
The Representation of Popper Measures

Wolfgang Spohn

ABSTRACT. The paper shows that all and only σ -additive Popper measures can be, indeed uniquely, represented by so called dimensionally well-ordered families of σ -additive probability measures.

Probabilities conditional on events having probability 0 cannot be defined within standard probability theory. Many considered this to be a serious defect and tried to do better. Essentially four ways to deal with this problem have emerged.

One method is due to Carnap; it consists in imposing his regularity condition on probability measures (cf. Carnap (1971, 1980), Section 7 and 21). However, this does not solve the problem; it is only a way of avoiding it as far as possible within standard probability theory. But the problem is unavoidable for measures on uncountable σ -fields.

The second method is the standard one within mathematics. It consists in generalizing the concept of conditional probability, i.e. in defining probabilities (and expectations) conditional on σ -fields (cf., e.g., Loève (1960), §§ 24, 25). This method meets most of the needs of applied mathematics and should make us aware of the fact that the problem is a rather theoretical one. However, the probabilities conditional on σ -fields are determined only almost uniquely (in the technical sense). Thus, I think, this method does not count as a fundamental solution of our problem.

Another method is to do probability theory within the framework of nonstandard analysis (cf. Loeb, 1979). Within this framework division by infinitesimals is defined. Thus, if non-empty events get at least infinitesimal instead of zero probability, probabilities conditional on them can also be defined. Without doubt, this method should be taken seriously. But nonstandard analysis is a very intricate matter, and I won't dwell upon this method here.

The fourth method, developed independently by Popper and Rényi, is the one I want to deal with. It consists in construing conditional probability not as a derived, but as a fundamental concept and in looking for a suitable set of axioms for it; absolute probabilities are then defined as probabilities conditional on the sure event. The simplest axiomatization that has emerged is comprised in the following definition:

DEFINITION 1. Let Ω be a non-empty set and \mathcal{A} , the set of *events*, a σ -field of subsets of Ω . Then $\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$ is called a *conditional probability space (c.p.s.)* iff \mathcal{B} , the set of *conditions*, is a non-empty subset of $\mathcal{A} \setminus \{\emptyset\}$ and P is a function from $\mathcal{A} \times \mathcal{B}$ into the closed interval $[0, 1]$ such that the following holds:

- (a) for each $B \in \mathcal{B}$ the function $P(\cdot | B)$ is a σ -additive probability measure on \mathcal{A} with $P(B | B) = 1$,
- (b) for each $A, B, C \in \mathcal{A}$ with $C, B \cap C \in \mathcal{B}$ we have

$$P(A \cap B | C) = P(A | B \cap C) \cdot P(B | C).$$

$\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$ is called an *additive conditional probability space* iff we have moreover:

- (c) if $A, B \in \mathcal{B}$, then also $A \cup B \in \mathcal{B}$.
- $\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$ is a *Popper space* iff in addition to (a) and (b) we have
- (d) for each $A \in \mathcal{A}$, if $P(A | B) > 0$ for some $B \in \mathcal{B}$, then $A \in \mathcal{B}$.

Finally, let's call $\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$ a *full conditional probability space* iff in addition to (a) and (b)

- (e) $\mathcal{B} = \mathcal{A} \setminus \{\emptyset\}$.

The function P of a c.p.s. is called, respectively, a (*additive, full*) *conditional probability measure (c.p.m.)* or a *Popper measure*.

Obviously, every full c.p.s. is a Popper space, and every Popper space is an additive c.p.s.

Some history: The definition of c.p.s.s is due to Rényi (1955). It scarcely needs any explanation; the point of clause (b) is, of course, to ensure the compatibility of probabilities with respect to varying conditions. The concept of an additive c.p.s. has also been invented by Rényi; its role will become clear in Theorem 1 below. (This use of 'additive' has nothing to do with the usual additivity of measures.) Popper spaces have, of course, been introduced by Popper; the axioms given by him in Popper (1959), new appendices *ii–*v, have been simplified by Stalnaker (1970), Harper (1976), and van Fraassen (1976). I have taken the axioms stated by van Fraassen (1976), p. 420.

In full c.p.s.s finally, there is no restriction to conditionalization at all.

I think the concept of a c.p.s. is an essentially satisfying solution to our problem, but I won't argue this, since I have no new reason for thinking so. The four concepts of a c.p.s. introduced above differ only in their demands on the set of conditions. I am not going to decide, which of these is the most preferable. But it is clear that the larger the set of conditions the fuller the solution to our problem.

C.p.m.s thus being useful and interesting, it is natural to ask how they relate to or may be constructed from usual (probability) measures. It is this question I want to deal with in this paper. For additive c.p.m.s, this question has been answered by Császár (1955); I shall reproduce his answer for introductory reasons as well as for the sake of completeness. With respect to Popper measures, I first thought that this question had been completely settled by van Fraassen (1976), so that nothing would be left to say about it. However, this is not quite so.

Van Fraassen proves there that Popper measures may be represented by well-ordered families of probability measures. But his representation has two minor blemishes. First, the representation given by him is not unique; the closest to uniqueness he comes to is the representation by what he calls minimal families of probability measures; but only Popper spaces which satisfy a rather restrictive condition and which he then calls dimensional can be represented by such minimal families (cf. van Fraassen (1976), p. 426). Secondly, the well-orderedness of a representing family is *not* substantial; what is essentially embodied in the representation is only a linear ordering of such a family (cf. p. 427, lines 5–7, and the theorem on p. 427).

As will become clear, these blemishes are due to the fact that van Fraassen is more generally concerned with finitely additive Popper measures (where clause (a) of Definition 1 is weakened to the effect that each $P(\cdot|B)$ is only finitely additive). Here every kind of measure will always be understood to be σ -additive; and restricting ourselves in such a way, we are able to establish more elegant representations. It will also be clear that, despite these differences, the following has been strongly inspired by van Fraassen's paper.

What might it mean to represent or to generate a c.p.s.? A first idea is given by

DEFINITION 2. A family $(\mu_i)_{i \in I}$ of measures on \mathcal{A} (i.e. of non-negative, σ -additive set functions on \mathcal{A} , possibly taking ∞ as value) is said to *generate* the c.p.s. $\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$ iff the following holds:

- (a) for each $B \in \mathcal{B}$ there is a $i \in I$ such that $0 < \mu_i(B) < \infty$,
- (b) for each $A \in \mathcal{A}$, $B \in \mathcal{B}$, and $i \in I$ with $0 < \mu_i(B) < \infty$ we have

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

However, this notion of generation is still uninformative; trivially, each c.p.s. can be generated in such a way, as has already been observed by Császár (1955), p. 340. Simply define for each $B \in \mathcal{B}$ the measure μ_B by $\mu_B(A) = P(A|B)$ for $A \subseteq B$ and, else, $\mu_B(A) = \infty$; the family $(\mu_B)_{B \in \mathcal{B}}$ then generates $\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$. In order to get more interesting representations, restrictions have to be imposed on generating families:

DEFINITION 3. The family $(\mu_i)_{i \in I}$ of measures on \mathcal{A} is called *dimensionally ordered* iff there is a linear ordering $<$ on I such that for all $A \in \mathcal{A}$ and $i, j \in I$: if $i < j$ and $\mu_i(A) < \infty$, then $\mu_j(A) = 0$.

This is a severe restriction. In a dimensionally ordered family $(\mu_i)_{i \in I}$ there is for each $A \in \mathcal{A}$ at most one $i \in I$ with $0 < \mu_i(A) < \infty$; if it generates the c.p.s. $\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$, there is for each $B \in \mathcal{B}$ exactly one $i \in I$ with $0 < \mu_i(B) < \infty$, and we then have $\mu_j(B) = 0$ for each $j > i$ and $\mu_j(B) = \infty$ for each $j < i$. Császár (1955), pp. 357f, has now proved the following

THEOREM 1. For each c.p.s. $\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$ there is a dimensionally ordered family of measures on \mathcal{A} generating it, if and only if it can be extended to an additive c.p.s., i.e. if there is an additive c.p.s. $\langle \Omega, \mathcal{A}, \mathcal{B}', P' \rangle$ such that $\mathcal{B} \subseteq \mathcal{B}'$ and $P \subseteq P'$.

This is a most satisfying representation result for c.p.s.s in general, and it explains the special role of additive c.p.s.s. Of course, this result also applies to Popper spaces. But since Popper spaces have stronger properties, stronger representation results should be, and indeed are, obtainable for them. To these we shall work up now.

The first step is to strengthen the concept of dimensional ordering. ($\alpha, \beta, \gamma, \dots, \zeta$ will in the sequel be used as variables for ordinal numbers.)

DEFINITION 4. The family $(\mu_\alpha)_{\alpha < \zeta}$ of measures on \mathcal{A} is said to be *dimensionally well-ordered* (d.w.o.) iff the following holds:

- (a) if $\mu_\beta(A) < \infty$, then for all $\alpha < \beta$ $\mu_\alpha(A) = 0$,
- (b) if $\mu_\beta(A) = \infty$, then there is a $\alpha < \beta$ with $0 < \mu_\alpha(A) < \infty$.

Trivially, each d.w.o. family is dimensionally ordered. In fact, d.w.o. families have a rather special nature, which is not very perspicuous from Definition 4. The following reformulation will make this clear. Preparatory to this is the

LEMMA 1. Let be $(\mu_\alpha)_{\alpha < \zeta}$ d.w.o. Then the finite part of each μ_α is bounded, i.e. for each $\alpha < \zeta$ there is a real x such that for each $A \in \mathcal{A}$ either $\mu_\alpha(A) < x$ or $\mu_\alpha(A) = \infty$.

Proof. Suppose that the finite part of μ_β is not bounded. That is, there is a sequence $A_1 \subseteq A_2 \subseteq \dots$ in \mathcal{A} such that $\mu_\beta(A_n) < \infty$ for all $n \in N$ (N being the set of natural numbers) and $\mu_\beta(A) = \infty$ for $A = \bigcup_{n \in N} A_n$. Clause (a) of Definition 4 now implies that for each $\alpha < \beta$ $\mu_\alpha(A_n) = 0$ for all $n \in N$ and hence, by σ -additivity, $\mu_\alpha(A) = 0$. But this contradicts clause (b) of Definition 4. Q.E.D.

Now, the reformulation is this:

DEFINITION 5. $(\pi_\alpha)_{\alpha < \zeta}$ is called a *dimensionally well-ordered (d.w.o.) family of probability measures on \mathcal{A}* , iff each π_β is a probability measure on \mathcal{A} for which there is a $C_\beta \in \mathcal{A}$ such that $\pi_\beta(C_\beta) = 1$ and $\pi_\alpha(C_\beta) = 0$ for all $\alpha < \beta$.

That this amounts to the same as Definition 4 is shown by

THEOREM 2. Each d.w.o. family $(\pi_\alpha)_{\alpha < \zeta}$ of probability measures on \mathcal{A} , together with a sequence $(x_\alpha)_{\alpha < \zeta}$ of positive reals, defines a d.w.o. family $(\mu_\alpha)_{\alpha < \zeta}$ of measures on \mathcal{A} in the following way:

(*) for all $\beta < \zeta$ and $A \in \mathcal{A}$, $\mu_\beta(A) = x_\beta \cdot \pi_\beta(A)$, if $\pi_\alpha(A) = 0$ for all $\alpha < \beta$, and $\mu_\beta(A) = \infty$ otherwise.

Conversely, each d.w.o. family of measures on \mathcal{A} is, according to (*), defined by some d.w.o. family of probability measures on \mathcal{A} and a suitable sequence of positive reals.

Proof. Let $(\mu_\alpha)_{\alpha < \zeta}$ be defined by (*). It is clear that each μ_β is a measure on \mathcal{A} . If $\mu_\beta(A) < \infty$, then, according to (*), for all $\alpha < \beta$ $\pi_\alpha(A) = 0$, i.e. $\mu_\alpha(A) = 0$. If $\mu_\beta(A) = \infty$, then, according to (*), there is a $\alpha < \beta$ with $\pi_\alpha(A) > 0$, i.e. with $0 < \mu_\alpha(A) < \infty$. Thus, $(\mu_\alpha)_{\alpha < \zeta}$ satisfies clauses (a) and (b) of Definition 4.

For the converse, let $(\mu_\alpha)_{\alpha < \zeta}$ be a d.w.o. family of measures on \mathcal{A} . Defining $x_\alpha = \sup \{\mu_\alpha(A) \mid \mu_\alpha(A) < \infty\}$, we know from Lemma 1 that $x_\alpha < \infty$ for each $\alpha < \zeta$. Hence, there is a sequence $A_1 \subseteq A_2 \subseteq \dots$ in $\{A \mid \mu_\alpha(A) < \infty\}$ with $\lim_{n \rightarrow \infty} \mu_\alpha(A_n) = x_\alpha$. Now define $C_\alpha = \bigcup_{n \in N} A_n$. Then, $\mu_\alpha(C_\alpha) = x_\alpha$, and for each A with $\mu_\alpha(A) < \infty$ we have

$\mu_\alpha(A \setminus C_\alpha) = 0$, i.e. $\mu_\alpha(A) = \mu_\alpha(A \cap C_\alpha)$; otherwise, x_α wouldn't be the specified supremum. So, we may define $\pi_\alpha(A) = \mu_\alpha(A \cap C_\alpha) / x_\alpha$ for all $A \in \mathcal{A}$. It is then obvious that each π_α is a probability measure on \mathcal{A} with $\pi_\alpha(C_\alpha) = 1$ and $\pi_\beta(C_\alpha) = 0$ for $\beta < \alpha$ and that (*) holds for (μ_α) , (π_α) , and (x_α) . Q.E.D.

Thus, we may, and are well advised to, go on working with the simpler d.w.o. families of probability measures. Note that two different d.w.o. families of measures definable through (*) by the same d.w.o. family of probability measures and therefore differing only in the appertaining sequence of positive reals generate the same c.p.s. according to Definition 2. Hence, we may transfer the concept of generation to d.w.o. families of probability measures:

DEFINITION 6. The d.w.o. family $(\pi_\alpha)_{\alpha < \zeta}$ of probability measures on \mathcal{A} generates the c.p.s. $\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$, if and only if $\mathcal{B} = \{B \in \mathcal{A} \mid \text{there is a } \alpha < \zeta \text{ with } \pi_\alpha(B) > 0\}$ and for all $A \in \mathcal{A}$ and $B \in \mathcal{B}$ $P(A \mid B) = \pi_\beta(A \mid B)$, where $\beta = \min \{\alpha \mid \pi_\alpha(B) > 0\}$.

Thereby, we have reached our first aim:

THEOREM 3. Each d.w.o. family of probability measures generates a Popper space.

Proof. Define \mathcal{B} and P from $(\pi_\alpha)_{\alpha < \zeta}$ as in Definition 6. Clause (a) of Definition 1 is trivially satisfied. Clause (b) is also satisfied: Let $A, B, C \in \mathcal{A}$ with $C, B \cap C \in \mathcal{B}$, and let $\gamma = \min \{\alpha \mid \pi_\alpha(C) > 0\}$ and $\beta = \min \{\alpha \mid \pi_\alpha(B \cap C) > 0\}$. If $\gamma < \beta$, we have $P(B \mid C) = P(A \cap B \mid C) = 0$; and if $\gamma = \beta$, then the equation in clause (b) turns into the equation $\pi_\gamma(A \cap B \mid C) = \pi_\gamma(A \mid B \cap C) \cdot \pi_\gamma(B \mid C)$, which is a truth of standard probability theory. Finally, clause (d) is met, too: Suppose that $P(A \mid B) > 0$ for some $B \in \mathcal{B}$, i.e. $\pi_\beta(A \mid B) > 0$ for $\beta = \min \{\alpha \mid \pi_\alpha(B) > 0\}$. This implies that $\pi_\beta(A \cap B) > 0$ and, thus, $\pi_\beta(A) > 0$, i.e. $A \in \mathcal{B}$. Q.E.D.

This representation of Popper measures is indeed unique:

THEOREM 4. Two d.w.o. families of probability measures generating the same Popper space are identical.

Proof. Let $(\pi_\alpha)_{\alpha < \zeta}$ and $(\pi'_\alpha)_{\alpha < \zeta'}$ generate the same Popper space $\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$. We'll show by transfinite induction that $\pi_\alpha = \pi'_\alpha$ for all $\alpha < \zeta$ and that $\zeta = \zeta'$. Thus, suppose that $\beta < \zeta, \zeta'$ and $\pi_\alpha = \pi'_\alpha$ for all $\alpha < \beta$. We then have to prove that $\pi_\beta = \pi'_\beta$:

According to Definition 5 there is a C and a C' such that $\pi_\beta(C) = \pi'_\beta(C') = 1$ and $\pi_\alpha(C) = \pi'_\alpha(C) = \pi'_\alpha(C') = \pi_\alpha(C') = 0$ for all $\alpha < \beta$. Now, if $\pi'_\beta(C) < 1$, then $P(C \mid C \cup C') = \pi'_\beta(C \mid C \cup$

$C') < 1$ and, contrary to that, $P(C|C \cup C') = \pi_\beta(C|C \cup C') = 1$. Hence, $\pi'_\beta(C) = 1$. Thus, for each $A \in \mathcal{A}$, $\pi'_\beta(A) = \pi'_\beta(A \cap C) + \pi'_\beta(A \setminus C) = \pi'_\beta(A|C) = P(A|C)$, and, similarly, $\pi_\beta(A) = P(A|C) = \pi'_\beta(A)$. Finally, if $\zeta < \zeta'$, there would be a C such that $\pi'_\zeta(C) = 1$ and $\pi'_\alpha(C) = \pi_\alpha(C) = 0$ for all $\alpha < \zeta$ and, hence, both $C \in \mathcal{B}$ and not $C \in \mathcal{B}$. $\zeta' < \zeta$ is impossible in the same way. So, $\zeta = \zeta'$. Q.E.D.

Our second aim is, of course, to reverse Theorem 3. This is a bit harder to get. For the following auxiliary definitions and lemmata we shall assume some fixed Popper space $\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$.

LEMMA 2. For all $A, C \in \mathcal{A}$ and $B \in \mathcal{B}$ we have: if $P(A|B) > 0$, then $P(A \cup C|B \cup C) > 0$.

Proof. If $P(C|B \cup C) > 0$, then also $P(A \cup C|B \cup C) > 0$. If $P(C|B \cup C) = 0$, then $P(B|B \cup C) = 1$, and, thus, $P(A \cup C|B \cup C) \geq P((A \cup C) \cap B|B \cup C) = P(A \cup C|B) \cdot P(B|B \cup C) > 0$ according to clause (b) of Definition 1. Van Fraassen (1976) has proved this lemma on p. 424. Q.E.D.

Now we define an order relation for the events in \mathcal{A} : For each $A, B \in \mathcal{A}$, $A \preceq B$ is to hold, if and only if either $A \cup B \in \mathcal{B}$ and $P(B|A \cup B) > 0$ or $A, B \in \mathcal{A} \setminus \mathcal{B}$. Moreover, $A \sim B$ iff $A \preceq B$ and $B \preceq A$, and $A \prec B$ iff $A \preceq B$ and not $B \preceq A$. (The idea to define these relations is contained in van Fraassen (1976), pp. 426f, and also in Császár (1955), p. 342.) \preceq proves to be a weak order in \mathcal{A} :

LEMMA 3. \preceq is connected, i.e. for all $A, B \in \mathcal{A}$, $A \preceq B$ or $B \preceq A$. (The proof is trivial.)

LEMMA 4. \preceq is transitive, i.e. for all $A, B, C \in \mathcal{A}$: if $A \preceq B$ and $B \preceq C$, then $A \preceq C$.

Proof. Again, the lemma is trivial, if $A, B \in \mathcal{A} \setminus \mathcal{B}$ or $B, C \in \mathcal{A} \setminus \mathcal{B}$. So suppose that $P(B|A \cup B) > 0$ and $P(C|B \cup C) > 0$. The first implies, by Lemma 2, that $P(B \cup C|A \cup B \cup C) > 0$. Thus we have $P(C|A \cup C) \geq P(C|A \cup B \cup C) = P(C|B \cup C) \cdot P(B \cup C|A \cup B \cup C) > 0$ according to clause (b) of Definition 1. Q.E.D.

The immediate consequence is that \sim is an equivalence relation in \mathcal{A} . The equivalence classes generated by \sim are even well-ordered. This results from

LEMMA 5. There is no infinite, strictly increasing sequence in \mathcal{A} , i.e. no sequence $A_1 \prec A_2 \prec \dots$.

Proof. Suppose that for all $m < n$ $A_m \prec A_n$, i.e. $P(A_m |$

$A_m \cup A_n) = 0$, and let $A = \bigcup_{n \in N} A_n$. Then $P(A_m | A) = 0$ holds for all $m \in N$, and this implies, by σ -additivity, that $P(A | A) = 0$, i.e. a contradiction. Q.E.D.

The next two lemmata detail the structure of the equivalence classes generated by \sim . For that purpose, define for each $A \in \mathcal{A}$ $\mathcal{R}_A = \{B \in \mathcal{A} | B \preceq A\}$ and $\mathcal{S}_A = \{B \in \mathcal{A} | B \prec A\}$; hence, $\{B \in \mathcal{A} | A \sim B\} = \mathcal{R}_A \setminus \mathcal{S}_A$. Remember that a σ -ideal in \mathcal{A} is a non-empty subset \mathcal{R} of \mathcal{A} such that $B \in \mathcal{R}$ for all $B \subseteq A$, if $A \in \mathcal{R}$, and $\bigcup_{n \in N} A_n \in \mathcal{R}$, if $A_n \in \mathcal{R}$ for all $n \in N$.

LEMMA 6. For each $A \in \mathcal{A}$, \mathcal{R}_A and \mathcal{S}_A are σ -ideals in \mathcal{A} .

Proof. Suppose first that $B \in \mathcal{R}_A$ and $C \subseteq B$. This implies that $B \preceq A$ and $C \preceq B$ and, thus, by Lemma 4, $C \preceq A$, i.e. $C \in \mathcal{R}_A$. Next suppose that $B_n \in \mathcal{R}_A$ for all $n \in N$ and $B = \bigcup_{n \in N} B_n$. This means that for all $n \in N$ $P(A | A \cup B_n) > 0$. Since $P(B|B) = 1$, there must be a $m \in N$ with $P(B_m | B) > 0$. Lemma 2 yields $P(A \cup B_m | A \cup B) > 0$. Therefore, we also have $P(A | A \cup B) = P(A \cup B_m | A \cup B) \cdot P(A | A \cup B_m) > 0$, i.e. $B \preceq A$. This shows that \mathcal{R}_A is a σ -ideal in \mathcal{A} . Since the equivalence classes generated by \sim are well-ordered, there must be some B such that $\mathcal{S}_A = \mathcal{R}_B$; thus, \mathcal{S}_A is also a σ -ideal. Q.E.D.

In fact, for $A \in \mathcal{B}$ \mathcal{R}_A is of a rather special character:

LEMMA 7. For each $A \in \mathcal{B}$ there is a $C \in \mathcal{R}_A$ such that $\mathcal{R}_A = \{B \in \mathcal{A} | \text{there is a } B' \in \mathcal{S}_A \text{ with } B \subseteq C \cup B'\}$.

Proof. First, we have $\inf\{P(A|A') | A' \in \mathcal{R}_A\} > 0$. For suppose there would be a sequence $A_1 \subseteq A_2 \subseteq \dots$ in \mathcal{R}_A with $\lim_{n \rightarrow \infty} P(A | A \cup A_n) = 0$. For $A' = \bigcup_{n \in N} A_n$ we would then have $P(A | A \cup A') = 0$, i.e. $A \prec A'$, and also $A' \in \mathcal{R}_A$, since \mathcal{R}_A is a σ -ideal; and that's a contradiction. Again, since \mathcal{R}_A is a σ -ideal, there must be a $C \in \mathcal{R}_A$ with $P(A|C) = \inf\{P(A|A') | A' \in \mathcal{R}_A\}$. Trivially, we then have $B \in \mathcal{R}_A$ for $B \subseteq C \cup B'$ and $B' \in \mathcal{S}_A$. For the converse, suppose $B \in \mathcal{R}_A$. Then $P(A|B \cup C) = P(A|C) > 0$. Clause (b) of Definition 1 says that $P(A|B \cup C) = P(A|C) \cdot P(C|B \cup C)$. Hence, $P(C|B \cup C) = 1$, i.e. $P(B \setminus C | B \cup C) = 0$, i.e. $B \setminus C \prec C$. Thus, there is a $B' \in \mathcal{S}_A$ with $B \subseteq C \cup B'$. Q.E.D.

Summarizing all these lemmata, we may define a well-ordered sequence $(\mathcal{R}_\alpha)_{\alpha < \zeta+1}$ of σ -ideals in \mathcal{A} in the following way:

$$\mathcal{R}_0 = \mathcal{A}, \quad \mathcal{R}_{\alpha+1} = \{B \in \mathcal{A} | B \prec A \text{ for some } A \in$$

$\mathcal{H}_\alpha\}$, and, in case α is a limit number, $\mathcal{H}_\alpha = \bigcap_{\beta < \alpha} \mathcal{H}_\beta$. For this sequence the following holds true: if $\beta < \alpha$, then $\mathcal{H}_\alpha \subset \mathcal{H}_\beta$; each equivalence class generated by \sim is of the form $\mathcal{H}_\alpha \setminus \mathcal{H}_{\alpha+1}$; $\mathcal{H}_\zeta = \mathcal{A} \setminus \mathcal{B}$; and, finally, for each $\alpha < \zeta$ there exists a $C_\alpha \in \mathcal{H}_\alpha$ such that $\mathcal{H}_\alpha = \{B \in \mathcal{A} \mid \text{there is a } B' \in \mathcal{H}_{\alpha+1} \text{ with } B \subseteq C_\alpha \cup B'\}$. Now our second aim is in reach:

THEOREM 5. Each Popper space $\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$ is generated by a d.w.o. family of probability measures.

Proof. Let the C_α be as just stated, and define for all $A \in \mathcal{A}$ $\pi_\alpha(A) = P(A \mid C_\alpha)$. Thus, each π_α is a probability measure on \mathcal{A} with $\pi_\alpha(C_\alpha) = 1$. If $\beta < \alpha$, then $\pi_\beta(C_\alpha) = P(C_\alpha \mid C_\beta) = 0$, since $C_\alpha < C_\beta$. Hence, the family $(\pi_\alpha)_{\alpha < \zeta}$ is d.w.o. It also generates $\langle \Omega, \mathcal{A}, \mathcal{B}, P \rangle$, as may be seen thus: First, it is obvious that for each $B \in \mathcal{A}$ there is a $\alpha < \zeta$ with $\pi_\alpha(B) > 0$, if and only if $B \in \mathcal{H}_\alpha$ for some $\alpha < \zeta$, i.e. if $B \in \mathcal{B}$. Secondly, let $B \in \mathcal{B}$ and $\beta = \min\{\alpha \mid \pi_\alpha(B) > 0\}$. According to our definition of the π_α , this means that $\beta = \min\{\alpha \mid P(B \mid C_\alpha) > 0\}$. This implies that $C_\beta \preceq B$ and $B < C_\alpha$ for all $\alpha < \beta$, i.e. $B \sim C_\beta$, and thus $B \in \mathcal{H}_\beta \setminus \mathcal{H}_{\beta+1}$. This in turn implies that $B \setminus C_\beta \in \mathcal{H}_{\beta+1}$, i.e. that $P(B \setminus C_\beta \mid B) = 0$ or, equivalently, $P(B \cap C_\beta \mid B) = 1$. And from this we finally get for $A \in \mathcal{A}$ that $\pi_\beta(A \mid B) = \pi_\beta(A \cap B) / \pi_\beta(B) = P(A \cap B \mid C_\beta) / P(B \mid C_\beta) = P(A \mid B \cap C_\beta) = P(A \mid B)$. Q.E.D.

A simple corollary to this is

THEOREM 6. A c.p.s. is full iff it is generated by a d.w.o. family $(\pi_\alpha)_{\alpha < \zeta}$ of probability measures such that \emptyset is the only event A for which $\pi_\alpha(A) = 0$ for all $\alpha < \zeta$.

This concludes our representation theorems for Popper measures. Retrospectively, it is quite clear why our σ -additive Popper measures lend themselves to nicer representations than the finitely additive Popper measures investigated by van Fraassen (1976). One crucial use of σ -additivity was in proving in Lemma 5 that well-orderings inhere in Popper spaces. The other crucial use was in Lemma 6 which then led to Lemma 7; and Lemma 7 proves a property corresponding to van Fraassen's dimensionality of finitely additive Popper spaces which he had to presuppose for their representation by his minimal families. These are just the two points mentioned right before Definition 2.

As a sort of appendix, let me sketch the connection between this representation of Popper measures and conditional logic. That such a connection exists is to be expected – since conditionalization is essential both to con-

ditional logic and probability theory – and in fact well known (cf., in particular, Stalnaker (1970), Section 2, and Harper (1976), Section II.2). With our representation, this connection may be made more perspicuous:

The concept fundamental to conditional logic is this:

DEFINITION 7. g is a *selection function* for Ω (where Ω may be taken as the set of possible worlds) iff g is a function from the power set of Ω (i.e. the set of propositions) into itself such that for all $A, B \subseteq \Omega$:

- (a) $g(A) \subseteq A$,
- (b) if $g(A) \cap B \neq \emptyset$, then $g(A \cap B) = g(A) \cap B$,
- (c) if $A \subseteq B$ and $g(A) \neq \emptyset$, then $g(B) \neq \emptyset$.

There is a realistic interpretation to g : B is true conditional on A (whatever this is to precisely mean) iff $g(A) \subseteq B$. Or a more easily graspable epistemic interpretation: B is believed conditional on A (by some person at a certain time) iff $g(A) \subseteq B$.

Several concepts of a selection function are in use, and it is well known that nearly every semantics for conditional logic is based on some such concept. The concept fixed in Definition 7 is the right one for our context, since there is a representation for it which, though being much simpler, resembles that of a Popper measure:

DEFINITION 8. $(C_\alpha)_{\alpha < \zeta+1}$ is a *selection sequence* for Ω , if and only if $(C_\alpha)_{\alpha < \zeta+1}$ is a strictly decreasing sequence of subsets of Ω (with respect to \subseteq) and if $C_0 = \Omega$.

THEOREM 7. Each selection sequence $(C_\alpha)_{\alpha < \zeta+1}$ defines a selection function g in the following way: $g(A) = \emptyset$, if $A \subseteq C_\zeta$; otherwise, $g(A) = A \cap C_\beta \setminus C_{\beta+1}$, where $\beta = \min\{\alpha \mid A \cap C_\alpha \setminus C_{\alpha+1} \neq \emptyset\}$. Conversely, each selection function is in this way defined by exactly one selection sequence.

Sketch of proof. It is easily verified that the function g defined in this way is a selection function. Conversely, define for a given selection function g by transfinite recursion: $C_\beta = \Omega \setminus \bigcup_{\alpha < \beta} g(C_\alpha)$, and $\zeta = \min\{\alpha \mid g(C_\alpha) = \emptyset\}$. Then it is again easily verified that $(C_\alpha)_{\alpha < \zeta+1}$ is a selection sequence and the only one defining the given g in the specified way. Q.E.D.

In fact, this theorem is well known, since it is the core of the equivalence of a selection functions semantics and a similarity spheres semantics for conditionals (cf. Lewis 1973, pp. 58f).

Here, a selection sequence directly corresponds to the sequence of σ -ideals defined just before Theorem 5. Thus,

the rather straightforwardly demonstrable fact that the selection functions familiar from conditional logic, if adapted to the algebraic framework of σ -fields, capture the 0-1-structure of Popper measures is made immediately evident by these representations.

Bibliography

- Carnap, R.: 1971, 1980, 'A basic system of inductive logic', Part I in R. Carnap and R. C. Jeffrey (eds.), *Studies in Inductive Logic and Probability*, Vol. I, University of California Press, Berkeley, 1971; Part II in R. C. Jeffrey (ed.), *ditto*, Vol. II, 1980.
- Császár, A.: 1955, 'Sur la Structure des Espaces de Probabilité Conditionnelle', *Acta Mathematica Academiae Scientiarum Hungaricae* 6, 337–361.
- van Fraassen, B. C.: 1976, 'Representation of conditional probabilities', *Journal of Philosophical Logic* 5, 417–430.
- Harper, W. L.: 1976, 'Rational belief change, Popper functions and counterfactuals', in W. L. Harper and C. A. Hooker (eds.), *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, Vol. I, Reidel, Dordrecht, pp. 73–115.
- Lewis, D.: 1973, *Counterfactuals*, Blackwell, Oxford.
- Loeb, P. A.: 1979, 'An introduction to nonstandard analysis and hyperfinite probability theory', in A. T. Bharucha-Reid (ed.), *Probabilistic Analysis and Related Topics*, Vol. 2, Academic Press, New York, pp. 105–142.
- Loève, M.: 1960, *Probability Theory*, Van Nostrand, Princeton.
- Popper, K. R.: 1959, *The Logic of Scientific Discovery*, Hutchinson, London.
- Rényi, A.: 1955, 'On a new axiomatic theory of probability', *Acta Mathematica Academiae Scientiarum Hungaricae* 6, 285–335.
- Stalnaker, R. C.: 1970, 'Probability and conditionality', *Philosophy of Science* 37, 64–80.

*Seminar für Philosophie, Logik und Wissenschaftstheorie,
Universität München,
8000 München 22, B.R.D.*