this order cannot possibly be represented by a numerical probability. The trick is to find a way to avoid the tedious task of exhaustively verifying Axiom 3 for the example.

Suppose we can find a probability measure P on X such that inequality (1) holds (and necessarily $\{d, e\} > \{a, b, c\}$) and such that all subsets A of X other than $\{d, e\}$ and $\{a, b, c\}$ are either strictly more probable than $\{d, e\}$ or strictly less probable than $\{a, b, c\}$, i.e., not$[P(\{d, e\}) \geqslant P(A) \geqslant P(\{a, b, c\})]$. Of course the ordering induced by P automatically satisfies the Axioms of Definition 4. Now, change that order only to the extent of changing $\{d, e\} > \{a, b, c\}$ to $\{a, b, c\} > \{d, e\}$. Since no set lies between these two, their relations to all other sets are unaffected by the inversion, and so the new ordering, which cannot possibly have a probability representation, satisfies Definition 4. Therefore, we need only construct a measure P for which inequality (1) holds and no event separates $\{d, e\}$ and $\{a, b, c\}$.

Let $0 < \epsilon < \frac{1}{3}$, and ignoring the normalizing factor $P(X) = 16 - 3\epsilon$, let

$$P(a) = 4 - \epsilon, \qquad P(b) = 1 - \epsilon, \qquad P(c) = 3 - \epsilon,$$
$$P(d) = 2, \qquad P(e) = 6.$$

Using additivity, it is easy to verify that inequality (1) holds, that $P(\{d, e\}) = 8$, and that $P(\{a, b, c\}) = 8 - 3\epsilon$. Since $3\epsilon < 1$, the only possibility for finding a distinct set A such that $P(\{d, e\}) \geqslant P(A) \geqslant P(\{a, b, c\})$ is for $P(A)$ to be of the form $8 - i\epsilon$, $i = 0, 1, 2, 3$. However, the number 8 can be expressed as a sum of distinct elements from $\{1, 2, 3, 4, 6\}$ in only two ways, $1 + 3 + 4$ or $2 + 6$. These correspond to $\{a, b, c\}$ and $\{d, e\}$, respectively. Thus, no A distinct from those two sets can have probability between 7 and 8 inclusive. This completes the example.

So further properties are needed in order to construct a representation.

5.2.3 Sufficient Conditions

Quite a number of conditions are known which, along with the properties of qualitative probability (Definition 4), are each sufficient to prove the existence of an additive probability representation. As would be expected, each condition postulates the existence of events with certain (strong) properties, and therefore limits the algebras of sets for which the theory establishes a representation. We present three conditions here which lead to unique, finitely additive measures, another in Section 5.4.2 which leads to a unique, countably additive measure, and a fifth in Section 5.4.3 which leads to a unique measure when X is finite.

AXIOM 5″. *Suppose $\langle X, \mathscr{E}, \succsim \rangle$ is a structure of qualitative probability. If $A, B \in \mathscr{E}$ and $A \succ B$, then there exists a partition $\{C_1, ..., C_n\}$ of X such that, for $i = 1, ..., n$, $C_i \in \mathscr{E}$ and $A \succ B \cup C_i$.*

Savage (1954) pointed out that if $\mathscr{E}$ does not seem to fulfill Axiom 5″, one can always enlarge it to an algebra that does by adjoining all finite sequences of heads and tails generated by a coin of one's own choice. For any n, this leads to a partition into 2^n events, and, as he said,

> It seems to me that you could easily choose such a coin and choose n sufficiently large so that you would continue to prefer to stake your gain on A, rather than on the union of B and any particular sequence of n heads and tails. For you to be able to do so, you need by no means consider every sequence of heads and tails equally probable (Savage, 1954, p. 38).

Savage established a reformulation of his axiom which is interesting and which we shall need again in Section 5.4.2. Suppose that $\langle X, \mathscr{E}, \succsim \rangle$ is a system of qualitative probability. It is called *fine* if, for every $A \succ \varnothing$, there exists a partition $\{C_1, ..., C_n\}$ of X such that $A \succsim C_i$ for $i = 1, ..., n$. For A, B in $\mathscr{E}$, define $A \sim^* B$ if for all $C, D \succ \varnothing$ such that $A \cap C = B \cap D = \varnothing$, then $A \cup C \succsim B$ and $B \cup D \succsim A$. The system is called *tight* if whenever $A \sim^* B$, then $A \sim B$.

THEOREM 1. *Suppose that $\langle X, \mathscr{E}, \succsim \rangle$ is a structure of qualitative probability. Axiom 5″ is true iff the structure is both fine and tight.*

This is Savage's Theorem 4, p. 38.

Although Savage's axiom is, in the presence of the axioms for qualitative probability, actually stronger than the next condition, which was invoked both by de Finetti (1937) and Koopman (1940a,b), Savage felt that his was the easier to justify intuitively.

AXIOM 5′. *Suppose that $\langle X, \mathscr{E}, \succsim \rangle$ is a structure of qualitative probability. For every positive integer n, there exists a partition $C_1, ..., C_n$ of X such that, for $i, j = 1, ..., n$, $C_i \in \mathscr{E}$ and $C_i \sim C_j$.*

If $\mathscr{E}$ includes all finite sequences of heads and tails generated by what you believe to be independent tosses of a fair coin, i.e., a head is just as probable as a tail, then Axiom 5′ must be fulfilled for you.

A common drawback of both axioms is that they force X to be infinite—the illustrative coins suggest why. The following condition, which is due to Luce (1967), is fulfilled in many infinite structures and in some finite ones, although not in most.

AXIOM 5. *Suppose that $\langle X, \mathscr{E}, \succsim \rangle$ is a structure of qualitative probability. If $A, B, C, D \in \mathscr{E}$ are such that $A \cap B = \varnothing$, $A \succ C$, and $B \succsim D$, then there exist $C', D', E \in \mathscr{E}$ such that:*

(i) $E \sim A \cup B$;

(ii) $C' \cap D' = \varnothing$;

(iii) $E \supset C' \cup D'$;

(iv) $C' \sim C$ *and* $D' \sim D$.

It is difficult to say just what this means other than to restate it in words. If A and B are disjoint, if A is more probable than C, and if B is at least as probable as D, then the axiom asserts that somewhere in the structure there is an event E that is equivalent in probability to $A \cup B$ such that E includes disjoint subsets C' and D' that are equivalent, respectively, to C and D. This axiom can be satisfied by some finite probability spaces, e.g., let $X = \{a, b, c, d\}$, let $P(a) = P(b) = P(c) = 0.2$, $P(d) = 0.4$, and $\mathscr{E}$ is all subsets. In some sense, however, most finite structures that have a probability representation violate the axiom, e.g., any structure whose equivalence classes fail to form a single standard sequence.

As was mentioned, Savage showed that, in the presence of the axioms for qualitative probability, Axiom 5″ is strictly stronger than 5′, and under the same conditions Luce (1967) showed that 5″ is strictly stronger than 5. Fine (1971a,b) used a still weaker structural condition—the existence, for every n, of an n-fold "almost uniform" partition—in conjunction with the (necessary) condition that the order topology on $\mathscr{E}$ have a countable base. None of these axiom systems permits finite models, so none is weaker than Axioms 1–5. Fine (1971a) proved that all are strictly stronger than Axioms 1–5.

Assuming that Axioms 1–5 are weaker than the other systems that have been mentioned and that adding the (necessary) Archimedean condition to those of qualitative probability is not objectionable, the appropriate result to prove is the following.

THEOREM 2. *Suppose that* $\langle X, \mathscr{E}, \succsim \rangle$ *is an Archimedean structure of qualitative probability for which Axiom 5 holds, then there exists a unique order-preserving function P from* $\mathscr{E}$ *into the unit interval* [0, 1] *such that* $\langle X, \mathscr{E}, P \rangle$ *is a finitely additive probability space.*

The proof, which is given in Section 5.3.2, involves reducing this structure to one of extensive measurement. It is evident that the crucial property, if $A \cap B = \varnothing$, then $P(A \cup B) = P(A) + P(B)$, is formally similar to the corresponding one in extensive measurement provided that $\bigcup$ is treated as a concatenation operation $\circ$. Clearly the condition $A \cap B = \varnothing$ will have to be rephrased as a restriction on the concatenation operation of the form: (A, B) in $\mathscr{B}$, where $\mathscr{B} \subset \mathscr{E} \times \mathscr{E}$.

In this connection, it should be noted that an ordering of an algebra of sets can be interpreted as a type of extensive measurement provided that the sets are collections of objects rather than events. Such an interpretation is especially natural for masses, in which case any collection of objects placed on one pan of a balance is an element of the algebra. Such an interpretation is less natural, though nonetheless possible, for length measurement. Observe that when we take one of the primitives to be an algebra of sets, we automatically assume the commutativity and associativity of concatenation, since both properties hold for unions of sets; these properties had to be listed more or less explicitly as axioms in the theories of Chapter 3. This is, of course, a general mathematical phenomena: the more structured the primitives, the weaker the axiomatic system need be in order to describe a given structure.

5.2.4 Preference Axioms for Qualitative Probability

If the ordering $\succsim$ on $\mathscr{E}$ is inferred from choices among acts, then the axioms of qualitative probability can be reformulated as assumptions concerning observable choices. For simplicity, we restrict attention to acts that have either two or three uncertain outcomes.

Let $\mathscr{E}$ be an algebra of sets on X and $\mathscr{C}$ a set of outcomes or consequences; let $a_A \cup b_B \cup c_C$ be the act that has consequence a if A occurs, b if B occurs, and c if C occurs, where it is understood that a, b, c are in $\mathscr{C}$, A, B, C are in $\mathscr{E}$, and A, B, C partition X. Denote the preference relation as $\succsim^*$. This is a weak ordering of both $\mathscr{C}$ and of the set $\mathscr{A}$ of all acts, i.e., formally, a subset of $(\mathscr{C} \times \mathscr{C}) \cup (\mathscr{A} \times \mathscr{A})$. We do not need to assume that a consequence and an act are compared.

We also identify acts that have the same event–outcome interpretation, e.g., permutations do not matter, $a_A \cup b_B \cup c_C = b_B \cup c_C \cup a_A$, and a three-outcome act with two identical outcomes reduces to a two-outcome act, $a_A \cup b_B \cup b_C = a_A \cup b_{B \cup C} = a_A \cup b_{-A}$.

We define a relation $\succsim$ on $\mathscr{E}$ as follows:

$A \succsim B$ *iff for all* $a, b \in \mathscr{C}$, *if* $a \succsim^* b$, *then* $a_A \cup b_{-A} \succsim^* a_B \cup b_{-B}$.

The following three axioms are sufficient (in the presence of the assumptions made informally above) to make $\langle X, \mathscr{E}, \succsim \rangle$ a structure of qualitative probability:

1. *If* $a \succ^* b$, $c \succsim^* d$, *and* $a_A \cup b_{-A} \succsim^* a_B \cup b_{-B}$, *then*
$$c_A \cup d_{-A} \succsim^* c_B \cup d_{-B}.$$
2. *If* $a \succ^* b$, *then* $a_X \cup b_\varnothing \succ^* a_\varnothing \cup b_X$ *and* $a_A \cup b_{-A} \succsim^* a_\varnothing \cup b_X$.
3. *If* $a_A \cup b_B \cup b_C \succsim^* a_{A'} \cup b_B \cup b_{C'}$, *then*
$$a_A \cup a_B \cup b_C \succsim^* a_{A'} \cup a_B \cup b_{C'}.$$

We also need to assume that $\succsim^*$ is nontrivial on $\mathscr{C}$, i.e., there exist a, b with $a \succ^* b$.

The first axiom implies that $\succsim$ is connected: for take $a \succ^* b$, then for any A, B, the connectedness of $\succsim^*$ yields some relation between $a_A \cup b_{-A}$ and $a_B \cup b_{-B}$, and this same relation then holds with c, d substituted for a, b, provided that $c \succsim^* d$. Transitivity of $\succsim$ follows from transitivity of $\succsim^*$, so $\succsim$ is a weak order.

The second axiom obviously implies $X \succ \varnothing$ and $A \succsim \varnothing$.

The third axiom is based on the idea that if two acts have a common consequence b when B occurs, then the contingency b_B is irrelevant to the decision between them, and so b can just as well be replaced by a. The axiom only assumes this in a very restricted situation: when the other two outcomes are b and a. The general principle, which this axiom exemplifies, is called the extended sure-thing principle. It leads to an expected utility representation (see Chapter 8, Savage's book, or Fishburn's book, referred to above). In this highly restricted form, however, it is already sufficient to yield Axiom 3 of qualitative probability: If A is disjoint from $B \cup C$, then $B \succsim C$ if and only if $A \cup B \succsim A \cup C$.

The proof is fairly obvious: Letting $a \succ^* b$, $B \succsim C$ translates into

$$a_B \cup b_A \cup b_{-(A \cup B)} \succsim^* a_C \cup b_A \cup b_{-(A \cup C)}$$

while $A \cup B \gtrsim A \cup C$ translates into

$$a_B \cup a_A \cup b_{-(A\cup B)} \gtrsim^* a_C \cup a_A \cup b_{-(A\cup C)} .$$

Thus the axiom (which is equivalent to its own converse) gives the desired result.

Axioms 1 and 3 are highly interesting propositions about choices. If Axiom 1 is violated, there would seem to be some kind of special interaction between outcomes and events, e.g., such that even though $a >^* b$, $c \gtrsim^* d$, and $a_A \cup b_{-A} \gtrsim^* a_B \cup b_{-B}$, the combination c_A is somehow much less valuable than c_B.

Violations of the extended sure-thing principle have been much discussed (see Chapter 8). From the present standpoint, the interesting thing to note is the light that Axiom 3 sheds on the principle of additivity of probabilities. If $a >^* b$ and $A \gtrsim A'$, then we must have

$$a_A \cup b_B \cup b_C \gtrsim^* a_{A'} \cup b_B \cup b_{C'} .$$

But if, somehow, B adds more to A' than it does to A, we could have the reversal,

$$a_{A'} \cup a_B \cup b_{C'} >^* a_A \cup a_B \cup b_C .$$

An example of a possible violation is offered by the following game (adapting an idea of Ellsberg, 1961). Consider three urns, containing 200 white balls, 200 black balls, and 100 red balls, respectively. One of the first two urns is selected by tossing a fair coin, and without informing the player which one was selected, white or black, its 200 balls are thoroughly mixed with the 100 red balls. Then a single ball is drawn from the mixture, and a prize is awarded or not depending on its color. Let a denote a valuable prize and b denote no prize. The player may prefer the payoff scheme

$$a_{\text{red}} \cup b_{\text{black}} \cup b_{\text{white}}$$

to the scheme

$$a_{\text{white}} \cup b_{\text{black}} \cup b_{\text{red}} .$$

In each case, the objective probability of winning the prize a is exactly one-third. The difference is that in the first case, he knows there are one-third red balls, whereas, in the second case, depending on the coin toss, there may be no white balls at all.

Now suppose that we change b to a on black, i.e., compare

$$a_{\text{red}} \cup a_{\text{black}} \cup b_{\text{white}}$$

to

$$a_{\text{white}} \cup a_{\text{black}} \cup b_{\text{red}} .$$

The objective probability of winning a is now two-thirds in each case, but the tables are turned. In the first scheme, the probability could be as low as one-third, depending on the coin toss, but in the second scheme, the probability is surely two-thirds. The point is that black adds a great deal more unambiguity to white than it does to red. Only if such considerations play no role—either because of the constancy of ambiguity over events or because of sophisticated decision-makers—may we expect the probability ordering inferred from decisions under uncertainty to lead to an additive representation.

Other discussions of the inference of a qualitative probability from decisions are found in Section 8.7.1 and in Anscombe and Aumann (1963), Davidson and Suppes (1956), Edwards (1962), Fishburn (1967g), Pratt, Raiffa, and Schlaifer (1964), Ramsey (1931), Savage (1954), and Suppes (1956).

5.3 PROOFS

5.3.1 Preliminary Lemmas

LEMMA 1. *If $\mathscr{E}$ is an algebra of sets on X (Definition 1, Section 5.1), then $\varnothing$, $X \in \mathscr{E}$ and $\mathscr{E}$ is closed under set difference and intersection. If $\mathscr{E}$ is a σ-algebra, then it is closed under countable intersections.*

PROOF. This is a standard result, easily proved using the formulas $X = A \cup -A$, $A \cap B = -(-A \cup -B)$, etc. ◇

The common hypotheses of Lemmas 2–4 is that $\langle X, \mathscr{E}, \gtrsim \rangle$ is a structure of qualitative probability (Definition 4, Section 5.2.1) and that $A, B, C, D \in \mathscr{E}$.

LEMMA 2. *Suppose that $A \cap B = C \cap D = \varnothing$. If $A \gtrsim C$ and $B \gtrsim D$, then $A \cup B \gtrsim C \cup D$; moreover, if either hypothesis is $>$, then the conclusion is $>$.*

PROOF. Let $A' = A - D$, $D' = D - A$. Note that A', D' are in $\mathscr{E}$ (Lemma 1) and that

$$A' \cap B = A' \cap D = A \cap D' = C \cap D' = \varnothing ,$$

and that $A' \cup D = A \cup D'$.

Using Axiom 3 twice,

$$A' \cup B \gtrsim A' \cup D = A \cup D' \gtrsim C \cup D'.$$

If either hypothesis is strict, then by the "if" part of Axiom 3,

$$A' \cup B \succ C \cup D'.$$

Observe that $A \cap D$ is disjoint from both $A' \cup B$ and $C \cup D'$; thus, by Axiom 3,

$$A \cup B = (A' \cup B) \cup (A \cap D) \succsim (C \cup D') \cup (A \cap D) = C \cup D;$$

and $\succ$ holds if $A' \cup B \succ C \cup D'$. ◇

COROLLARY. *If* $A \cap B = C \cap D = \varnothing$, $A \sim C$, *and* $B \sim D$, *then* $A \cup B \sim C \cup D$.

LEMMA 3. *If* $A \supset B$, *then* $A \succsim B$.

PROOF. Use Axiom 3 on $A - B \succsim \varnothing$ (Axiom 2), noting that B is disjoint from $A - B$ and from $\varnothing$ and that $(A - B) \cup B = A$. ◇

COROLLARY 1. $X \succsim A$.

COROLLARY 2. *If* $A \supset B$, *then* $A \succ B$ *iff* $A - B \succ \varnothing$.

LEMMA 4. *If* $A \succsim B$, *then* $-B \succsim -A$.

PROOF. If $A \succsim B$, $-A \succ -B$, then by Lemma 2, $X \succ X$, contradicting Axiom 1. ◇

5.3.2 Theorem 2 (p. 208)

If $\langle X, \mathscr{E}, \succsim \rangle$ *is an Archimedean structure of qualitative probability for which Axiom 5 is true, then it has a unique, finitely additive probability representation.*

PROOF. Let $\mathbf{A}$ denote the equivalence class that includes A, and let $\mathscr{E}$ be the set of all equivalence classes, excluding $\boldsymbol{\varnothing}$. By Axiom 2, $\mathbf{X} \in \mathscr{E}$. If $\mathbf{X}$ is the only element of $\mathscr{E}$, then we let $P(A) = 0$ if $A \sim \varnothing$, $P(A) = 1$ if $A \sim X$, and P fulfills the assertions of the theorem (note that $\boldsymbol{\varnothing}$ is closed under disjoint union, by Lemma 2). We assume, therefore, that $\mathscr{E}$ contains $\mathbf{A} \neq \mathbf{X}$.

We construct an extensive structure on $\mathscr{E}$ letting

$$\mathscr{B} = \{(\mathbf{A}, \mathbf{B}) \mid A \succ \varnothing, B \succ \varnothing, \text{ and there exist } A' \in \mathbf{A},\ B' \in \mathbf{B} \text{ with } A' \cap B' = \varnothing\}.$$

Thus $\mathscr{B}$ is nonempty because if $\mathbf{A} \neq$ both $\mathbf{X}$ and $\boldsymbol{\varnothing}$, then $-A \succ \varnothing$ (Lemma 4), so the pair $(\mathbf{A}, -\mathbf{A}) \in \mathscr{B}$. We define $\circ$ on $\mathbf{B}$ by letting $\mathbf{A} \circ \mathbf{B} = \mathbf{A \cup B}$, if $A \cap B = \varnothing$. By the corollary of Lemma 2, $\circ$ is well defined. We let $\succsim$ be the induced simple order on $\mathscr{E}$.

We shall now prove that $\langle \mathscr{E}, \succsim, \mathscr{B}, \circ \rangle$ is an extensive structure with no essential maximum (Definition 3.3, Section 3.4.3). The six axioms are established in corresponding numbered paragraphs. It is convenient in some places to use boldface Greek capitals for elements of $\mathscr{E}$.

1. $\succsim$ is a weak order; in fact, it is a simple order.

5. Positivity follows from Corollary 2 of Lemma 3.

2. Associativity. Suppose that $(\boldsymbol{\Gamma}, \boldsymbol{\Delta})$, $(\boldsymbol{\Gamma} \circ \boldsymbol{\Delta}, \boldsymbol{\Lambda}) \in \mathscr{B}$. The essence of the proof is to construct A, B, C pairwise disjoint such that $\mathbf{A} = \boldsymbol{\Gamma}$, $\mathbf{B} = \boldsymbol{\Delta}$, $\mathbf{C} = \boldsymbol{\Lambda}$. For then, the rest will follow via the associativity of the union operation. The construction is as follows. Take disjoint sets $D' \in \boldsymbol{\Gamma} \circ \boldsymbol{\Delta}$, $C' \in \boldsymbol{\Lambda}$. By positivity $\mathbf{D'} \succ \boldsymbol{\Delta}$, thus, by Axiom 5 we can take $E \sim D' \cup C'$ and B, C disjoint with $E \supset B \cup C$ and $B \in \boldsymbol{\Delta}$, $C \in \boldsymbol{\Lambda}$. Then let $A = E - (B \cup C)$. It remains only to show that $A \in \boldsymbol{\Gamma}$; but since $(\mathbf{A} \circ \boldsymbol{\Delta}) \circ \boldsymbol{\Lambda} = \mathbf{E} = (\boldsymbol{\Gamma} \circ \boldsymbol{\Delta}) \circ \boldsymbol{\Lambda}$, this follows from Lemma 2, applied twice.

3. We need to show that if $(\mathbf{A}, \mathbf{C}) \in \mathscr{B}$ and $\mathbf{A} \succsim \mathbf{B}$, then $(\mathbf{C}, \mathbf{B}) \in \mathscr{B}$; the rest follows by the obvious commutativity of $\circ$ and by Lemma 2. If $\mathbf{A} = \mathbf{B}$, there is nothing to show; if $\mathbf{A} \succ \mathbf{B}$, apply Axiom 5.

4. Solvability. If $\mathbf{A} \succ \mathbf{B}$, apply Axiom 5 to $A \succ B$ and $\varnothing \sim \varnothing$ to obtain A', B' with $A' \supset B'$, $A \sim A'$, $B \sim B'$, and let $C = A' - B'$; then $\mathbf{A} = \mathbf{B} \circ \mathbf{C}$.

6. Finally, we show that $\{n \mid \mathbf{B} \succ n\mathbf{A}\}$ is finite, for $\mathbf{A} \in \mathscr{E}$, where $1\mathbf{A} = \mathbf{A}$, $n\mathbf{A} = (n - 1)\,\mathbf{A} \circ \mathbf{A}$. We do this by showing that the existence of $n\mathbf{A}$ implies the existence of an n-term standard sequence relative to $A \succ \varnothing$. For suppose that $n\mathbf{A}$ exists and that we have a k-term standard sequence relative to A, $A_1, \ldots, A_k$, with $k < n$, and with $A_i \in i\mathbf{A}$ for $i \leqslant k$. Since $(k\mathbf{A}, \mathbf{A}) \in \mathscr{B}$, we have $B_k \in k\mathbf{A}$ and $C_k \in \mathbf{A}$ such that $B_k \cap C_k = \varnothing$. Let $A_{k+1} = B_k \cup C_k$. Then clearly, $A_1, \ldots, A_{k+1}$ is a $(k + 1)$-term standard sequence relative to A, and $A_{k+1} \in (k + 1)\,\mathbf{A}$. Since a 1-term standard sequence can be obtained by $A_1 = A$, we have the required result by induction.

Now, by Theorem 3.3 (Section 3.4.3), there is a positive-valued ratio scale ϕ on $\mathscr{E}$ such that

$$\mathbf{A} \succsim \mathbf{B} \quad \text{iff} \quad \phi(\mathbf{A}) \geqslant \phi(\mathbf{B}),$$

and

$$\text{for} \quad (\mathbf{A}, \mathbf{B}) \in \mathscr{B}, \qquad \phi(\mathbf{A} \circ \mathbf{B}) = \phi(\mathbf{A}) + \phi(\mathbf{B}).$$

Choose the unit so $\phi(\mathbf{X}) = 1$. For $A \in \mathscr{E}$, let

$$P(A) = \begin{cases} \phi(\mathbf{A}), & \text{if } A \succ \varnothing, \\ 0, & \text{if } A \sim \varnothing. \end{cases}$$

It is easy to see that P fulfills the assertions of the theorem. Moreover, P is unique, since if another such function P' existed, then $\phi'(A) = P'(A)$ would be a representation of $\langle \mathscr{E}, \succsim, \mathscr{B}, \circ \rangle$. Since $\phi' = \alpha\phi$ and $\phi'(X) = \phi(X) = 1$, $\phi' = \phi$ and $P' = P$. ◇

5.4 MODIFICATIONS OF THE AXIOM SYSTEM

5.4.1 QM-Algebra of Sets

The notions of event and probability given in Definitions 1 and 2 have proved satisfactory for almost all scientific purposes. The one outstanding exception is quantum mechanics. In that theory both $P(A)$ and $P(B)$ may exist and yet $P(A \cap B)$ need not. For example, suppose that A is the event that, at some time t, a certain elementary particle is located in some region α of space and B is the event that, at the same time t, this same particle has a momentum whose value lies in some interval β. Then the event $A \cap B$ means that, at time t, the particle is both in the region α and has a momentum in β. One basic feature of quantum mechanics is that we may be able to observe whether event A occurred or whether event B occurred, but not necessarily be able to observe whether both events A and B, i.e., event $A \cap B$, occurred. In the theory this means that no probability is assignable to $A \cap B$ even when $P(A)$ and $P(B)$ are specified.

Such an incompatibility between two very fundamental sets of ideas must be resolved, presumably by modifying probability theory in some way. Those interested in a full understanding of the physical issues involved and in some of the proposed modifications of both classical logic and probability theory should consult Birkhoff and von Neumann (1936), Kochen and Specker (1965), Reichenbach (1944), Suppes (1961, 1963, 1965, 1966), and Varadarajan (1962). Suffice it to say that the simplest proposal (Suppes, 1966) is to restrict further our notion of an algebra of sets (Definition 1) to:

DEFINITION 5. *Suppose that X is a nonempty set and that $\mathscr{E}$ is a nonempty family of subsets of X. Then $\mathscr{E}$ is a* QM-algebra of sets *on X iff, for every $A, B \in \mathscr{E}$:*

1. $-A \in \mathscr{E}$;
2. *If $A \cap B = \varnothing$, then $A \cup B \in \mathscr{E}$.*

Furthermore, if $\mathscr{E}$ is closed under countable unions of mutually disjoint sets, then $\mathscr{E}$ is called a QM σ-algebra.

Since the definition is still stated in terms of unions, and the difficulty

seemed to be about intersections, it is not clear that we have coped with the problem; however, as is shown in Lemma 5 (Section 5.5.1), all is well in the sense that when A, B are in $\mathscr{E}$, $A \cap B$ is in $\mathscr{E}$ if and only if $A \cup B$ is in $\mathscr{E}$, and if $A \supset B$, then $A - B$ is in $\mathscr{E}$.

As a simple example, if $\langle X, \mathscr{E}, P \rangle$ is a finitely additive probability space, then for any A in $\mathscr{E}$, the set of all events probabilistically independent of A is a QM-algebra of sets on X; see Section 5.8.

What is not clear about this suggestion is just how badly probability theory suffers. For one thing, how many of the important theorems of probability theory cannot now be proved? This is not the place to enter into that discussion. For another, to what extent have we impaired our ability to construct a numerical probability representation from a qualitative ordering? If one carefully reexamines Axioms 1–4 of Definition 4 and Axiom 5 of Section 5.2.3, the proofs of the lemmas used in the proof of Theorem 2, and the proof of that theorem, then one finds that, with the exception of Lemmas 1 and 2, all unions are of disjoint sets and that all differences are of one set that includes another. As we mentioned above, Lemma 5 (Section 5.5.1) replaces Lemma 1. As for Lemma 2, we note that only disjoint unions enter into its statement, that it is a necessary condition for the representation (the argument for necessity is very similar to the argument for Axiom 3), and that it is strictly stronger than Axiom 3. Therefore, we adopt it as Axiom 3′, in place of Axiom 3.

AXIOM 3′. *Suppose that $A \cap B = C \cap D = \varnothing$. If $A \succsim C$ and $B \succsim D$, then $A \cup B \succsim C \cup D$; moreover, if either hypothesis is $\succ$, then the conclusion is $\succ$.*

We shall encounter this axiom again in the treatment of qualitative conditional probability (Sections 5.6 and 5.7). Except for the proof of Lemma 5, we have established the following:

THEOREM 3. *If $\mathscr{E}$ is a* QM-algebra *of sets on X and if $\langle X, \mathscr{E}, \succsim \rangle$ satisfies Axioms* 1, 2, 3′, 4, *and* 5, *then there is a unique function P on $\mathscr{E}$ that satisfies the Kolmogorov axioms (Definition* 2) *and preserves the order $\succsim$ on $\mathscr{E}$.*

5.4.2 Countable Additivity

Theorem 2, and others like it, establish conditions under which finite additivity hold, but nothing has been said about countable additivity (see Definition 2), which is needed in many parts of probability theory. Countably additive representations of qualitative probability structures have been studied by Villegas (1964, 1967) and Fine (1971a). The main result (from our point of view) is embodied in the following definition and theorem. The definition follows Villegas (1964).

DEFINITION 6. *Suppose that* $\langle X, \mathscr{E}, \succsim\rangle$ *is a structure of qualitative probability and that* $\mathscr{E}$ *is a* σ*-algebra. We say that* $\succsim$ *is* monotonically continuous *on* $\mathscr{E}$ *iff for any sequence* $A_1, A_2, \ldots$ *in* $\mathscr{E}$ *and any* B *in* $\mathscr{E}$, *if* $A_i \subset A_{i+1}$ *and* $B \succsim A_i$ *for all* i, *then* $B \succsim \bigcup_{i=1}^{\infty} A_i$.

THEOREM 4. *A finitely additive probability representation of a structure of qualitative probability, on a* σ*-algebra, is countably additive iff the structure is monotonically continuous.*

This theorem is proved in Section 5.5.2. An immediate corollary is that Theorem 2 gives a set of sufficient conditions for a countably additive probability representation for a binary relation $\succsim$ on a σ-algebra: Axioms 1–5 and monotone continuity.

Given Axioms 1–3 (qualitative probability) and monotone continuity, Axioms 4 and 5 can be dispensed with if a different structural condition is used. To formulate the condition, we need a qualitative definition of a probability *atom*. The numerical definition is that A is an atom if $P(A) > 0$ and if $A \supset B$, then $P(B) = P(A)$ or $P(B) = 0$. Corresponding to this, we have:

DEFINITION 7. *Let* $\succsim$ *be a weak ordering of an algebra of sets* $\mathscr{E}$. *An event* $A \in \mathscr{E}$ *is an* atom *iff* $A \succ \varnothing$ *and for any* $B \in \mathscr{E}$, *if* $A \supset B$, *then* $A \sim B$ *or* $B \sim \varnothing$.

The structural condition which replaces Axioms 4 and 5 is that $\mathscr{E}$ has no atoms. Recall the definitions of *fine* and *tight* (Section 5.2.3).

THEOREM 5. *Suppose that* $\langle X, \mathscr{E}, \succsim\rangle$ *is a structure of qualitative probability,* $\mathscr{E}$ *is a* σ*-algebra,* $\succsim$ *is monotonically continuous on* $\mathscr{E}$, *and there are no atoms. Then the structure is both fine and tight, there is a unique order-preserving probability, and it is countably additive.*

For the proof of this and other related results, we refer to Villegas (1964). This structural condition turns out, of course, to be a special case of Axiom 5 (given σ-algebras and monotone continuity), since a structure of qualitative probability that is both fine and tight satisfies Axiom 5 (see Section 5.2.3).

5.4.3 Finite Probability Structures with Equivalent Atoms

As in the chapters on extensive measurement and difference measurement, it is possible here to give a separate and elementary axiomatization of a single finite standard sequence. To do so, we replace the structural restriction (Axiom 5) by some sort of equal-spacing axiom.

In the case of probability structures, it is impossible for a single standard sequence to exhaust $\mathscr{E} - \{\varnothing\}$. For if $A_1, A_2, \ldots, A_n$ is a standard sequence, then by Definition 3, we have $A_2 \sim B_1 \cup C_1$, where B_1, C_1 are disjoint and both are $\sim A_1$; thus, either B_1 or C_1 must be outside the sequence. The most that one can expect is for a single standard sequence to exhaust the set of nonnull equivalence classes $\mathscr{E}$. However, there are a great variety of structures of $\mathscr{E}$ compatible with a single standard sequence in $\mathscr{E}$, e.g., any finite structure satisfying Axiom 5. Perhaps the simplest structure is one with a finite number of equivalent atoms—corresponding to a uniform distribution on a finite set of categories. This structure can be characterized by the following axiom:

AXIOM 6. *If* $A \succsim B$, *then there exists* $C \in \mathscr{E}$ *such that* $A \sim B \cup C$.

This axiom, and the following theorem, are due to Suppes (1969, pp. 5–8).

THEOREM 6. *Suppose that* $\langle X, \mathscr{E}, \succsim\rangle$ *is a structure of qualitative probability (Axioms 1–3 of Definition 4) such that* X *is finite and Axiom 6 holds. Then there exists a unique order-preserving function* P *on* $\mathscr{E}$ *such that* $\langle X, \mathscr{E}, P\rangle$ *is a finitely additive probability space. Moreover, all atoms in* $\mathscr{E}$ *are equivalent.*

The proof is found in Section 5.5.3.

Since any event in a finite probability space is a union of disjoint atoms and a null event, the probabilities are all multiples of the atomic probability.

5.5 PROOFS

5.5.1 Structure of QM-Algebras of Sets

LEMMA 5. *Suppose that* $\mathscr{E}$ *is a* QM*-algebra of sets on* X *(Definition 5), then for all* $A, B \in \mathscr{E}$:

(i) $\varnothing$ *and* $X \in \mathscr{E}$;

(ii) *if* $A \supset B$, *then* $A - B \in \mathscr{E}$;

(iii) $A \cup B \in \mathscr{E}$ *iff* $A \cap B \in \mathscr{E}$.

PROOF. (i) The proof is the same as Lemma 1, since $A \cap -A = \varnothing$.

(ii) Since $A \supset B$ implies $-A \cap B = \varnothing$, and since $-A \in \mathscr{E}$, it follows that $-A \cup B \in \mathscr{E}$. So $A - B = -(-A \cup B) \in \mathscr{E}$.

(iii) If $A \cup B \in \mathscr{E}$, then by part (ii), $A - B = (A \cup B) - B$ and

$B - A = (A \cup B) - A \in \mathscr{E}$. Thus, the symmetric difference,

$$(A - B) \cup (B - A) \in \mathscr{E};$$

and again, by part (ii), the intersection, which is the union minus the symmetric difference, is in $\mathscr{E}$. Conversely, if the intersection is in $\mathscr{E}$, then $A - B = A - (A \cap B)$ and $B - A = B - (A \cap B) \in \mathscr{E}$ by part (ii), so $A \cup B$, which is the disjoint union of $A \cap B$, $A - B$, and $B - A$, is in $\mathscr{E}$. ◇

5.5.2 Theorem 4 (p. 216)

Suppose that a structure of qualitative probability $\langle X, \mathscr{E}, \succsim \rangle$ has an order-preserving probability measure P and that $\mathscr{E}$ is a σ-algebra. Then $\langle X, \mathscr{E}, P \rangle$ is countably additive iff $\succsim$ is monotonically continuous on $\mathscr{E}$.

PROOF. If P is countably additive, let $\{A_i\}$ satisfy $A_i \subset A_{i+1}$ and $A_i \precsim B$. By countable additivity,

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_{i+1} - A_i).$$

Since the partial sums of the right-hand expression are $P(A_{i+1})$, they are bounded by $P(B)$, and so $P(\bigcup_{i=1}^{\infty} A_i) \leqslant P(B)$, implying $B \succsim \bigcup_{i=1}^{\infty} A_i$.

For the converse, note first that by finite additivity alone, if $\{A_i\}$ are pairwise disjoint, then

$$\sum_{i=1}^{\infty} P(A_i) \leqslant P\left(\bigcup_{i=1}^{\infty} A_i\right).$$

This is true because each partial sum on the left is $P(\bigcup_{i=1}^{n} A_i) \leqslant P(\bigcup_{i=1}^{\infty} A_i)$. Suppose that for some disjoint union, strict inequality holds, in violation of countable additivity, i.e., let

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) - \sum_{i=1}^{\infty} P(A_i) = \epsilon > 0.$$

We consider two cases. First, suppose that for some A_m in the sequence,

$$\epsilon \geqslant P(A_m) > 0.$$

Let

$$B_k = \bigcup_{i=1}^{k} A_i, \qquad B = \left(\bigcup_{i=1}^{\infty} A_i\right) - A_m.$$

For each k, $B_k \subset B_{k+1}$, and

$$\begin{aligned} P(B_k) &\leqslant \sum_{i=1}^{\infty} P(A_i) \\ &= P\left(\bigcup_{i=1}^{\infty} A_i\right) - \epsilon \\ &\leqslant P\left(\bigcup_{i=1}^{\infty} A_i\right) - P(A_m) \\ &= P(B). \end{aligned}$$

Thus, $B \succsim B_k$. But since $A_m \succ \varnothing$, we have

$$\bigcup_{k=1}^{\infty} B_k = \bigcup_{i=1}^{\infty} A_i = B \cup A_m \succ B.$$

Thus, monotone continuity is contradicted.

The other case is where no such A_m exists. Then clearly, $P(A_i) = 0$ for all but finitely many i. Let J be the subset of integers i for which $P(A_i) = 0$, and let

$$A = \bigcup_{i \in J} A_i.$$

By finite additivity,

$$\begin{aligned} P\left[\left(\bigcup_{i=1}^{\infty} A_i\right) - A\right] &= P\left(\bigcup_{i \notin J} A_i\right) \\ &= \sum_{i \notin J} P(A_i) \\ &= \sum_{i=1}^{\infty} P(A_i) \\ &= P\left(\bigcup_{i=1}^{\infty} A_i\right) - \epsilon. \end{aligned}$$

Therefore, $P(A) = \epsilon$. Let

$$C_k = \bigcup_{\substack{i \in J \\ i \leqslant k}} A_i.$$

Then $P(C_k) = 0$ by finite additivity; so we have

$$C_k \subset C_{k+1}, \qquad \varnothing \succsim C_k,$$

but

$$\bigcup_{k=1}^{\infty} C_k = A \succ \varnothing,$$

violating monotone continuity. ◇

5.5.3 Theorem 6 (p. 217)

A structure $\langle X, \mathscr{E}, \succsim \rangle$ of qualitative probability for which X is finite and which satisfies Axiom 6 (Section 5.4.3) has a unique probability representation; moreover, all atoms are equivalent.

PROOF. Define a new structure $\langle X', \mathscr{E}', \succsim' \rangle$ in which the nonempty null events are eliminated. (If N is the union of all null events, $N \sim \varnothing$; put $X' = X - N$, $\mathscr{E}' = \{A - N \mid A \in \mathscr{E}\}$, $A - N \succsim' B - N$ iff $A \succsim B$.) It is easy to show that $\langle X', \mathscr{E}', \succsim' \rangle$ also satisfies the hypotheses of the theorem; moreover, any nonempty event in $\mathscr{E}'$ is expressible in a unique way as the union of finitely many pairwise disjoint atoms.

Let A_1 be a minimal atom with respect to $\succsim'$ (such exists by the finiteness of X) and let $\{A_1, ..., A_n\}$ be the set of distinct atoms equivalent to A_1. We show there are no other atoms in $\mathscr{E}'$. If there are others, let A be minimal among the atoms $\succ' A_1$. Define B and C by

$$B = A \cup (A_2 \cup \cdots \cup A_n),$$
$$C = A_1 \cup (A_2 \cup \cdots \cup A_n).$$

(If $n = 1$, $B = A$, $C = A_1$.) By Axiom 3, $B \succ' C$. By Axiom 6, there exists $D \in \mathscr{E}'$ such that $B \sim' C \cup D$. With no loss in generality, take D disjoint from C. Since D contains no minimal atoms, $D \succsim' A$; but then

$$B \sim' C \cup D \succsim' C \cup A = B \cup A_1 \succ' B,$$

a contradiction. Therefore $\{A_1, ..., A_n\}$ is the set of all atoms.

Now construct μ on $\mathscr{E}$ by letting $\mu(A)$ be the cardinality of the set

$$\{A_i \mid A_i \subset A - N\}.$$

Let $P(A) = \mu(A)/n$. It is easy to verify that P is an order-preserving probability and that P is unique. ◇

5.6 A REPRESENTATION BY CONDITIONAL PROBABILITY

Most work in the theory of probability requires the important defined notion of conditional probability. Intuitively, $P(A \mid B)$ is the probability of the event A when we have the added information that the outcome is one

of the elements of the event B—in other words, when we know that the event B occurred. Thus, the outcome must be in both A and B, that is, in $A \cap B$. This immediately suggests $P(A \cap B) = P(A \mid B)$; however, this cannot be correct since, given B, either A or $-A$ must certainly occur so $P(A \mid B) + P(-A \mid B) = 1$, whereas, $P(A \cap B) + P(-A \cap B) = P(B)$. However, if we set

$$P(A \mid B) = P(A \cap B)/P(B), \tag{2}$$

all is well, provided that $P(B) > 0$. This is the accepted definition of *conditional probability*.

Because this concept is so important and because ordinary probability is the special case of it in which $B = X$, several authors have treated it as the basic notion to be axiomatized and have given generalizations of Kolmogorov's axioms (Definition 2). Copeland (1941, 1956) and Rényi (1955) are of particular interest. Császár (1955) studied conditions under which a real-valued function of two set-valued arguments, which is what conditional probability is, can be expressed in the quotient form of Equation (2) in terms of a function of one set-valued argument, which is what unconditional probability is.

Our concerns are somewhat different. We wish to know conditions under which a qualitative relation of the form $A \mid B \succsim C \mid D$, meaning "$A$ given B is qualitatively at least as probable as C given D," can be represented by a probability measure of the form of Equation (2) in the sense that

$$A \mid B \succsim C \mid D \quad \text{iff} \quad P(A \cap B)/P(B) \geqslant P(C \cap D)/P(D).$$

To our knowledge, the only published attempts to do this are Koopman (1940 a,b), Aczél (1961, 1966, p. 319), Luce (1968), and Domotor (1969). Domoter gave necessary and sufficient conditions in the finite case similar to the axiom systems in Chapter 9. The other three systems have much in common. Koopman treated $\varnothing$ as the only null event and did not assume that every pair $A \mid B$ and $C \mid D$ are necessarily comparable; whereas Aczél and Luce admitted other null events and did require comparability of all pairs. The lack of comparability forced Koopman to introduce some extra axioms. Aczél and Luce postulated an additivity requirement, similar to Axiom 3 of Definition 4; whereas, Koopman invoked the property that if $A \mid B \succsim C \mid D$, then $-C \mid D \succsim -A \mid B$, to serve an analogous function. The major differences lie in the choice of sufficient conditions and in the proofs. Aczél postulated a real-valued function on the pairs $A \mid B$, which gives the ordering, and imposed continuity conditions. His proof used results in functional equations. Koopman used a kind of uniform-partition postulate and constructed the probability function by a limiting process. Luce, whose necessary conditions were essentially the same as those of Aczél, reduced the result to known results in measurement theory, using a solvability axiom. He also used a functional-equation argument. Here, we present a modification of Luce's system.

In addition to the above studies, Copeland (1956) stated a system of axioms that is a good deal more transparent than Koopman's, but he did not construct a numerical representation. Moreover, he assumed that his basic system of elements is closed under the conditioning operation, which is to say, if A and B are events, where B is nonnull (see below), then $A \mid B$ is also treated as an event. This seems a trifle odd intuitively, although, as a matter of fact, our structural restriction entails that every $A \mid B$ is equivalent in probability to some event.

5.6.1 Necessary Conditions: Qualitative Conditional Probability

At first glance, one might expect to begin with an algebra of sets $\mathscr{E}$ and a relation $\gtrsim$ on $\mathscr{E} \times \mathscr{E}$; however, the proviso that $P(B) > 0$ in Equation (2) makes it clear that matters are not quite this simple. Events with probability 0—null events—must be excluded from the conditioning position. Such events form a subset $\mathscr{N}$ of $\mathscr{E}$, and the relation $\gtrsim$ is on $\mathscr{E} \times (\mathscr{E} - \mathscr{N})$. We use the suggestive notion $A \mid B$ for a typical element of $\mathscr{E} \times (\mathscr{E} - \mathscr{N})$. Since we are familiar with the general procedure for finding necessary conditions, we first state the axioms in the form of a definition and then discuss those that are novel to this context.

DEFINITION 8. *Suppose that X is a nonempty set, $\mathscr{E}$ is an algebra of sets on X, $\mathscr{N}$ is a subset of $\mathscr{E}$, and $\gtrsim$ is a binary relation on $\mathscr{E} \times (\mathscr{E} - \mathscr{N})$. The quadruple $\langle X, \mathscr{E}, \mathscr{N}, \gtrsim \rangle$ is a* structure of qualitative conditional probability *iff for every A, B, C, A', B', and $C' \in \mathscr{E}$ (or $\in \mathscr{E} - \mathscr{N}$, whenever the symbol appears to the right of $\mid$) the following six axioms hold:*

1. *$\langle \mathscr{E} \times (\mathscr{E} - \mathscr{N}), \gtrsim \rangle$ is a weak order.*
2. *$X \in \mathscr{E} - \mathscr{N}$; and $A \in \mathscr{N}$ iff $A \mid X \sim \varnothing \mid X$.*
3. *$X \mid X \sim A \mid A$ and $X \mid X \gtrsim A \mid B$.*
4. *$A \mid B \sim A \cap B \mid B$.*
5. *Suppose that $A \cap B = A' \cap B' = \varnothing$. If $A \mid C \gtrsim A' \mid C'$ and $B \mid C \gtrsim B' \mid C'$, then $A \cup B \mid C \gtrsim A' \cup B' \mid C'$; moreover, if either hypothesis is $>$, then the conclusion is $>$.*
6. *Suppose that $A \supset B \supset C$ and $A' \supset B' \supset C'$. If $B \mid A \gtrsim C' \mid B'$ and $C \mid B \gtrsim B' \mid A'$, then $C \mid A \gtrsim C' \mid A'$; moreover, if either hypothesis is $>$, then the conclusion is $>$.*

The structure is Archimedean *provided that:*

7. *Every standard sequence is finite, where $\{A_i\}$ is a* standard sequence *iff for all i, $A_i \in \mathscr{E} - \mathscr{N}$, $A_{i+1} \supset A_i$, and $X \mid X > A_i \mid A_{i+1} \sim A_1 \mid A_2$.*

Of course, as in all our axiom systems, the choice of necessary axioms depends somewhat on the nature of the nonnecessary structural restrictions. In the case of Definition 4 (qualitative probability), the three axioms are extremely natural and have long been accepted as the definition of qualitative probability. We added only the Archimedean axiom and a structural restriction. Here, we must be somewhat more tentative about the choice of axioms. Lemmas 9–12 (Section 5.7.1) and their corollaries are also necessary conditions; they are derived from Axioms 1–6 only with the help of the structural axiom introduced below. Some of these properties could well be included in the concept of qualitative conditional probability. Indeed, in Section 5.6.3 we shall relabel Lemma 12 as Axiom 6′ and use it instead of Axiom 6 in the development of a nonadditive conditional probability representation (Section 5.6.4).

Some properties, which it would seem natural to include in qualitative conditional probability, follow from Axioms 1–6 without any structural restriction. For example, Axiom 3 could be expanded into

$$X \mid X \gtrsim A \mid B \gtrsim \varnothing \mid X,$$

but this can be proved (Corollary 1 to Lemma 6).

The proof that Axioms 1–5 are necessary for the desired representation is quite easy. Note, for example, that $P(A \mid B) = P(A \cap B \mid B)$, which yields Axiom 4. The derivation of Axiom 5 is similar to that of its analog, Axiom 3 of Definition 4.

To show that Axiom 6 is necessary, note that if $B \mid A \gtrsim C' \mid B'$, with $A \supset B$ and $B' \supset C'$, then

$$P(B)/P(A) \geqslant P(C')/P(B').$$

Similarly, $C \mid B \gtrsim B' \mid A'$, $B \supset C$, and $A' \supset B'$ yield

$$P(C)/P(B) \geqslant P(B')/P(A').$$

Since the P-values are nonnegative, these inequalities can be multiplied to yield

$$P(C)/P(A) \geqslant P(C')/P(A'),$$

or $C \mid A \gtrsim C' \mid A'$, as required. Moreover, we know that $P(B)$ and $P(B')$ are positive, hence, by the second inequality, $P(C)$ is positive. Thus, if either inequality is strict, multiplication preserves the strict inequality.

The Archimedean property, 7, is also derived by multiplying fractions: Since $P(A_n) \leqslant 1$ and $P(A_1) > 0$,

$$0 < P(A_1) \leqslant \frac{P(A_1)}{P(A_n)} = \frac{P(A_1)}{P(A_2)}\frac{P(A_2)}{P(A_3)}\cdots\frac{P(A_{n-1})}{P(A_n)} = \left[\frac{P(A_1)}{P(A_2)}\right]^{n-1}.$$

Since $X \mid X \succ A_1 \mid A_2$, it follows that $P(A_1)/P(A_2) < 1$. Thus, the inequality $P(A_1) \leqslant [P(A_1)/P(A_2)]^{n-1}$ can hold for only finitely many n.

See also the further discussion of Axiom 6 (as contrasted with Axiom 6′) in Sections 5.6.3 and 5.6.4 (also, following Lemma 12 in Section 5.7.1) and of Axiom 7 in Section 5.7.2.

5.6.2 Sufficient Conditions

Continuing the numbering of Definition 8, we add the nonnecessary property:

AXIOM 8. *If $A \mid B \succsim C \mid D$, then there exists C' in $\mathscr{E}$ such that $C \cap D \subset C'$ and $A \mid B \sim C' \mid D$.*

This simply says that $\mathscr{E}$ is sufficiently rich so that whenever $A \mid B \succ C \mid D$, we can add just enough to C to make $C' \mid D$ equivalent to $A \mid B$. This is a rather strong structural restriction. In particular, Axioms 1–6 and 8 yield Axiom 5 of Section 5.2.3 (see 5.7.2), and moreover, it can be shown that, apart from trivial cases, Axiom 5′ of Section 5.2.3 holds. That is, X can be partitioned into arbitrarily fine equivalent events. Koopman's system included the nonnecessary postulate that for every positive integer n there is a nonnull event that can be partitioned into n equiprobable events. It would, of course, be desirable to replace Axiom 8 by something weaker, even if additional necessary axioms had to be added.

THEOREM 7. *Suppose that $\langle X, \mathscr{E}, \mathscr{N}, \succsim\rangle$ is an Archimedean structure of qualitative conditional probability (Definition 8) for which Axiom 8 holds. Then there exists a unique real-valued function P on $\mathscr{E}$, such that for all $A, C \in \mathscr{E}$ and all $B, D \in \mathscr{E} - \mathscr{N}$:*

(i) *$\langle X, \mathscr{E}, P\rangle$ is a finitely additive probability space;*

(ii) *$A \in \mathscr{N}$ iff $P(A) = 0$;*

(iii) *$A \mid B \succsim C \mid D$ iff $P(A \cap B)/P(B) \geqslant P(C \cap D)/P(D)$.*

The proof (Section 5.7) involves three major steps. First, a number of elementary properties, such as $A \mid B \succsim C \mid D$ implies $-C \mid D \succsim -A \mid B$, are shown to follow from the axioms (Section 5.7.1). Second, we introduce an ordering $\succsim'$ on $\mathscr{E}$, by letting $A \succsim' B$ if and only if $A \mid X \succsim B \mid X$ (Section 5.7.2). The structure $\langle X, \mathscr{E}, \succsim'\rangle$ is shown to satisfy the hypotheses of Theorem 2. By the proofs of Theorems 2 and 3.3 (reduction to extensive measurement, and thence, to Hölder's theorem) we obtain an Archimedean, positive, regular, ordered local semigroup (Definition 2.2), whose elements are equivalence classes in $\mathscr{E} - \mathscr{N}$.

Third, we obtain a semiring structure by using the conditional structure to define a multiplication operation (Section 5.7.3). If $A \mid X \sim C \mid B$, with $B \supset C$, then the representation that we aim for gives $P(A)\,P(B) = P(C)$. Therefore we define $\mathbf{A} * \mathbf{B} = \mathbf{C}$ (where boldface letters denote equivalence classes). It is then easy to show that the axioms of Definition 2.4 hold. The required representation follows from the semiring theorem, 2.6.

5.6.3 Further Discussion of Definition 8 and Theorem 7

In this subsection we discuss three topics: the role of various axioms in the proof of Theorem 7; alternative proofs and axiomatizations; and the testability of the axioms.

Axioms 1–3 of Definition 8 are obvious ground rules for qualitative conditional probability, and are often used in the proof without explicit reference. For example, we may say, "since $A \mid A \sim B \mid B$...," without noting that this uses Axioms 3 ($X \mid X \sim A \mid A$) and 1 (transitivity).

By restricting the relation $\succsim$ to hold between $A \mid B$ and $C \mid D$ only if $B \supset A$ and $D \supset C$, we could bypass Axiom 4 altogether. In effect, this is what we do during the proof, establishing the representation conditions for $B \supset A$, $D \supset C$:

(iii)′ *$A \mid B \succsim C \mid D$ iff $P(A)/P(B) \geqslant P(C)/P(D)$.*

We then use $A \mid B \sim A \cap B \mid B$ only to pass from this to the more usual representation, involving condition (iii) of Theorem 7.

Axiom 5 is, of course, the key assumption for an additive probability measure. It is analogous to Lemma 2 of Section 5.3.1 and Axiom 3′ of Section 5.4.1. From the standpoint of semirings, it has two principal uses. First, it is used to show that $\oplus$ is well defined and monotone, where if $A \cap B = \varnothing$, $\mathbf{A} \oplus \mathbf{B} = \mathbf{A \cup B}$. This only uses part of the strength of the axiom, i.e., if $A \mid X \succsim A' \mid X$ and $B \mid X \succsim B' \mid X$, then $A \cup B \mid X \succsim A' \cup B' \mid X$ (of course, $A \cap B = A' \cap B' = \varnothing$). This use of Axiom 5 is not really explicit, being buried in the reductions to Theorems 2 and 3.3. Rather, 5 is used heavily in the preliminary lemmas to derive other needed features of qualitative conditional probability. The second major use in the semiring development is to prove right distributivity of multiplication $*$ over addition $\oplus$. Here, we use the form of the axiom with $\mid C$ instead of $\mid X$. The structure of the axiom suggests right distributivity, i.e., if $A \mid X \sim A' \mid C$ $(\mathbf{A} * \mathbf{C} = \mathbf{A}')$ and $B \mid X \sim B' \mid C$ $(\mathbf{B} * \mathbf{C} = \mathbf{B}')$, then $A \cup B \mid X \sim A' \cup B' \mid C$ $[(\mathbf{A} \oplus \mathbf{B}) * \mathbf{C} = \mathbf{A}' \oplus \mathbf{B}' = (\mathbf{A} * \mathbf{C}) \oplus (\mathbf{B} * \mathbf{C})]$.

Axiom 6 also plays two main roles in the semiring development, besides its use in some preliminary lemmas. First, it is used to show that multiplication $*$ is commutative. This leads immediately to the left distributive law. Second, Axiom 6 leads to Lemma 12 and its corollary (Section 5.7.1), which are used to prove left and right monotonicity of multiplication.

Except for commutativity of multiplication, we would be better off using Lemma 12 as an axiom, instead of Axiom 6. To further this discussion, we state Lemma 12 here as an alternative axiom, 6′.

AXIOM 6′. *Suppose that* $A \supset B \supset C$ *and* $A' \supset B' \supset C'$. *If* $B \mid A \succsim B' \mid A'$ *and* $C \mid B \succsim C' \mid B'$, *then* $C \mid A \succsim C' \mid A'$; *moreover, if either hypothesis is* $\succ$ *and* $C \in \mathscr{E} - \mathscr{N}$, *then the conclusion is* $\succ$.

It can be seen that Axiom 6′ is very similar to Axiom 6. The hypotheses of Axiom 6′ are "uncrossed," with $B \mid A$ dominating $B' \mid A'$, etc. Axiom 6′ is very analogous to the weak monotonicity condition in difference structures (Chapter 4), whereas Axiom 6 is analogous to what has been called the strong 6-tuple condition (Block & Marschak, 1960). Here, we use Axiom 6, plus other conditions, including Axiom 8, to derive 6′; then 6′ does most of the work, except for commutativity of multiplication. We could use 6′ as an axiom, if we were willing to add another axiom to guarantee left distributivity. For example, we could assume:

Suppose that $A \supset A'$, $B \supset B'$, *and* $A \cap B = \varnothing$. *If* $A' \mid A \sim B' \mid B$, *then* $A' \cup B' \mid A \cup B \sim A' \mid A \sim B' \mid B$.

Alternatively, we could try to push through the analogy with difference structures, deriving 6 from 6′ via a representation theorem from Chapter 4. This is done in the next subsection (proof in 5.7.4), but we seem to require a stronger structural condition (see Axiom 8′, below) in order to obtain the solvability axiom of positive-difference structures.

Luce's (1968) proof relied on the positive-difference structure of conditional probability. The exponential of a difference representation gave a ratio representation Q. Normalizing so that $Q(X) = 1$, Q fulfilled all the assertions of Theorem 7, except possibly additivity. The normalized Q was unique up to $Q \to Q^{\alpha}$, and Luce used a functional-equation argument to show that α could be chosen such that additivity also held. Our use of the ordered local semiring incorporates the essence of a slightly different functional-equation argument, since the proof of Theorem 2.6 is based precisely on finding the unique additive representation for a semiring that is also multiplicative.

Archimedean axioms are familiar to the reader by this time, so we need not comment on the role of Axiom 7; but we do note that we have defined *standard sequence* in terms of the ratio structure, rather than in terms of the additive structure. We do so because the resulting definition is much simpler (compare Definition 3) and it leads directly to the Archimedean axiom for positive differences (see Section 5.7.4). However, the proof that Axioms 1–8 imply the Archimedean axiom for the additive structure is a bit tricky (5.7.2).

The structural condition, 8, plays many different roles. We note here that it is used to define multiplication: $\mathbf{A} * \mathbf{B} = \mathbf{C}$ where C is the solution to $A \mid X \sim C \mid B$, with $B \supset C$. The solution exists by Axiom 8 (note Axiom 4 is also used), since $A \mid X \succsim \varnothing \mid B$ (Corollaries 1 and 2 of Lemma 6, below).

This completes the discussion of the roles of the various axioms, and of alternative axiomatizations and proofs. We now comment briefly on the testability of the axioms.

The status of the various axioms depends critically on the method used to establish the empirical ordering over $\mathscr{E} \times (\mathscr{E} - \mathscr{N})$. If the ordering is estimated from relative frequency counts—by the proportion of the times that A occurs when B occurs—we would most likely attribute any failure of the axioms to error. If the ordering is obtained by having a subject rate the likelihood of A, given B, then it would seem that the most interesting axioms to test are 5 and 6. Axiom 4 could fail if the subject did not perceive set relations correctly, but this would not bear strongly on the issue of a probability representation. The possibility that Axiom 5 might fail while 6 holds motivates our development, in the next section, of a nonadditive conditional representation. If Axiom 6 fails, then the properties of "subjective conditional probability" are indeed very far removed from the representation here axiomatized.

In the case where the relation $\succsim$ is inferred from decisions under uncertainty, Axiom 1 also comes into question, for the same reasons as in unconditional representations (see Section 5.2.4 and Chapter 8).

It seems hard to envisage situations in which tests of Axioms 2, 3, 7, or 8 would be empirically interesting. In most, if not all, practical situations, $\mathscr{N} = \{\varnothing\}$, which simplifies matters considerably.

5.6.4 A Nonadditive Conditional Representation

As we mentioned in the preceding subsection, there may be cases where an ordering of "conditional probability" satisfies the multiplicative property expressed by Axiom 6 or 6′, but does not satisfy the additive property of Axiom 5. We therefore want a nonadditive representation to be obtained without using Axiom 5.

We retain Axioms 1–4 and 7 of Definition 8; we replace Axiom 6 by Axiom 6′, introduced in the preceding subsection; and we replace Axioms 5 and 8 by the following:

5′. *If* $A \in \mathcal{N}$ *and* $A \supset B$, *then* $B \in \mathcal{N}$.

8′. *If* $A \mid B \succsim C \mid D$, *then there exists* $C' \in \mathscr{E}$ *such that* $C \cap D \subset C'$ *and* $A \mid B \sim C' \mid D$; *if, in addition,* $C \in \mathscr{E} - \mathcal{N}$, *then there exists* D' *such that* $D \supset D' \supset C \cap D$ *and* $A \mid B \sim C \mid D'$.

Axiom 5′ corresponds to Lemma 8 (iii) below, i.e., it can be proved from Axioms 1–6 of Definition 8. The first part of Axiom 8′ is just a restatement of Axiom 8; the second part has the additional proviso that just enough can be removed from $D - C$ to obtain a subset D' such that $A \mid B \sim C \mid D'$. Naturally, this must be restricted to the case where C is not in $\mathcal{N}$.

THEOREM 8. *Let* $\langle X, \mathscr{E}, \mathcal{N}, \succsim \rangle$ *satisfy the conditions of Theorem 7, except that Axioms 5, 6, and 8 are replaced by 5′, 6′, and 8′, respectively. Then there exists a function Q from* $\mathscr{E} - \mathcal{N}$ *to* (0, 1], *such that*:

(i) $Q(X) = 1$;

(ii) *if* $A \cap B,\ C \cap D \in \mathscr{E} - \mathcal{N}$, *then* $A \mid B \succsim C \mid D$ *iff*

$$Q(A \cap B)/Q(B) \geqslant Q(C \cap D)/Q(D).$$

If Q' *is any other function with the same properties, then* $Q' = Q^{\alpha}$, $\alpha > 0$.

Note that this theorem deals entirely with the restriction of $\succsim$ to $(\mathscr{E} - \mathcal{N}) \times (\mathscr{E} - \mathcal{N})$; we do not have Q defined on elements of $\mathcal{N}$. In order to extend Q to $\mathcal{N}$, so that (ii) still holds, and so that A is in $\mathcal{N}$ if and only if $Q(A) = 0$, we would need to add certain additional conditions, which are provable with the help of Axiom 5, but not provable from the present axioms. For example, we would need properties such as $A \mid B \succsim \varnothing \mid X$, $\varnothing \mid X \sim \varnothing \mid B$, and $A \mid B \sim \varnothing \mid B$ if and only if A is in $\mathcal{N}$. Since the extension to $\mathcal{N}$ complicates matters and seems of little interest, we omit it.

Theorem 8 is proved in Section 5.7.4.

5.7 PROOFS

5.7.1 Preliminary Lemmas

The common hypothesis of the following four lemmas is that $\langle X, \mathscr{E}, \mathcal{N}, \succsim \rangle$ is a structure of qualitative conditional probability (Definition 8, Axioms 1–6). Axiom 7 is not used, and Axiom 8 is assumed only in Lemma 9.

LEMMA 6. *If* $A \mid B \succsim C \mid D$, *then* $-C \mid D \succsim -A \mid B$.

PROOF. Suppose that, on the contrary, $-A \mid B \succ -C \mid D$. By the strict inequality clause of Axiom 5, $(A \cup -A) \mid B \succ (C \cup -C) \mid D$. By Axioms 1 and 4, $B \mid B \succ D \mid D$, contradicting $B \mid B \sim D \mid D$ derived from Axioms 1 and 3. ◇

COROLLARY 1. $A \mid B \succsim \varnothing \mid X$.

PROOF. Axiom 3 and the lemma. ◇

COROLLARY 2. $\varnothing \mid A \sim \varnothing \mid B$.

PROOF. By Axioms 1 and 3, $A \mid A \sim B \mid B$; by the lemma, $-A \mid A \sim -B \mid B$; and the conclusion follows by Axioms 4 and 1. ◇

COROLLARY 3. *If* $A \mid C \sim -A \mid C$ *and* $B \mid D \sim -B \mid D$, *then*

$$A \mid C \sim B \mid D.$$

PROOF. If $A \mid C \succ B \mid D$, we can use Axiom 1 to derive

$$-A \mid C \succ -B \mid D;$$

this contradicts Lemma 6. ◇

Of course, in the representation, $A \mid C \sim -A \mid C$ will yield $P(A \mid C) = \frac{1}{2}$. Corollary 3 is the qualitative version of that result.

LEMMA 7. *If* $A \supset B$, *then* $A \mid C \succsim B \mid C$.

PROOF. By Corollaries 1 and 2 of Lemma 6, $(A - B) \mid C \succsim \varnothing \mid C$. Also, $B \mid C \sim B \mid C$. By Axiom 5, $A \mid C \succsim B \mid C$. ◇

We collect together a number of properties of the set $\mathcal{N}$ in one lemma:

LEMMA 8.

(i) $\varnothing \in \mathcal{N}$.

(ii) *If* $A \in \mathcal{N}$, *then* $-A \in \mathscr{E} - \mathcal{N}$.

(iii) *If* $A \in \mathcal{N}$, *and* $A \supset B$, *then* $B \in \mathcal{N}$.

(iv) *If* $A, B \in \mathcal{N}$, *then* $A \cup B \in \mathcal{N}$.

PROOF.

(i) $\varnothing \mid X \sim \varnothing \mid X$.

(ii) If $A,\ -A \in \mathcal{N}$, then $A \mid X \sim \varnothing \mid X \sim -A \mid X$. By Lemma 6, $-A \mid X \sim X \mid X \sim A \mid X$, i.e., $X \mid X \sim \varnothing \mid X$, contradicting Axiom 2.

(iii) By Lemma 7, $A \mid X \sim \varnothing \mid X$, $A \supset B$ imply $\varnothing \mid X \gtrsim B \mid X$. By Corollary 1 to Lemma 6, $B \mid X \sim \varnothing \mid X$, or B is in $\mathscr{N}$.

(iv) By part (iii), $(A - B) \mid X \sim \varnothing \mid X$; this, with $B \mid X \sim \varnothing \mid X$, yields $(A \cup B) \mid X \sim \varnothing \mid X$, by Axiom 5. ◇

LEMMA 9. *If $A \mid B \sim \varnothing \mid B$ and $B \supset A$, then $A \in \mathscr{N}$. Conversely, if Axiom 8 holds, $B \in \mathscr{E} - \mathscr{N}$, and $A \in \mathscr{N}$, then $A \mid B \sim \varnothing \mid B$.*

PROOF. Suppose that $A \mid B \sim \varnothing \mid B$, $B \supset A$, but $A \in \mathscr{E} - \mathscr{N}$. By Corollary 2 of Lemma 6, $\varnothing \mid A \sim \varnothing \mid X$. Since $B \in \mathscr{E} - \mathscr{N}$, Axiom 2 and Corollary 1 of Lemma 6 yield $\varnothing \mid X \prec B \mid X$. We have

$$\varnothing \mid A \prec B \mid X, \qquad A \mid B \sim \varnothing \mid B,$$

whence, by Axiom 6, $\varnothing \mid B \prec \varnothing \mid X$, contradicting Corollary 2 of Lemma 6.

For the converse, suppose that $B \in \mathscr{E} - \mathscr{N}$ and $A \mid B \succ \varnothing \mid B$. We apply Axiom 8 to $A \mid B \succ \varnothing \mid X$ to obtain

$$A \mid B \sim C \mid X \succ \varnothing \mid X.$$

Thus, $C \in \mathscr{E} - \mathscr{N}$. Since also $B \in \mathscr{E} - \mathscr{N}$, we have $B \mid X \succ \varnothing \mid C$. By Axiom 4, $A \cap B \mid B \sim C \mid X$. By Axiom 6, $A \cap B \mid X \succ \varnothing \mid X$, so $A \cap B \in \mathscr{E} - \mathscr{N}$, and by Lemma 8 (iii), $A \in \mathscr{E} - \mathscr{N}$. ◇

For the remaining lemmas, we assume that $\langle X, \mathscr{E}, \mathscr{N}, \gtrsim \rangle$ satisfies Axioms 1–6 and Axiom 8.

The next two lemmas have the content of Axiom 3 of positive-difference structures (Definition 4.1): if ab, bc are positive intervals, then ac is strictly greater than either of them. Here, AB will be a positive interval if $A \supset B$ and $X \mid X \succ B \mid A$. Thus, we want to establish that if $A \supset B \supset C$, with AB, BC positive, then $C \mid A$ is strictly smaller than both $B \mid A$ and $C \mid B$.

LEMMA 10. *If $A \supset B \supset C$ and $B \in \mathscr{E} - \mathscr{N}$, then $C \mid B \gtrsim C \mid A$, and $\succ$ holds unless either C or $A - B \in \mathscr{N}$.*

PROOF. By Lemma 8 (iii), $A \in \mathscr{E} - \mathscr{N}$. If $C \in \mathscr{N}$, then Lemma 9 gives

$$C \mid B \sim \varnothing \mid B \sim \varnothing \mid A \sim C \mid A,$$

as required. So assume that $C \in \mathscr{E} - \mathscr{N}$.

We have $C \mid B \sim C \mid B$ and $C \mid C \gtrsim B \mid A$, whence, Axiom 6 yields $C \mid B \gtrsim C \mid A$. Moreover, $\succ$ holds unless $C \mid C \sim B \mid A$; but in that case, Lemma 6 gives $\varnothing \mid C \sim (A - B) \mid A$, and now, from Lemma 9 and $(A - B) \mid A \sim \varnothing \mid A$, we have $A - B \in \mathscr{N}$. ◇

LEMMA 11. *If $A \supset B \supset C$ and $A \in \mathscr{E} - \mathscr{N}$, then $B \mid A \gtrsim C \mid A$, and $\succ$ holds unless either B or $B - C \in \mathscr{N}$.*

PROOF. If $B \in \mathscr{N}$, then $C \in \mathscr{N}$ and $B \mid A \sim \varnothing \mid A \sim C \mid A$, as required. So assume that $B \in \mathscr{E} - \mathscr{N}$.

We have $A \mid A \gtrsim C \mid B$, $B \mid A \sim B \mid A$, so Axiom 6 yields $B \mid A \gtrsim C \mid A$, with $\succ$ unless $A \mid A \sim C \mid B$. In the latter case, $B - C \in \mathscr{N}$. ◇

The next lemma (Axiom 6′, Section 5.6.3) corresponds to Axiom 4 of positive differences, the weak monotonicity condition, with an additional strict inequality proviso.

LEMMA 12. *If $A \supset B \supset C$, $A' \supset B' \supset C'$, $B \mid A \gtrsim B' \mid A'$, and $C \mid B \gtrsim C' \mid B'$, then $C \mid A \gtrsim C' \mid A'$. Moreover, the conclusion is $\succ$ unless either $C \in \mathscr{N}$ or both hypotheses are $\sim$.*

PROOF. Suppose that the hypotheses of the lemma hold, but that $C' \mid A' \gtrsim C \mid A$. We must show that in fact $C' \mid A' \sim C \mid A$ and that either $C \in \mathscr{N}$ or both hypotheses are $\sim$.

By Lemma 10, we have $C' \mid B' \gtrsim C \mid A$. By Axiom 8, choose B'' with $A \supset B'' \supset C$ and $C' \mid B' \sim B'' \mid A$. If $B'' \in \mathscr{N}$, then Lemma 9 yields that C, $C' \in \mathscr{N}$ and $C' \mid A' \sim C \mid A \sim \varnothing \mid X$, so we are done. Assume $B'' \in \mathscr{E} - \mathscr{N}$. We show that $C \mid B'' \gtrsim B \mid A$. Otherwise, $B \mid A \succ C \mid B''$ and

$$C \mid B \gtrsim C' \mid B' \sim B'' \mid A$$

yield, by Axiom 6, that $C \mid A \succ C \mid A$. But now, $C \mid B'' \gtrsim B' \mid A'$ and $B'' \mid A \sim C' \mid B'$ yield $C \mid A \gtrsim C' \mid A'$, by Axiom 6. Thus, we have $C \mid A \sim C' \mid A'$, and by the strict clause of Axiom 6, $C \mid B'' \sim B' \mid A'$. We now have $C \mid B'' \sim B \mid A$, and can now infer, by the strict clause of Axiom 6, that $B'' \mid A \sim C \mid B$ (otherwise, $C \mid A \succ C \mid A$). Hence, $\sim$ holds in both hypotheses. ◇

COROLLARY. *Suppose that C, $D \in \mathscr{E} - \mathscr{N}$ and that $D \supset C \supset A \cup B$. Then $A \mid C \gtrsim B \mid C$ iff $A \mid D \gtrsim B \mid D$.*

PROOF. If $A \mid C \gtrsim B \mid C$, then since $C \mid D \sim C \mid D$, the lemma implies $A \mid D \gtrsim B \mid D$.

Conversely, if $B \mid C \succ A \mid C$, then the strict clause in the lemma yields $B \mid D \succ A \mid D$. ◇

This corollary is one of the main results to be used in the proof of

Theorem 7. It is used first to establish the Archimedean axiom for a structure of qualitative (unconditional) probability, in Section 5.7.2, on the basis of the Archimedean axiom of Definition 8. Subsequently, it is used to show monotonicity of multiplication in the ordered local semiring. The main line of development thus consists of Lemmas 6, 7, 8 (iii), 9, 10, and 12, with corollaries.

One more lemma is required, that will be used to obtain the solvability axiom for unconditional probability.

LEMMA 13. *If* $A \mid X \succsim B \mid X$, *then there exists* $B' \subset A$ *such that* $B' \mid X \sim B \mid X$.

PROOF. By Lemma 6, $-B \mid X \succsim -A \mid X$. By Axiom 8, there exists $-B' \supset -A$ such that $-B \mid X \sim -B' \mid X$. We have $B' \subset A$ and, by Lemma 6, $B' \mid X \sim B \mid X$. ◇

5.7.2 An Additive Unconditional Representation

THEOREM 9. *Suppose that* $\langle X, \mathscr{E}, \mathscr{N}, \succsim\rangle$ *is an Archimedean structure of qualitative conditional probability (Definition 8) for which Axiom 8 holds. Let* $\succsim'$ *on* $\mathscr{E}$ *be defined by*

$$A \succsim' B \quad \text{iff} \quad A \mid X \succsim B \mid X.$$

Then $\langle X, \mathscr{E}, \succsim'\rangle$ *is an Archimedean structure of qualitative probability for which Axiom 5 of Section 5.2.3 holds.*

PROOF. We verify the five axioms.

1. $\succsim'$ is a weak order on $\mathscr{E}$, since $\succsim$ is a weak order.

2. $A \succsim' \varnothing$ by Corollary 1 of Lemma 6; $X \succ' \varnothing$ by Axiom 2 (if $X \sim' \varnothing$, then $X \in \mathscr{N}$, but $X \in \mathscr{E} - \mathscr{N}$).

3. This follows as a special case of Axiom 5 of Definition 8 (note that the strict clause of that Axiom corresponds to the "if" clause in Axiom 3 of Definition 4).

4. Suppose that $A_1', A_2', \ldots$ is a standard sequence in $\langle X, \mathscr{E}, \succsim'\rangle$ relative to some $A \succ' \varnothing$ (see Definition 3). We first construct another standard sequence $A_1, A_2, \ldots$, with the properties that, for all i, $A_i \sim A_i'$ and $A_i \subset A_{i+1}$. To do this, let $A_1 = A$ and assuming that A_{i-1} has been defined, choose A_i by applying Axiom 8. We have $A_i' \succ' A_{i-1}'$ (use Axiom 5, Definition 3, and $A \succ' \varnothing$), hence, $A_i' \mid X \succ A_{i-1} \mid X$. By Axiom 8, there exists $A_i \supset A_{i-1}$ with $A_i' \mid X \sim A_i \mid X$.

Next we show that the subsequence A_{2^i}, $i = 0, 1, \ldots$, is a standard sequence (in the sense of Definition 8) of $\langle X, \mathscr{E}, \mathscr{N}, \succsim\rangle$. Let $C_1 = A_1$ and for $i > 1$, $C_i = A_i - A_{i-1}$. The C_i are pairwise disjoint and

$$A_{2^i} = \bigcup_{k=1}^{2^i} C_k,$$

$$A_{2^{i+1}} - A_{2^i} = \bigcup_{k=2^i+1}^{2^{i+1}} C_k.$$

By Definition 3 and Axiom 5, we have $C_i \sim' A$ for all i. Thus, by Axiom 5 and induction, any two unions of m C_i's are $\sim'$, and in particular,

$$A_{2^i} \sim' A_{2^{i+1}} - A_{2^i}.$$

By Lemma 8, each $A_i \in \mathscr{E} - \mathscr{N}$. By the corollary of Lemma 12, and Axiom 4,

$$A_{2^i} \mid A_{2^{i+1}} \sim (A_{2^{i+1}} - A_{2^i}) \mid A_{2^{i+1}} \sim -A_{2^i} \mid A_{2^{i+1}}.$$

By Corollary 3 to Lemma 6, we have $A_{2^i} \mid A_{2^{i+1}} \sim A_1 \mid A_2$, all i. By Axioms 3 and 5, we have $X \mid X \sim A_2 \mid A_2 \succ A_1 \mid A_2$ (note that by Lemma 9, $C_2 \mid A_2 \succ \varnothing \mid A_2$). Thus, $\{A_{2^i}\}$ is indeed a standard sequence (Definition 8) and since it is finite, so must be the sequences $\{A_i\}$ and $\{A_i'\}$.

5. Suppose that $A \cap B = \varnothing$, $A \succ' C$, and $B \succsim' D$. Applying Lemma 13 to $A \mid X \succ C \mid X$, we have $C' \subset A$ with $C' \mid X \sim C \mid X$. Similarly, we have $D' \subset B$ with $D' \mid X \sim D \mid X$. Since $A \cap B = \varnothing$, $C' \cap D' = \varnothing$ also, and so C', D', and $E = A \cup B$ fulfill the condition. ◇

COROLLARY. *Suppose that the hypotheses of Theorem 9 hold. Let* $\mathbf{\mathscr{E}}$ *be the set of* $\sim'$ *equivalence classes, excluding* $\mathscr{N}$; *and for* $A \in \mathscr{E} - \mathscr{N}$, *let* $\mathbf{A}$ *denote the equivalence class of* A. *Let* $\boldsymbol{\succsim}$ *be the induced simple order on* $\mathbf{\mathscr{E}}$; *and if* $\mathbf{A}, \mathbf{B} \in \mathbf{\mathscr{E}}$, *with* $A \cap B = \varnothing$, *let* $\mathbf{A} \oplus \mathbf{B}$ *be* $\mathbf{A \cup B}$. *Let* $\mathscr{B}$ *be the set of* $(\mathbf{A}, \mathbf{B})$ *with* $A \cap B = \varnothing$. *Then* $\langle \mathbf{\mathscr{E}}, \boldsymbol{\succsim}, \mathscr{B}, \oplus\rangle$ *is an Archimedean, positive, regular, ordered local semigroup (Definition 2.2).*

PROOF. From Section 5.3.2 and Theorem 9 we know that $\langle \mathbf{\mathscr{E}}, \boldsymbol{\succsim}, \mathscr{B}, \oplus\rangle$ is an extensive structure with no essential maximum (Definition 3.2); but since $\boldsymbol{\succsim}$ is already a simple order, the argument in Section 3.5.3 shows that the structure is an Archimedean, positive, regular, ordered local semigroup. ◇

5.7.3 Theorem 7 (p. 224)

An Archimedean structure of qualitative probability $\langle X, \mathscr{E}, \mathscr{N}, \succsim\rangle$ *has a unique representation P such that* $\langle X, \mathscr{E}, P\rangle$ *is a finitely additive probability* *space, $\mathcal{N}$ is the class of events with probability zero, and $A \mid B \gtrsim C \mid D$ iff $P(A \cap B)/P(B) \geqslant P(C \cap D)/P(D)$.*

PROOF. We start by introducing a semiring operation into the semigroup $\langle \mathcal{E}, \gtrsim, \mathcal{B}, \oplus \rangle$ defined in the corollary to Theorem 9, in the preceding subsection. Take $\mathbf{A}, \mathbf{B} \in \mathcal{E}$. We have $A \mid X > \varnothing \mid B$ (since $\mathcal{N} \notin \mathcal{E}$ and since $\varnothing \mid X \sim \varnothing \mid B$ by Corollary 2 to Lemma 6). By Axioms 8 and 4, there exists C with $B \supset C$ and $A \mid X \sim C \mid B$. We now define $\mathbf{A} * \mathbf{B} = \mathbf{C}$.

To show that $*$ is well defined, suppose that $A \mid X \sim A' \mid X$, $B \mid X \sim B' \mid X$, $C \subset B$, $C' \subset B'$, $A \mid X \sim C \mid B$, and $A' \mid X \sim C' \mid B'$. Applying Lemma 12 to $C \subset B \subset X$, $C' \subset B' \subset X$, we have $C \mid X \sim C' \mid X$; thus, $\mathbf{C} = \mathbf{C}'$. By Lemma 9, $\mathbf{A} * \mathbf{B} \in \mathcal{E}$, since in the above definition, $C \mid B > \varnothing \mid X \sim \varnothing \mid B$.

The set $\mathcal{B}^*$ for which $*$ is defined consists of all of $\mathcal{E} \times \mathcal{E}$ (this is because the product of two numbers less than or equal to 1 is again less than or equal to 1; $\mathbf{X}$ is a natural identity for $*$). This greatly simplifies the proof of the other axioms of Definition 2.4.

To show that $*$ is associative, choose $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathcal{E}$. Let $\mathbf{A} * \mathbf{B} = \mathbf{C}'$, $C' \subset B$; and let $\mathbf{C}' * \mathbf{C} = \mathbf{D}$, $D \subset C$. Thus, $(\mathbf{A} * \mathbf{B}) * \mathbf{C} = \mathbf{D}$. Similarly, choose $\mathbf{A}' = \mathbf{B} * \mathbf{C}$, $A' \subset C$, and $\mathbf{D}' = \mathbf{A} * \mathbf{A}'$, $D' \subset A'$, so that $\mathbf{A} * (\mathbf{B} * \mathbf{C}) = \mathbf{D}'$. We have $A \mid X \sim C' \mid B$ and $A \mid X \sim D' \mid A'$, hence $D' \mid A' \sim C' \mid B$. Also $A' \mid C \sim B \mid X$. Applying Lemma 12 to $D' \subset A' \subset C$ and $C' \subset B \subset X$, we have $D' \mid C \sim C' \mid X$. Since $C' \mid X \sim D \mid C$, we have $D' \mid C \sim D \mid C$. By the corollary to Lemma 12, this implies $\mathbf{D}' = \mathbf{D}$, as required.

Before establishing that the other axioms of Definition 2.4 are satisfied, we note that $*$ is commutative: if $A \mid X \sim C \mid B$, with $C \subset B$, and $B \mid X \sim D \mid A$, with $D \subset A$, then by Axiom 6 applied to $D \subset A \subset X$, $C \subset B \subset X$, we have $D \mid X \sim C \mid X$, or $\mathbf{D} = \mathbf{C}$.

We now show that $*$ is right distributive over $\oplus$. Left distributivity then follows by commutativity of $*$. Let $A \cap B = \varnothing$, $A \cup B \mid X \sim C' \mid C$, $C' \subset C$. Then $(\mathbf{A} \oplus \mathbf{B}) * \mathbf{C} = \mathbf{C}'$. By Axiom 8, choose $A' \subset C'$ such that $A \mid A \cup B \sim A' \mid C'$. (Note that $A \cup B$, $C' \in \mathcal{E} - \mathcal{N}$.) Let $B' = C' - A'$. By Lemma 6, $B \mid A \cup B \sim B' \mid C'$. Now apply Lemma 12 to $A \subset A \cup B \subset X$, $A' \subset C' \subset C$, obtaining $A \mid X \sim A' \mid C$; and similarly, by Lemma 12, $B \mid X \sim B' \mid C$. Hence, $\mathbf{A} * \mathbf{C} = \mathbf{A}'$, $\mathbf{B} * \mathbf{C} = \mathbf{B}'$. By construction, $A' \cap B' = \varnothing$, therefore,

$$\begin{aligned}(\mathbf{A} * \mathbf{C}) \oplus (\mathbf{B} * \mathbf{C}) &= \mathbf{A}' \cup \mathbf{B}' \\ &= \mathbf{C}' \\ &= (\mathbf{A} \oplus \mathbf{B}) * \mathbf{C}.\end{aligned}$$

Finally, we note that $\mathbf{A} \gtrsim \mathbf{B}$ implies $\mathbf{A} * \mathbf{C} \gtrsim \mathbf{B} * \mathbf{C}$ follows immediately from the corollary of Lemma 12; left monotonicity follows by commutativity; also it could easily be proved directly from Lemma 12.

We now assert that the first three axioms of Definition 2.4 hold. Axiom 1, that $\langle \mathcal{E}, \gtrsim, \mathcal{B}, \oplus \rangle$ is an ordered local semigroup, was proved in the previous section; Axiom 2, that $\langle \mathcal{E}, \gtrsim, \mathcal{E} \times \mathcal{E}, * \rangle$ is an ordered local semigroup, has just been established (the assertions about pairs in $\mathcal{E} \times \mathcal{E}$ are all trivial, since $*$ is defined everywhere); and Axiom 3, distributivity, was also just proved. The final axiom asserts that for any $\mathbf{A} \in \mathcal{E}$, there exists $(\mathbf{B}, \mathbf{C})$ in $\mathcal{B}$ such that $(\mathbf{A}, \mathbf{B} \oplus \mathbf{C}) \in \mathcal{E} \times \mathcal{E}$. This is trivial unless $\mathcal{B}$ is empty. But $\mathcal{B}$ can be empty only if $\mathcal{E}$ contains exactly one element, $\mathbf{X}$; in which case, Theorem 7 is trivially true, letting $P(A) = 1$ or 0 accordingly as $A \sim' X$ or $A \sim' \varnothing$.

If $\mathcal{B}$ is nonempty, then $\langle \mathcal{E}, \gtrsim, \mathcal{B}, \mathcal{E} \times \mathcal{E}, \oplus, * \rangle$ is an Archimedean, positive, regular, ordered local semiring. By Theorem 2.6, there exists a unique function ϕ from $\mathcal{E}$ to Re$^+$ which is order preserving, additive, and multiplicative. Define

$$P(A) = \begin{cases} \phi(\mathbf{A}), & \text{if } A \in \mathcal{E} - \mathcal{N}, \\ 0, & \text{if } A \in \mathcal{N}. \end{cases}$$

We must show that P satisfies the conclusions of Theorem 7.

Suppose that $A \mid B \gtrsim C \mid D$. We want to show that $P(A \cap B)/P(B) \geqslant P(C \cap D)/P(D)$. Without loss of generality, suppose $A \cap B$, $C \cap D \in \mathcal{E} - \mathcal{N}$ (since $C \cap D \in \mathcal{N}$ is trivial, and $A \cap B \in \mathcal{N}$ implies $C \cap D \in \mathcal{N}$, using Lemma 9). Choose $E, F \in \mathcal{E} - \mathcal{N}$ such that

$$E \mid X \sim A \mid B \gtrsim C \mid D \sim F \mid X.$$

We have $\mathbf{B} * \mathbf{E} = \mathbf{A} \cap \mathbf{B}$ and $\mathbf{F} * \mathbf{D} = \mathbf{C} \cap \mathbf{D}$. Therefore, since ϕ is multiplicative and order preserving and since $\mathbf{E} \gtrsim \mathbf{F}$, we have

$$\begin{aligned}\phi(\mathbf{A} \cap \mathbf{B})/\phi(\mathbf{B}) &= \phi(\mathbf{E}) \\ &\geqslant \phi(\mathbf{F}) \\ &= \phi(\mathbf{C} \cap \mathbf{D})/\phi(\mathbf{D}).\end{aligned}$$

The required inequality for P follows immediately. The strict case $>$ similarly leads to $>$ in the P-inequality; so $A \mid B \gtrsim C \mid D$ if and only if $P(A \cap B)/P(B) \geqslant P(C \cap D)/P(D)$.

We need only show now that $\langle X, \mathcal{E}, P \rangle$ is a finitely additive probability space. Since $\mathbf{X}$ is a multiplicative identity, we have $P(X) = 1$; clearly $P \geqslant 0$; and additivity of P follows from additivity of ϕ on $\langle \mathcal{E}, \mathcal{B}, \oplus \rangle$.

For uniqueness, note that another representation P' would lead to another representation ϕ' of $\langle \mathcal{E}, \gtrsim, \mathcal{B}, \mathcal{E} \times \mathcal{E}, \oplus, * \rangle$, by letting $\phi'(\mathbf{A}) = P'(A)$. For example, to show that ϕ' is multiplicative, let $\mathbf{A} * \mathbf{B} = \mathbf{C}$. We can choose C such that $A \mid X \sim C \mid B$ and $C \subset B$; and then, from the representation, $P'(A)\,P'(B) = P'(C)$, implying $\phi'(\mathbf{A} * \mathbf{B}) = \phi'(\mathbf{A})\phi'(\mathbf{B})$. ◇

5.7.4 Theorem 8 (p. 228)

If $\langle X, \mathscr{E}, \mathscr{N}, \succsim\rangle$ satisfies Axioms 1–4, 5′, 6′, 7, and 8′ (Section 5.6), then there exists $Q: (\mathscr{E} - \mathscr{N}) \to (0, 1]$, such that (i) $Q(X) = 1$ *and* (ii) $A \mid B \succsim C \mid D$ *iff* $Q(A \cap B)/Q(B) \geqslant Q(C \cap D)/Q(D)$. *This Q is unique up to $Q \to Q^{\alpha}$, $\alpha > 0$.*

PROOF. We construct a positive-difference structure $\langle \mathscr{E} - \mathscr{N}, \mathscr{A}^*, \succsim^*\rangle$ (see Definition 4.1) as follows. Let $AB \in \mathscr{A}^*$ iff $A, B \in \mathscr{E} - \mathscr{N}$, $A \supset B$, and $X \mid X \succ B \mid A$. For AB, $CD \in \mathscr{A}^*$, let $AB \succsim^* CD$ iff $D \mid C \succsim B \mid A$. We verify the six axioms of Definition 4.1.

1. Weak ordering is immediate.

2 & 3. If $AB, BC \in \mathscr{A}^*$, then clearly, A, $C \in \mathscr{E} - \mathscr{N}$ and $A \supset C$. We shall show that $B \mid A$ and $C \mid B$ are both $\succ C \mid A$; this will show that $X \mid X \succ C \mid A$, so $AC \in \mathscr{A}^*$, and also, that AB, $BC \prec^* AC$.

By the strict clause of Axiom 6′, $B \mid B \succ C \mid B$ and $B \mid A \sim B \mid A$ yield $B \mid A \succ C \mid A$. Similarly, $C \mid B \sim C \mid B$ and $B \mid B \succ B \mid A$ yield $C \mid B \succ C \mid A$. (Note that Axioms 1 and 3 were used implicitly in the above proof.)

4. This follows immediately from Axiom 6′.

5. Suppose that AB, $CD \in \mathscr{A}^*$, with $AB \succ^* CD$. By Axiom 8′, applied to $D \mid C \succ B \mid A$, we have C' and D', with $A \supset C'$, D' and C', $D' \supset B$, such that

$$B \mid C' \sim D \mid C \sim D' \mid A.$$

By Axiom 5′ or 8′, $C' \in \mathscr{E} - \mathscr{N}$; by Axiom 5′, $D' \in \mathscr{E} - \mathscr{N}$. Thus, AD', $C'B \in \mathscr{A}^*$, and $AD' \sim^* CD \sim C'B$. To complete the proof we must show that AC', and $D'B \in \mathscr{A}^*$, i.e., that $X \mid X \succ$ both $C' \mid A$ and $B \mid D'$. If $X \mid X \sim C' \mid A$, then apply the strict clause of Axiom 6′ to $B \mid C' \succ B \mid A$, $C' \mid A \sim A \mid A$, to obtain $B \mid A \succ B \mid A$, a contradiction. Similarly, $B \mid D' \sim B \mid B$ and $D' \mid A \succ B \mid A$ would yield $B \mid A \succ B \mid A$; hence, $X \mid X \succ B \mid D'$.

6. $\{A_i\}$ is a standard sequence, in the sense of Definition 4.1, if and only if it is a standard sequence in the sense of Definition 8. (This is immediate from the definitions.) Thus, any standard sequence in $\langle \mathscr{E} - \mathscr{N}, \mathscr{A}^*, \succsim^*\rangle$ is finite.[2]

[2] The reader may be puzzled by the fact that all standard sequences are finite. This is because all standard sequences are ascending, hence, all are strictly bounded by XA_1. If we permitted descending standard sequences, $A_1, A_2, \ldots$, such that $A_i \supset A_{i+1}$ and $A_{i+1} \mid A_i \sim A_i \mid A_{i-1}$, there could very well be infinite standard sequences. But it would still be true that any strictly bounded standard sequence is finite. By solvability, any strictly bounded descending standard sequence could be converted into an ascending standard sequence. This is really the reason why only one direction needs to be considered.

This completes the verification of the axioms for positive differences. By Theorem 4.1, there exists a positive-valued function ψ on $\mathscr{E} - \mathscr{N}$ such that $AB \succsim^* CD$ iff $\psi(AB) \geqslant \psi(CD)$, and for AB, $BC \in \mathscr{A}^*$, $\psi(AC) = \psi(AB) + \psi(BC)$. We define a function Q as

$$Q(A) = \begin{cases} 1, & \text{if } X \mid X \sim A \mid X, \\ \exp[-\psi(XA)], & \text{if } XA \in \mathscr{A}^*. \end{cases}$$

We must show that Q is well defined and that its domain is $\mathscr{E} - \mathscr{N}$; that is, the two conditions in its definition are mutually exclusive, and $A \in \mathscr{E} - \mathscr{N}$ iff one of them holds.

By definition of $\mathscr{A}^*$, $XA \in \mathscr{A}^*$ excludes both $X \mid X \sim A \mid X$ and $A \in \mathscr{N}$. Also, $X \mid X \sim A \mid X$ excludes $A \in \mathscr{N}$ (otherwise, $X \mid X \sim A \mid X \sim \varnothing \mid X$ yields $X \in \mathscr{N}$). Conversely, if $A \notin \mathscr{N}$ and $X \mid X \succ A \mid X$, then $XA \in \mathscr{A}^*$. Thus, Q is well defined on $\mathscr{E} - \mathscr{N}$.

Next we note that if $AB \in \mathscr{A}^*$, then

$$Q(B)/Q(A) = \exp[-\psi(AB)].$$

For if $XA \in \mathscr{A}^*$, then $XB \in \mathscr{A}^*$, so

$$\begin{aligned} Q(B)/Q(A) &= \exp[-\psi(XB) + \psi(XA)] \\ &= \exp[-\psi(AB)]. \end{aligned}$$

Or if $XA \notin \mathscr{A}^*$, then $X \mid X \sim A \mid X$, and by Axiom 6′, $B \mid X \sim B \mid A$ (otherwise, the strict clause of Axiom 6′ yields $B \mid X \succ B \mid X$). Consequently, $XB \in \mathscr{A}^*$ and $\psi(AB) = \psi(XB)$; the equation follows, since $Q(A) = 1$.

We can now show that (ii) of Theorem 8 is satisfied. First, if AB, $CD \in \mathscr{A}^*$, then the fact that $\exp[-\psi(AB)]$ is a strictly decreasing function of $\psi(AB)$ gives the required result (since ψ is order preserving on $\mathscr{A}^*$, and the ordering $\succsim^*$ on $\mathscr{A}^*$ is the converse of the corresponding $\succsim$ ordering). This extends easily to $A, B, C, D \in \mathscr{E} - \mathscr{N}$ with $A \supset B$, $C \supset D$, but AB or $CD \notin \mathscr{A}^*$. For if $AB \notin \mathscr{A}^*$, then $X \mid X \sim B \mid A$; but this entails $A \mid X \sim B \mid X$. (If $A \mid X \succ B \mid X$, then $B \mid A \sim B \mid B$ and Axiom 6′ yield $B \mid X \succ B \mid X$; etc.) Hence, $Q(A) = Q(B)$. Similarly, if $CD \notin \mathscr{A}^*$, $X \mid X \sim D \mid C$ and $Q(C) = Q(D)$. Thus, if both are not in $\mathscr{A}^*$, $B \mid A \sim D \mid C$ and $Q(B)/Q(A) = 1 = Q(D)/Q(C)$; if only $AB \notin \mathscr{A}^*$, then $B \mid A \succ D \mid C$ and $Q(B)/Q(A) = 1 > \exp[-\psi(CD)] = Q(D)/Q(C)$. Finally, the representation extends to A, B, C, D arbitrary in $\mathscr{E} - \mathscr{N}$ by using the representation for $A \cap B$, A, $C \cap D$, C, and employing Axiom 4.

To prove uniqueness, note that if Q' is another representation, then $\psi'(AB) = -\log[Q'(B)/Q'(A)]$ gives another representation for $\langle \mathscr{E} - \mathscr{N},$

$\mathcal{A}^*, \succsim^*\rangle$, hence, by Theorem 4.1, $\psi' = \alpha\psi$. Thus,

$$\begin{aligned} Q'(A) &= Q'(A)/Q'(X) \\ &= \exp[-\psi'(XA)] \\ &= \exp[-\alpha\psi(XA)] \\ &= [Q(A)]^\alpha, \end{aligned}$$

if $X \mid X \succ A \mid X$. Since $1^\alpha = 1$, we have $Q' = Q^\alpha$ on $\mathscr{E} - \mathscr{N}$. ◇

5.8 INDEPENDENT EVENTS

Intimately connected with the concept of conditional probability is that of (statistical) independence. Given a probability measure P, events A, B are independent if and only if $P(A \cap B) = P(A)\,P(B)$ [or, for $P(B) \neq 0$, if and only if $P(A \mid B) = P(A)$]. In qualitative terms, using the primitives of the preceding sections, we can define A, B to be $\sim$-independent if and only if either B is in $\mathscr{N}$ or $A \mid B \sim A \mid X$. (Or again, using $*$ defined in Section 5.7, A, B are $\sim$-independent when $\mathbf{A} * \mathbf{B} = \mathbf{A} \cap \mathbf{B}$.)

An interesting idea, introduced by Domotor (1969), is to accept independence of events as a primitive notion, in which case the primitives for a theory of conditional probability consist of a set X, an algebra $\mathscr{E}$ on X, an ordering $\succsim^*$ on $\mathscr{E}$, and a new binary relation (interpreted as independence) $\perp$ on $\mathscr{E}$. Of course, Theorem 2 is then to be enriched by adding to the constraints on the representation the following one:

$$A \perp B \quad \text{iff} \quad P(A \cap B) = P(A)\,P(B).$$

It will be noted that this property of the representation together with the axiom $A \perp X$ (see Definition 9 below) implies $P(X) = 1$. Without such an additional constraint, we can obtain an absolute scale of probability only by the fiat of normalization. From the standpoint of unconditional probability, only ratios of probabilities are meaningful. This further constraint, based on the new primitive $\perp$, yields a genuine absolute scale. Another way to understand the absolute scale is from the standpoint of conditional probability: in Section 5.7.3, $\mathbf{X}$ is the natural multiplicative identity with respect to $*$, forcing $P(X) = 1$.

We start by introducing the axioms of a pure independence structure $\langle X, \mathscr{E}, \perp\rangle$.

DEFINITION 9. *Suppose $\mathscr{E}$ is an algebra of sets on X and $\perp$ is a binary*

relation on $\mathscr{E}$. Then $\perp$ is an independence relation[3] *iff*

1. *$\perp$ is symmetric.*
2. *For $A \in \mathscr{E}$, the set $\{B \mid B \in \mathscr{E}$ and $A \perp B\}$ is a QM-algebra of sets on X (Definition 5), i.e.,*

   (i) *$A \perp X$;*

   (ii) *if $A \perp B$, then $A \perp -B$; and*

   (iii) *if $A \perp B$, $A \perp C$, and $B \cap C = \varnothing$, then $A \perp B \cup C$.*

It is easy to demonstrate that these axioms are necessary for the representation, e.g., for 2 (ii), if $P(A \cap B) = P(A)\,P(B)$, then

$$\begin{aligned} P(A \cap -B) &= P(A) - P(A \cap B) \\ &= P(A) - P(A)\,P(B) \\ &= P(A)[1 - P(B)] \\ &= P(A)\,P(-B). \end{aligned}$$

It is definitely not true that $\{B \mid A \perp B\}$ is an ordinary subalgebra of $\mathscr{E}$; there can be events A, B, and C for which $A \perp B$ and $A \perp C$ but not $A \perp B \cap C$ and not $A \perp B \cup C$. This corresponds to the fact that pairwise independence of events, in the usual probabilistic sense, does not entail independence of the entire set of three or more events. Only if A, B, C form an independent family of sets is each event independent of the subalgebra generated bv the other events. This motivates the definition of an abstract independence relation for sets of three or more events, based on $\perp$, using subalgebras.

If $\mathscr{S}$ is a subset of $\mathscr{E}$, then the *smallest subalgebra containing* $\mathscr{S}$ is given as the intersection of all subalgebras of $\mathscr{E}$ that contain $\mathscr{S}$. (Note that $\mathscr{E}$ itself is a subalgebra containing $\mathscr{S}$, so this intersection is defined and it contains $\mathscr{S}$; and the intersection of arbitrarily many subalgebras is easily seen to be a subalgebra.) The smallest subalgebra containing A consists of $\{\varnothing, A, -A, X\}$.

Two subalgebras of $\mathscr{E}$, $\mathscr{E}_1$ and $\mathscr{E}_2$ are said to be *$\perp$-independent* if and only if $A_1 \perp A_2$ whenever A_1 is in $\mathscr{E}_1$ and A_2 is in $\mathscr{E}_2$. Note that if $A \perp B$, then $\{\phi, A, -A, X\}$ and $\{\varnothing, B, -B, X\}$ are $\perp$-independent. This leads to the generalization: $A_1, \ldots, A_m$ are $\perp$-independent if and only if the smallest subalgebras containing every proper subset of the A_i and its complement are $\perp$-independent. Formally, we have:

DEFINITION 10. *Let $\mathscr{E}$ be an algebra of sets and $\perp$ an independence*

[3] This concept is different from that of a relation being independent implicitly defined in Definition 1.3 of Section 1.3.2.

relation on $\mathscr{E}$. For $m \geqslant 2$, $A_1, \ldots, A_m \in \mathscr{E}$ are $\perp$-independent iff, for every proper subset M of $\{1, \ldots, m\}$, every B in the smallest subalgebra containing $\{A_i \mid i \in M\}$, and every C in the smallest subalgebra containing $\{A_i \mid i \notin M\}$, we have $B \perp C$.

As we noted in motivating this definition, if $m = 2$, we have A and B are $\perp$-independent if and only if $A \perp B$.

We can now define what we mean by a structure of qualitative probability with independence. We take a structure of qualitative probability $\langle X, \mathscr{E}, \succsim \rangle$; an independence relation $\perp$ on $\mathscr{E}$; and an axiom interlocking $\succsim$ and $\perp$. The interlocking axiom is very natural, given the representation: probabilities of independent events multiply, and multiplication is order preserving, therefore, intersections of independent events preserve order. In addition, we include in the following definition a very strong structural condition that uses the defined notion of $\perp$-independence of sets of events.

DEFINITION 11. *Suppose that X is a nonempty set, $\mathscr{E}$ is an algebra of sets on X, and $\succsim$ and $\perp$ are binary relations on $\mathscr{E}$. The quadruple $\langle X, \mathscr{E}, \succsim, \perp \rangle$ is a* structure of qualitative probability with independence *iff the following axioms hold:*

1. *$\langle X, \mathscr{E}, \succsim \rangle$ is a structure of qualitative probability (Definition 4).*
2. *$\perp$ is an independence relation (Definition 9).*
3. *Suppose that $A, B, C, D \in \mathscr{E}$, $A \perp B$, and $C \perp D$. If $A \succsim C$ and $B \succsim D$, then $A \cap B \succsim C \cap D$; moreover, if $A \succ C$, $B \succsim D$, and $B \succ \varnothing$, then $A \cap B \succ C \cap D$.*

The structure $\langle X, \mathscr{E}, \succsim, \perp \rangle$ is complete *iff the following additional axiom holds:*

4. *For any $A_1, \ldots, A_m, A \in \mathscr{E}$, there exists $A' \in \mathscr{E}$ with $A' \sim A$ and $A' \perp A_i$, $i = 1, \ldots, m$. Moreover, if $A_1, \ldots, A_m$ are $\perp$-independent, then A' can be chosen so that $A_1, \ldots, A_m, A'$ are also $\perp$-independent.*

In a complete structure of qualitative probability with independence, we can introduce a defined relation $\succsim'$ on pairs $A \mid B$ and $C \mid D$. Clearly, if $A \subset B$, $C \subset D$, $A \perp D$, and $C \perp B$, we will want to assert that $A \mid B \succsim' C \mid D$ if and only if $P(A)\,P(D) \geqslant P(B)\,P(C)$, or, in qualitative terms, when $A \cap D \succsim C \cap B$. If the requisite independence relations for A, D and C, B do not hold, then using completeness (Axiom 4), we may replace these sets by the equivalent ones for which independence does hold. The theorem we obtain is that, with suitable definitions of $\mathscr{N}$ and $\succsim'$, a complete structure of qualitative probability with independence gives rise to a structure of qualitative conditional probability.

DEFINITION 12. *Suppose $\langle X, \mathscr{E}, \succsim, \perp \rangle$ is a structure of qualitative probability with independence. Let $\mathscr{N} = \{A \mid A \in \mathscr{E}, A \sim \varnothing\}$. If $A, C \in \mathscr{E}$ and $B, D \in \mathscr{E} - \mathscr{N}$, define*

$$A \mid B \succsim' C \mid D$$

iff there exist $A', B', C', D' \in \mathscr{E}$, with

$$A' \sim A \cap B, \qquad B' \sim B, \qquad C' \sim C \cap D, \qquad D' \sim D;$$
$$A' \perp D' \quad \text{and} \quad C' \perp B';$$

and

$$A' \cap D' \succsim C' \cap B'.$$

THEOREM 10. *Suppose that $\langle X, \mathscr{E}, \succsim, \perp \rangle$ is a complete structure of qualitative probability with independence and that $\mathscr{N}$ and $\succsim'$ are given by Definition 12. Then $\langle X, \mathscr{E}, \mathscr{N}, \succsim' \rangle$ is a structure of qualitative conditional probability (Definition 8).*

If we now make the additional assumptions that the Archimedean and solvability axioms of Section 5.6 are satisfied by $\langle X, \mathscr{E}, \mathscr{N}, \succsim' \rangle$, then Theorem 10 and these assumptions allow us to deduce the representation P of Theorem 7. This gives the desired representation for $\perp$. For if $A \perp B$, with B not in $\mathscr{N}$, then Definition 12 gives $A \mid B \sim' A \mid X$, hence,

$$P(A \cap B)/P(B) = P(A).$$

(If A, B are both $\sim \varnothing$, then $A \cap B \sim \varnothing$, so the desired result is trivial.)

Of course, the Archimedean and solvability axioms, stated in terms of $\succsim'$, can be translated into terms of $\succsim$ and $\perp$. The resulting axioms are not very natural. A simple axiomatization directly in terms of $\succsim$ and $\perp$ would be preferable.

5.9 PROOF OF THEOREM 10

A complete structure of conditional probability with independence induces, by Definition 12, a structure of qualitative conditional probability.

PROOF. We remark first that by Axiom 3 of Definition 11, the definition of $\succsim'$ is independent of the particular choice of A', B', C', D' in Definition 12. For if $A' \sim A''$, $B' \sim B''$, $C' \sim C''$, and $D' \sim D''$, with $A' \perp D'$, $A'' \perp D''$, $C' \perp B'$, and $C'' \perp B''$, then $A' \cap D' \sim A'' \cap D''$ and $C' \cap B' \sim C'' \cap B''$.

By completeness (Axiom 4), $\succsim'$ is a connected relation; for let $A' = A \cap B$, $C' = C \cap D$, and use the axiom twice with $m = 1$ to obtain $D' \sim D$ with $A' \perp D'$ and $B' \sim B$ with $C' \perp B'$; then compare $A' \cap D'$ with $C' \cap B'$.

To show that $\succsim'$ is transitive, use completeness successively with $m = 1, 2, 3, 4,$ and 5 to obtain $A' = A \cap B$, $B' \sim B$, $C' \sim C \cap D$, $D' \sim D$, $E' \sim E \cap F$, and $F' \sim F$, such that A', B', C', D', E', F' are $\perp$-independent. By the remark above on the definition of $\succsim'$, we infer, from $A \mid B \succsim' C \mid D$ and $C \mid D \succsim' E \mid F$, that $A' \cap D' \succsim C' \cap B'$ and $C' \cap F' \succsim D' \cap E'$. By the definition of $\perp$-independence, we have

$$A' \cap D' \perp C' \cap F', \qquad C' \cap B' \perp D' \cap E'.$$

Therefore, Axiom 3 yields

$$A' \cap F' \cap C' \cap D' \succsim E' \cap B' \cap C' \cap D'.$$

If not $A \mid B \succsim' E \mid F$, then $E' \cap B' \succ A' \cap F'$. If $C' \cap D' \succ \varnothing$, then the strict clause of Axiom 3 yields $E' \cap B' \cap C' \cap D' \succ A' \cap F' \cap C' \cap D'$ (since $\perp$-independence gives $E' \cap B' \perp C' \cap D'$ and $A' \cap F' \perp C' \cap D'$). This contradiction shows that $A \mid B \succsim E \mid F$ as required. On the other hand, if $C' \cap D' \sim \varnothing$, then the strict clause of Axiom 3, applied to $C' \perp D'$, $\varnothing \perp \varnothing$, and $D' \succ \varnothing$ $(D' \in \mathscr{E} - \mathscr{N})$ yields $C' \sim \varnothing$; thence,

$$C' \cap F' \sim \varnothing \succsim D' \cap E';$$

and since $D' \succ \varnothing$, the same arguments give $E' \sim \varnothing$ and $E' \cap B' \sim \varnothing$. Thus, in this case also, $A' \cap F' \succsim E' \cap B'$, as required.

The second axiom of Definition 8, that $X \in \mathscr{E} - \mathscr{N}$ and $A \in \mathscr{N}$ iff $A \mid X \sim' \varnothing \mid X$, is immediate; obviously, $A \mid X \sim' B \mid X$ iff $A \sim B$. Similarly, the third axiom of Definition 8, that $X \mid X \succsim' A \mid B$ and $X \mid X \sim' A \mid A$, is immediate; and the fourth axiom, $A \mid B \sim' A \cap B \mid B$, follows from the definition of $\succsim'$. Thus, only the additivity and strong 6-tuple conditions need to be established.

Suppose that $A \mid C \succsim' D \mid F$ and $B \mid C \succsim' E \mid F$, where $A \cap B = \varnothing = D \cap E$. Without loss of generality, assume $A \cup B \subset C$, $D \cup E \subset F$. By completeness, choose $C' \sim C$ with $C' \perp D, E$ and choose $F' \sim F$ with $F' \perp A, B$. By definition of $\succsim'$ we have

$$A \cap F' \succsim D \cap C'$$
$$B \cap F' \succsim E \cap C'.$$

By Lemma 2, which uses only qualitative probability,

$$(A \cup B) \cap F' \succsim (D \cup E) \cap C',$$

and if either antecedent inequality is strict, so is the conclusion. By the disjoint union property of $\perp$ [Axiom 2 (iii) of Definition 9] we have $C' \perp D \cup E$ and $F' \perp A \cup B$. Hence, $A \cup B \mid C \succsim' D \cup E \mid F$, with $\succ'$ holding if either antecedent is strict.

Finally, suppose that $A \supset B \supset C$, $D \supset E \supset F$, $B \mid A \succsim F \mid E$ and $C \mid B \succsim' E \mid D$. Choose A', B', C', D', E', F', respectively $\sim A, B, C, D, E, F$ and $\perp$-independent (completeness with $m = 1, 2, 3, 4, 5$). We have $B' \cap E' \succsim F' \cap A'$ and $C' \cap D' \succsim E' \cap B'$ (with strict $\succsim'$ going to strict $\succsim$). By transitivity of $\succsim$, $C' \cap D' \succsim F' \cap A'$ (or strict inequality, if there is a strict antecedent), hence $C \mid A \succsim$ (or $\succ$, as required) $F \mid D$. ◇

EXERCISES

1. For each of the following three families $\mathscr{E}$ of sets, either show that $\mathscr{E}$ is an algebra of sets (Definition 1) or show a violation of one of the axioms.

(i) Let X be a nonempty set and A a nonempty subset of X. Then $\mathscr{E}$ consists of all subsets of X that either include A or are disjoint from A.

(ii) Let X be a nonempty set and A a nonempty subset of X. Then $\mathscr{E}$ consists of $\varnothing$ and all subsets of X that intersect both A and $-A$.

(iii) Let $\mathscr{F}$ and $\mathscr{G}$ be algebras of sets on X and Y, respectively. Then $\mathscr{E}$ consists of those subsets of $X \cup Y$ of the form $A \cup B$ where A is in $\mathscr{F}$ and B is in $\mathscr{G}$. (5.1)

2. Suppose that $\langle X, \mathscr{E}, P\rangle$ is a finitely additive probability space (Definition 2). Prove that, for all $A, B \in \mathscr{E}$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{5.1}$$

3. Prove the independence of the axioms of qualitative probability (Definition 4). (5.2.1)

4. In a structure of qualitative probability (Definition 4), either prove that the relation $\sim^*$ (defined on p. 206 in connection with the definition of tight) is an equivalence relation or give a counterexample. (5.2.3)

5. Suppose that $\langle X, \mathscr{E}, \succsim\rangle$ is a structure of qualitative probability (Definition 4) and that $A, B, C, D \in \mathscr{E}$ are such that $A \cup C = B \cup D = X$. Prove that if $A \succsim B$ and $C \succsim D$, then $A \cap C \succsim B \cap D$. (5.2.1, 5.3.1)

6. Suppose that $\langle X, \mathscr{E}, \succsim\rangle$ is a structure of qualitative probability (Definition 4) and that $\mathscr{N} = \{A \mid A \in \mathscr{E}$ and $A \sim \phi\}$. Prove that $\mathscr{N}$ has the following properties:

(i) If $A \in \mathscr{N}$, $B \in \mathscr{E}$, and $B \subset A$, then $B \in \mathscr{N}$.

(ii) If $A, B \in \mathscr{N}$, then $A \cup B$, $A \cap B \in \mathscr{N}$.

(iii) If $A \in \mathscr{N}$, then $-A \in \mathscr{E} - \mathscr{N}$. (5.2.1, 5.3.1)

7. Let $X = \{a, b, c, d, e\}$ and $\mathscr{E} = 2^X$. Suppose that $\succsim$ is a weak ordering of $\mathscr{E}$ such that:

(i) $\{a\} \sim \{b\}$;

(ii) $\{c\} \sim \{d\}$;

(iii) $\{a, b\} \sim \{c, d, e\}$;

(iv) $\{a, e\} \sim \{c, d\}$.

Find a probability representation consistent with $\succsim$. Is it unique? (5.4.3, 5.5.3)

8. Does the probability space in Exercise 7 satisfy the hypotheses of Theorem 6? Can you define general circumstances in which a finite structure of qualitative probability has a unique representation? (5.4.3, 9.2.2).

9. Suppose that if there exists an A with $0 < P(A) < 1$, then $\langle X, \mathscr{E}, P\rangle$ is a finitely additive probability space (Definition 2) with the following property: for every A in $\mathscr{E}$ with $0 < P(A) < 1$, there exists a B in $\mathscr{E}$ with $0 < P(B) < 1$ that is independent of A (p. 238). Prove that this space has no atoms (p. 216). (5.1, 5.4.2, 5.8)

In the following three exercises we suppose that $\langle X, \mathscr{E}, \mathscr{N}, \succsim\rangle$ *is a structure of qualitative conditional probability (Definition* 8, *Section* 5.6.1). *In the proofs you may use the axioms of Definition* 8 *and the lemmas of Section* 5.7.1, *but not Theorem* 7.

10. If $A \in \mathscr{E}$, $B, C \in \mathscr{E} - \mathscr{N}$, $A \quad B = A \quad C$, and $B \quad C$, then $A \mid B \succsim A \mid C$. (5.6.1, 5.7.1)

11. If $A, A', B, B' \in \mathscr{E}$, $C, D \in \mathscr{E} - \mathscr{N}$, $A \mid C \succsim A' \mid D$, and $B \mid C \precsim B' \mid D$, then $(A - B) \mid C \succsim (A' - B') \mid D$. (5.6.1, 5.7.1)

12. Suppose that $A, B, C \in \mathscr{E}$ and that $C \cap B$, $C - B \in \mathscr{E} - \mathscr{N}$. Prove that $A \mid C \sim A \mid C \cap B$ iff $A \mid C \sim A \mid (C - B)$. Thus, letting $C = X$, if A is independent of B (see Section 5.8), then $A \mid B \sim A \mid -B$. (5.6.1, 5.7.1, 5.8)

13. Prove Exercises 10–12 using the representation of Theorem 7. (5.6.2)

Chapter 6 Additive Conjoint Measurement

□ 6.1 SEVERAL NOTIONS OF INDEPENDENCE

Not all attributes that we wish to measure have an internal, additive structure of the sort we studied in Chapters 3 and 5 (temperature and density are just two examples), and when no extensive concatenation is apparent we must use some other structure in measuring the attribute. One of the most common alternative structures is for the underlying entities to be composed of two or more components, each of which affects the attribute in question. A simple example is the attribute momentum which is exhibited by physical objects and which is affected both by their mass and by their velocity. Another example, discussed in Section 1.3.2, is the judged comfort of various humidity and temperature combinations.

The objects in these examples cannot readily be concatenated, nevertheless, they can be treated as composite entities. The rest of this volume, except for the last chapter, concerns theories leading to construction of measurement scales for composite objects which preserve their observed order with respect to the relevant attribute (e.g., preference, comfort, momentum) and where the scale value of each object is a function of the scale values of its components. Since such theories lead to simultaneous measurement of the objects and their components, they are called conjoint measurement theories. The present chapter deals with the additive case; more complicated rules of combination are studied in Chapters 7, 8, and 9.

The first part of this chapter (Sections 6.1–6.3) deals with the two-component case. Empirical examples from physics and psychology are presented