Lexicographic probability, conditional probability, and nonstandard probability

Joseph Y. Halpern
Dept. Computer Science
Cornell University
Ithaca, NY 14853
halpern@cs.cornell.edu
http://www.cs.cornell.edu/home/halpern

May 3, 2001

Abstract

The relationship between *Popper spaces* (conditional probability spaces that satisfy some regularity conditions), lexicographic probability systems (LPS's) [Blume, Brandenburger, and Dekel 1991a; Blume, Brandenburger, and Dekel 1991b], and nonstandard probability spaces (NPS's) is considered. If countable additivity is assumed, Popper spaces and a subclass of LPS's are equivalent; without the assumption of countable additivity, the equivalence no longer holds. If the state space is finite, LPS's are equivalent to NPS's. However, if the state space is infinite, NPS's are shown to be more general than LPS's.

1 Introduction

Probability is certainly the most commonly-used approach for representing uncertainty and conditioning the standard way of updating probabilities in the light of new information. Unfortunately, there is a well-known problem with conditioning: Conditioning on events of measure 0 is not defined. That makes it unclear how to proceed if an agent learns something to which she initially assigned probability 0. Although conditioning on events of measure 0 may seem to be of little practical interest, it turns out to play a critical role in game theory (see, for example, [Blume, Brandenburger, and Dekel 1991a; Blume, Brandenburger, and Dekel 1991b; Hammond 1994; Kreps and Wilson 1982; Myerson 1986; Selten 1965]), the analysis of conditional statements (see [Adams 1966; McGee 1994]), and in dealing with nonmonotonicity (see, for example, [Lehmann and Magidor 1992]).

There have been various attempts to deal with the problem of conditioning on events of measure 0. Perhaps the best known, which goes back to Popper [1968] and de Finetti [1936], is to take as primitive not probability, but conditional probability. If μ is a conditional probability measure, then $\mu(V|U)$ may still be undefined for some pairs V and U, but it is also possible that $\mu(V|U)$ is defined even if $\mu(U) = 0$. Another approach, which goes back to at least Robinson [1973] and has been explored in the economics literature [Hammond 1994], the AI literature [Lehmann and Magidor 1992; Wilson 1995], and the philosophy literature (see [McGee 1994] and the references therein) is to consider *nonstandard probability spaces* (NPS's), where there are infinitesimals that can be used to model events that, intuitively, have infinitesimally small probability yet may still be learned or observed.

There is another approach to this problem, which uses sequences of probability measures to represent uncertainty. The most recent exemplar of this approach, which I focus on here, are the *lexicographic probability systems* of Blume, Brandenburger, and Dekel [1991a, 1991b] (BBD from now on). However, the idea of using a system of measures to represent uncertainty actually was explored as far back as the 1950s by Rényi [1956]. A *lexicographic probability system* is a sequence $\langle \mu_0, \mu_1, \ldots \rangle$ of probability measures. Intuitively, the first measure in the sequence, μ_0, is the most important one, followed by μ_1, μ_2, and so on. Roughly speaking, the probability assigned to an event U by a sequence such as $\langle \mu_0, \mu_1 \rangle$ can be taken to be $\mu_0(U) + \epsilon\mu_1(U)$, where ϵ is an infinitesimal. Thus, even if the probability of U according to μ_0 is 0, U still has a positive (although infinitesimal) probability if $\mu_1(U) > 0$.

How are all these approaches related? This question, which is the focus of the paper, has been considered before. For example, Hammond [1994] shows that conditional probability spaces are equivalent to a subclass of LPS's called conditional LPS's if the state space is finite and it is possible to condition on any nonempty set. As shown by Spohn [1986], Hammond's result can be extended to arbitrary countably additive *Popper spaces*, where a Popper space is a conditional probability space that satisfies certain regularity conditions. The extension is nontrivial and, indeed, does not work without the assumption of countable additivity. Rényi [1956] and van Fraassen [1976] provide other representations of conditional probability spaces as sequences of measures, although not LPS's. Their results apply even if the underlying state space is infinite, but countable additivity does not play a role in their representations. (See Section 3 for further discussion of this issue.)

I show that if the state space is finite, then LPS's are equivalent to NPS's, using a strong notion of equivalence. This equivalence breaks down if the state space is infinite; in this case, NPS's are strictly more general than LPS's (whether or not countable additivity is assumed).

Finally, I consider the relationship between Popper spaces and NPS's, and show that NPS's are more general. (The theorem I prove is a generalization of one proved by McGee [1994], but my interpretation of it is quite different; see Section 5.)

The remainder of the paper is organized as follows. In the next section, I review all the relevant definitions for the three representations of uncertainty considered here. Section 3 considers the relationship between Popper spaces and LPS's. Section 4 considers the relationship between LPS's and NPS's. Finally, Section 5 considers the relationship between Popper spaces and NPS's. I conclude with some discussion of these results in Section 6.

2 Conditional, lexicographic, and nonstandard probability spaces

In this section I briefly review the three approaches to representing likelihood discussed in the introduction.

2.1 Popper spaces

A *conditional probability measure* takes *pairs* U, V of subsets as arguments; $\mu(V, U)$ is generally written $\mu(V|U)$ to stress the conditioning aspects. The first argument comes from some algebra $\mathcal{F}$ of subsets of a space W; if W is infinite, $\mathcal{F}$ is often taken to be a σ-algebra. (Recall that an algebra of subsets of

W is a set of subsets containing W and closed under union and complementation. A σ-algebra is an algebra that is closed under union countable.) The question is what constraints, if any, should be placed on the second argument. I start with three minimal requirements, and later add a fourth.

Definition 2.1: A *Popper algebra* over W is a set $\mathcal{F} \times \mathcal{F}'$ of subsets of $W \times W$ such that (a) $\mathcal{F}$ is an algebra over W, (b) $\mathcal{F}'$ is a nonempty subset of $\mathcal{F}$ (not necessarily an algebra over W), and (c) $\mathcal{F}'$ is closed under supersets in $\mathcal{F}$, in that if $V \in \mathcal{F}'$, $V \subseteq V'$, and $V' \in \mathcal{F}$, then $V' \in \mathcal{F}'$. (Popper algebras are named after Karl Popper.) ∎

Definition 2.2: A *conditional probability space (cps) over* $(W, \mathcal{F})$ is a tuple $(W, \mathcal{F}, \mathcal{F}', \mu)$ such that $\mathcal{F} \times \mathcal{F}'$ is a Popper algebra over W and $\mu : \mathcal{F} \times \mathcal{F}' \to [0, 1]$ satisfies the following conditions:

CP1. $\mu(U|U) = 1$ if $U \in \mathcal{F}'$.

CP2. $\mu(V_1 \cup V_2|U) = \mu(V_1|U) + \mu(V_2|U)$ if $V_1 \cap V_2 = \emptyset$, $U \in \mathcal{F}'$, and $V_1, V_2 \in \mathcal{F}$.

CP3. $\mu(V|U) = \mu(V|X) \times \mu(X|U)$ if $V \subseteq X \subseteq U$, $U, X \in \mathcal{F}'$, $V \in \mathcal{F}$.

A *Popper space over* $(W, \mathcal{F})$ is a conditional probability space $(W, \mathcal{F}, \mathcal{F}', \mu)$ that satisfies an additional condition: if $U \in \mathcal{F}'$ and $\mu(V|U) \neq 0$ then $V \cap U \in \mathcal{F}'$. If $\mathcal{F}$ is a σ-algebra and μ is countably additive (that is, if $\mu(\cup V_i|U) = \sum_{i=1}^{\infty} \mu(V_i|U)$ if the V_i's are pairwise disjoint elements of $\mathcal{F}$ and $U \in \mathcal{F}'$), then the Popper space is said to be *countably additive*. Let $Pop(W, \mathcal{F})$ denote the set of Popper spaces over $(W, \mathcal{F})$; if $\mathcal{F}$ is a σ-algebra, let $Pop^c(W, \mathcal{F})$ denote the set of countably additive Popper spaces over $(W, \mathcal{F})$. The probability measure μ in a Popper space is called a *Popper measure*. ∎

The additional regularity condition on $\mathcal{F}'$ required in a Popper space corresponds to the observation that for an unconditional probability measure μ, if $\mu(V|U) \neq 0$ then $\mu(V \cap U) \neq 0$, so conditioning on $V \cap U$ should be defined.

Popper [1968] was the first to consider formally conditional probability as the basic notion. Although his definition of conditional probability space is not quite the same as that used here. CP1–3 are essentially due to Rényi [1955]. De Finetti [1936] also did some early work, apparently independently, taking conditional probabilities as primitive. Indeed, as Rényi [1964] points out, the idea of taking conditional probability as primitive seems to go back as far as Keynes [1921]. Van Fraassen [1976] defined what I have called Popper measures; he called them Popper functions, reserving the name Popper measure for what I am calling a countably additive Popper measure. Hammond [1994] discusses the use of conditional probability spaces in philosophy and game theory, and provides an extensive list of references.

2.2 Lexicographic probability spaces

Definition 2.3: A *lexicographic probability space (LPS) (of length* α*) over* $(W, \mathcal{F})$ is a tuple $(W, \mathcal{F}, \vec{\mu})$ where, as before, W is a set of possible worlds and $\mathcal{F}$ is an algebra over W, and $\vec{\mu}$ is a sequence of probability measures on $(W, \mathcal{F})$ indexed by ordinals $< \alpha$. (Technically, $\vec{\mu}$ is a function from the ordinals less than α to probability measures on $(W, \mathcal{F})$.) I typically write $\vec{\mu}$ as $(\mu_0, \mu_1, \ldots)$ or as $(\mu_\beta : \beta < \alpha)$. If $\mathcal{F}$ is a σ-algebra and each of the probability measures in $\vec{\mu}$ is countably additive, then $\vec{\mu}$ is a *countably additive LPS*. Let $LPS(W, \mathcal{F})$ denote the set of LPS's over $(W, \mathcal{F})$; if $\mathcal{F}$ is a σ-algebra, let $LPS^c(W, \mathcal{F})$ denote the set of countably additive LPS's over $(W, \mathcal{F})$. When $(W, \mathcal{F})$ are understood, I often refer to $\vec{\mu}$ as the LPS. ∎

BBD define a *conditional lexicographic probability space (CLPS)* to be an LPS such that the probability measures in the sequence have disjoint supports; that is, there exist sets $U_i \in \mathcal{F}$ such that $\mu_i(U_i) = 1$ and the sets U_i are pairwise disjoint for $i < \alpha$. Let a *structured LPS (SLPS)* be an LPS such that there exist sets $U_i \in \mathcal{F}$ such that $\mu_i(U_i) = 1$ and $\mu_i(U_j) = 0$ for $j > i$. (Spohn [1986] calls SLPS's *dimensionally well-ordered families of probability measures*; they are also the "probabilified ordinal conditional functions" (OCFs) briefly discussed in [Spohn 1988].) Let *SLPS*$(W, \mathcal{F})$ denote the set of SLPS's over $(W, \mathcal{F})$; if $\mathcal{F}$ is a σ-algebra, let *SLPS*$^c(W, \mathcal{F})$ denote the set of countably additive SLPS's over $(W, \mathcal{F})$.

Clearly every CLPS is an SLPS; moreover, if α is countable, then every countably additive SLPS is a CLPS: Given an SLPS $\vec{\mu}$ with associated sets $U_i, i < \alpha$, define $U_i' = U_i - (\cup_{j>i} U_j)$. (Clearly this is true even without the assumption of countable additivity if α is finite.) The sets U_i' are clearly pairwise disjoint elements of $\mathcal{F}$, and U_i' is a support for μ_i. However, in general, CLPS's are a strict subset of SLPS's, as the following example shows.

Example 2.4: Consider a well-ordering of the interval $[0, 1]$, that is, an isomorphism from $[0, 1]$ to an initial segment of the ordinals. Suppose that the initial segment of the ordinals has length α. Let $([0, 1], \mathcal{F}, \vec{\mu})$ be an LPS of length α where $\mathcal{F}$ consists of the Borel subsets of $[0, 1]$. Let μ_0 be the standard Borel measure on $[0, 1]$, and let μ_β be the measure that gives probability 1 to r_β, the βth real in the well-ordering. This clearly gives an SLPS, since the support of μ_0 is $[0, 1]$ and the support of μ_β for $0 < \beta < \alpha$ is $\{r_\beta\}$. However, this SLPS is not equivalent to any CLPS; there is no support of μ_0 which is disjoint from the supports of μ_β for all β with $0 < \beta < \alpha$. ∎

The difference between CLPS's and SLPS's does not arise in the work of BBD, since they consider only finite sequences of measures. The restriction to finite sequences, in turn, is due to their restriction to finite sets W of possible worlds. Clearly, if W is finite, then all CLPS's over W must be finite, since the support of each of the measures must be disjoint.

We can put an obvious lexicographic order $<_L$ on sequences $(x_0, x_1, \ldots)$ of numbers in $[0, 1]$ of length α: $(x_0, x_1, \ldots) <_L (y_0, y_1, \ldots)$ if there exists $\beta < \alpha$ such that $x_\beta < y_\beta$ and $x_\gamma = y_\gamma$ for all $\gamma < \beta$. That is, we compare two sequences by comparing their components at the first place they differ. (Even if α is infinite, because we are dealing with ordinals, there will be a least ordinal at which the sequences differ if they differ at all.) This lexicographic order will be used to define decision rules.

BBD define conditioning in LPS's as follows. Given $\vec{\mu}$ and $U \in \mathcal{F}$ such that $\mu_i(U) > 0$ for some index i, let $\vec{\mu}|U = (\mu_{k_0}(\cdot|U), \mu_{k_1}(\cdot|U), \ldots)$, where $(k_0, k_1, \ldots)$ is the subsequence of all indices for which the probability of U is positive. Note that $\vec{\mu}|U$ is undefined if $\mu_\beta(U) = 0$ for all $\beta < \alpha$.

2.3 Nonstandard probability spaces

It is well known that there exist *non-Archimedean fields*—fields that include the real numbers as a subfield but also have *infinitesemals*, numbers that are positive but still less than any positive real number. The smallest such non-Archimedean field, commonly denoted $I\!R(\epsilon)$, is the smallest field generated by adding to the reals a single infinitesimal ϵ.[1] The *hyperreals*, nonstandard models of the reals which satisfy all the first-order properties that hold of the real numbers (see [Davis 1977]), are

[1]The construction of $I\!R(\epsilon)$ apparently goes back to Robinson [1973]. It is reviewed by Hammond [1994] and Wilson [1995] (who calls $I\!R(\epsilon)$ the *extended reals*).

also instances of non-Archimedean fields. For most of this paper, I use only the following properties of non-Archimedean fields:

1. If $I\!R^*$ is a non-Archimedean field, then for all $b \in I\!R^*$ such that $-r < b < r$ for some standard real $r > 0$, there is a unique closest real number a such that $|a - b|$ is an infinitesimal. (Formally, a is the inf of the set of real numbers that are at least as large as b.) Let $st(b)$ denote the closest standard real to b; $st(b)$ is sometimes read "the standard part of b".

2. If $st(\epsilon/\epsilon') = 0$, then $a\epsilon < \epsilon'$ for all positive standard real numbers a. (If $a\epsilon$ were greater than ϵ', then ϵ/ϵ' would be greater than $1/a$, contradicting the assumption that $st(\epsilon/\epsilon') = 0$.)

Given a non-Archimedean field $I\!R^*$, a *nonstandard probability space (NPS) over* $(W, \mathcal{F})$ *(with range* $I\!R^*$*)* is a tuple $(W, \mathcal{F}, \mu)$, where W is a set of possible worlds, $\mathcal{F}$ is an algebra of subsets of W, and μ assigns to sets in $\mathcal{F}$ an element of $I\!R^*$ such that $\mu(W) = 1$ and $\mu(U \cup V) = \mu(U) + \mu(V)$ if U and V are disjoint. If W is infinite, we may also require that $\mathcal{F}$ be a σ-algebra and that μ be countably additive. (There are some subtleties involved with countable additivity in nonstandard probability spaces; see Section 4.3.)

3 Relating Popper Spaces to (S)LPS's

In this section, I provide an isomorphism $F_{S\to P}$ from Popper spaces over $(W, \mathcal{F})$ to SLPS's over $(W, \mathcal{F})$, for each fixed W and $\mathcal{F}$. Given an SLPS $(W, \mathcal{F}, \vec{\mu})$ of length α, consider the cps $(W, \mathcal{F}, \mathcal{F}', \mu)$ such that $\mathcal{F}' = \cup_{\beta<\alpha}\{V \in \mathcal{F} : \mu_\beta(V) > 0\}$. For $V \in F'$, let j_V be the smallest index such $\mu_{j_V}(V) > 0$. Define $\mu(U|V) = \mu_{j_V}(U|V)$. I leave it to the reader to check that $(W, \mathcal{F}, \mathcal{F}', \mu)$ is a Popper space.

There are many isomorphisms between two infinite spaces. Why is $F_{S\to P}$ of interest? Suppose that $F_{S\to P}(W, \mathcal{F}, \vec{\mu}) = (W, \mathcal{F}, \mathcal{F}', \mu)$. It is easy to check that the following two important properties hold:

- $\mathcal{F}'$ consists precisely of those events for which conditioning in the LPS is defined; that is, $\mathcal{F}' = \{U : \mu_\beta(U) \neq 0 \text{ for some } \mu_\beta \in \vec{\mu}\}$.
- For $U \in \mathcal{F}'$, $\mu(\cdot|U) = \mu'(\cdot|U)$, where μ' is the first probability measure in the sequence $\vec{\mu}|U$. That is, the Popper measure agrees with the most significant probability measure in the conditional LPS given U. Given that an LPS assigns to an event U a sequence of numbers and a Popper measure assigns to U just a single number, this is clearly the best single number to take.

It seems that these are minimal properties that we would want an isomorphism to satisfy. Moreover, it is easy to see that these two properties completely characterize $F_{S\to P}$.

BBD claim without proof that $F_{S\to P}$ is an isomorphism from CLPS's to conditional probability spaces. They work in finite spaces (so that CLPS's are equivalent to SLPS's) and restrict attention to LPS's where (in the notation used here), W is finite, $\mathcal{F} = 2^W$, and $\mathcal{F}' = 2^W - \emptyset$ (so that conditioning is defined for all nonempty sets). Since $\mathcal{F}' = 2^W - \emptyset$, the cps's they consider are all Popper spaces. Hammond [1994] provides a careful proof of this result, under the restrictions considered by BBD. Hammond's result holds for arbitrary finite Popper spaces, with essentially no change in proof.

Theorem 3.1: [Hammond 1994] *If W is finite, then $F_{S\to P}$ is an isomorphism from* SLPS$(W, \mathcal{F})$ *to* Pop$(W, \mathcal{F})$.

The situation is much different in the infinite case (which is not considered by either BBD or Hammond). It is easy to see that $F_{S\rightarrow P}$ is an injection from from SLPS's to Popper spaces. However, as the following example shows, if we do not require countable additivity, it is not an isomorphism.

Example 3.2: (This example is essentially due to Robert Stalnaker [private communication, 2000].) Let $W = I\!N$, the natural numbers, let $\mathcal{F}$ consist of the finite and cofinite subsets of $I\!N$, and let $\mathcal{F}' = \mathcal{F} - \{\emptyset\}$. Define $\mu^1(V|U) = 1$ if U and V are both cofinite. If U is finite, define $\mu^1(V|U) = |V \cap U|/|U|$. I leave it to the reader to check that $(I\!N, \mathcal{F}, \mathcal{F}', \mu^1)$ is a Popper space. Suppose there were some LPS $(I\!N, \mathcal{F}, \vec{\mu})$ which was mapped by $F_{S\rightarrow P}$ to this Popper space. Then it is easy to check that if μ_i is the first measure in $\vec{\mu}$ such that $\mu_i(U) > 0$ for some finite set U, then $\mu_i(U') > 0$ for all finite sets U. To see this, note that for any finite set U', since $\mu_i(U) > 0$, it follows that $\mu_i(U \cup U') > 0$. Since $U \cup U'$ is finite, it must be the case that μ_i is the first measure in $\vec{\mu}$ such that $\mu_i(U \cup U') > 0$. Thus, by definition, $\mu^1(U'|U \cup U') = \mu_i(U'|U \cup U')$. Since $\mu^1(U'|U \cup U') > 0$, it follows that $\mu_i(U') > 0$. Moreover, essentially the same argument shows that $\mu_i(U)$ must be proportional to $|U|$. But there is no probability measure μ_i that makes the probability of every finite set proportional to its cardinality. ∎

As the following theorem, proved by Spohn [1986], shows, there is no such counterexample if we restrict to countably additive SLPS's and countably additive Popper spaces.

Theorem 3.3: [Spohn 1986] *For all W, the map $F_{S\rightarrow P}$ is an isomorphism from* $\mathrm{SLPS}^c(W, \mathcal{F})$ *to* $\mathrm{Pop}^c(W, \mathcal{F})$.

It is important in Theorem 3.3 that we consider SLPS's and not CLPS's. $F_{S\rightarrow P}$ is in fact not an isomorphism from CLPS's to Popper spaces.

Example 3.4: Consider the Popper space $([0, 1], \mathcal{F}, \mathcal{F}', \mu)$ which is the image under $F_{S\rightarrow P}$ of the SLPS constructed in Example 2.4. It is easy to see that this Popper space cannot be the image under $F_{S\rightarrow P}$ of some CLPS. ∎

It is interesting to contrast these results to those of Rényi [1956] and van Fraassen [1976]. Renyi considers what he calls *dimensionally-ordered* systems. A dimensionally ordered system (over $(W, \mathcal{F})$ has the form $(W, \mathcal{F}, \mathcal{F}', \{\mu_i : i \in I\})$, where $\mathcal{F}$ is a an algebra of subsets of W, $\mathcal{F}'$ is a subset of $\mathcal{F}$ closed under finite unions, I is a totally ordered set (but not necessarily well-founded, so it may not, for example, have a first element) and μ_i is a measure on $(W, \mathcal{F})$ (not necessarily a probability measure) such that

- for each $U \in \mathcal{F}'$, there is some $i \in I$ such that $0 < \mu_i(U) < \infty$ (note that the measure of a set may, in general, be ∞),
- if $\mu_i(U) < \infty$ and $j < i$, then $\mu_j(U) = 0$.

Note that it follows from these conditions that for each $U \in \mathcal{F}'$, there is exactly one $i \in I$ such that $0 < \mu_i(U) < \infty$.

There is an obvious analogue of the map $F_{S\rightarrow P}$ mapping dimensionally ordered system to cps's. Namely, let $F_{D\rightarrow C}$ map the dimensionally ordered system $(W, \mathcal{F}, \mathcal{F}', \{\mu_i : i \in I\})$ to the cps

$(W, \mathcal{F}, \mathcal{F}', \mu)$, where $\mu(V|U) = \mu_i(V|U)$, where i is the unique element of I such that $0 < \mu_i(U) < \infty$. Rényi shows that $F_{D \to C}$ is an isomorphism from dimensionally-ordered systems to cps's where the set $\mathcal{F}'$ is closed under finite unions. (Csaszar [1955] extends this result to cases where the set $\mathcal{F}'$ is not necessarily closed under finite union.) Rényi assumes that all measures involved are countably additive and that $\mathcal{F}$ is a σ-algebra, but these are inessential assumptions. That is, his proof goes through without change if $\mathcal{F}$ is an algebra and the measures are additive; all that happens is that the resulting conditional probability measure is additive rather than σ-additive.

It is critical in Rényi's framework that the μ_i's are arbitrary measures, and not just probability measures. His result does not hold if the μ_i's are required to be probability measures. If we consider only finitely additive measures, the Popper space constructed in Example 3.2 already shows why. It corresponds to a dimensionally ordered space (μ_1, μ_2) where $\mu_1(U) = |U|$ (i.e., the measure of a set is its cardinality) and $\mu_2(U)$ is 1 if U is cofinite and 0 if U is finite. It cannot be captured by a dimensionally ordered space where all the elements are probability measures, for the same reason that it is not the image of an SLPS under $F_{S \to P}$. (Rényi [1956] actually provides a general characterization of when the μ_i's can be taken to be (countably additive) probability measures.)

Van Fraassen [1976] proved a result whose assumptions are somewhat closer to Theorem 3.3. Van Fraassen considers what he calls *ordinal families of probability measures*. An ordinal family over $(W, \mathcal{F})$ is a sequence of the form $\{(W_\beta, \mathcal{F}_\beta, \mu_\beta) : \beta < \alpha\}$ such that

- $W_\beta \subseteq W$;
- $\mathcal{F}_\beta$ is an algebra over W_β;
- μ_β is a probability measure with domain $\mathcal{F}_\beta$;
- $\cup_{\beta<\alpha} \mathcal{F}_\beta = \mathcal{F}$;
- if $U \in \mathcal{F}$ and $V \in \mathcal{F}_\beta$, then $U \cap V \in \mathcal{F}_\beta$;
- if $U \in \mathcal{F}$, $U \cap V \in \mathcal{F}_\beta$, and $\mu_\beta(U \cap V) > 0$, then there exists γ such that $U \in \mathcal{F}_\gamma$ and $\mu_\gamma(U) > 0$.

Given an ordinal family $\{(W_\beta, \mathcal{F}_\beta, \mu_\beta) : \beta < \alpha\}$ over $(W, \mathcal{F})$, consider the map $F_{O \to C}$ which associates with it the cps $(W, \mathcal{F}, \mathcal{F}', \mu)$, where $\mathcal{F}' = \{U \in \mathcal{F} : \mu_\gamma(U) > 0 \text{ for some } \gamma < \alpha\}$ and $\mu(V|U) = \mu_\beta(V|U)$, where β is the smallest ordinal such that $U \in \mathcal{F}_\beta$ and $\mu_\beta(U) > 0$. Van Fraassen shows that $F_{O \to C}$ is an isomorphism from ordinal families over $(W, \mathcal{F})$ to Popper spaces over $(W, \mathcal{F})$. Again, for van Fraassen, countable additivity does not play a significant role. If $\mathcal{F}$ is a σ-algebra, a *countably additive* ordinal family over $(W, \mathcal{F})$ is defined just as an ordinal family, except that now $\mathcal{F}_\beta$ is required to be a σ-algebra over W_β for all $\beta < \alpha$ and $\mathcal{F}$ is required to be the least σ-algebra containing $\cup_{\beta<\alpha} \mathcal{F}_\beta$ (since $\cup_{\beta<\alpha} \mathcal{F}_\beta$ is not in general a σ-algebra). The same map $F_{O \to C}$ is also an isomorphism from countably additive ordinal families to countably additive Popper spaces.

Spohn's result, Theorem 3.3, can be viewed as a strengthening of van Fraassen's result in the countably additive case, since for Theorem 3.3 all the $\mathcal{F}_\beta$'s are required to be identical. This is a nontrivial requirement. The fact that it cannot be met in the case that W is infinite and the measures are not necessarily finitely additive is an indication of this.

It is worth seeing how van Fraassen's approach handles the finitely additive examples which do not correspond to SLPS's. The Popper space in Example 3.2 corresponds to the ordinal family

$\{(W_n, \mathcal{F}_n, \mu_n) : n \leq \omega\}$ where, for $n < \omega$, $W_n = \{1, \ldots, n\}$, $\mathcal{F}_n$ consists of all subsets of W_n, and μ_n is the uniform measure, while $W_\omega = \mathbb{N}$, $\mathcal{F}_\omega$ consists of the finite and cofinite subsets of $\mathbb{N}$, and $\mu_\omega(U)$ is 1 if U is cofinite and 0 if U is finite. It is easy to check that this ordinal family has the desired properties. The key point to observe here is the leverage obtained by allowing each probability measure to have a different domain.

4 Relating LPS's to NPS's

In this section, I show that LPS's and NPS's are isomorphic in a strong sense. Again, I separate the results for the finite case and the infinite case.

4.1 The finite case

Consider an LPS of the form (μ_1, μ_2, μ_3). Roughly speaking, the corresponding NPS should be $(1 - \epsilon - \epsilon^2)\mu_1 + \epsilon\mu_2 + \epsilon^2\mu_3$, where ϵ is some infinitesimal. That means that μ_2 gets infinitesimal weight relative to μ_1 and μ_3 gets infinitesimal weight relative to μ_2. The question is, which infinitesimal ϵ should be chosen? Intuitively, it shouldn't matter. No matter which infinitesimal is chosen, the resulting NPS should be equivalent to the original LPS. How can we make this intuition precise?

Suppose that we want to use an LPS or an NPS to compute which of two bounded, *real-valued* random variables has higher expected value. (The intended application here is decision making, where the functions can be thought of as the utilities corresponding to two actions; the one with higher expected utility is preferred.) The idea is that two measures of uncertainty (each of which can be an LPS or an NPS) are equivalent if the preference order they place on random variables (according to their expected value) is the same. Note that, given an LPS $\vec{\mu}$, the expected value of a random variable X is $\sum_x x\vec{\mu}(X = x)$, where $\vec{\mu}(X = x)$ is a sequence of probability values and the multiplication and addition are pointwise. Thus, the expected value is a sequence; these sequences can be compared using the lexicographic order $<_L$ defined in Section 2.2. If ν is either an LPS or NPS, then let $E_\nu(X)$ denote the expected value of random variable X according to ν.

Definition 4.1: If each of ν_1 and ν_2 is either an NPS over $(W, \mathcal{F})$ or an LPS over $(W, \mathcal{F})$, then ν_1 is *equivalent to* ν_2, denoted $\nu_1 \approx \nu_2$, if, for all random variables X and Y measurable with respect to $\mathcal{F}$, $E_{\nu_1}(X) \leq E_{\nu_1}(Y)$ iff $E_{\nu_2}(X) \leq E_{\nu_2}(Y)$. (As usual, X is said to be measurable with respect to $\mathcal{F}$ if $\{w : X(w) = x\} \in \mathcal{F}$ for all x in the range of X.) ∎

This notion of equivalence satisfies analogues of the two key properties of the map $F_{S \to P}$ considered at the beginning of Section 3.

Proposition 4.2: *If* $\nu \in \mathrm{NPS}(W, \mathcal{F})$, $\vec{\mu} \in \mathrm{LPS}(W, \mathcal{F})$, *and* $\nu \approx \vec{\mu}$, *then* $\nu(U) > 0$ *iff* $\vec{\mu}(U) > \vec{0}$. *Moreover, if* $\nu(U) > 0$, *then* $\mathrm{st}(\nu(V|U)) = \mu_j(V|U)$, *where* μ_j *is the first probability measure in* $\vec{\mu}$ *such that* $\mu_j(U) > 0$.

The next result justifies restricting to finite LPS's if the state space is finite. Given an algebra $\mathcal{F}$, let $Basic(\mathcal{F})$ consist of the *basic sets* in $\mathcal{F}$, that is, the nonempty sets $\mathcal{F}$ that themselves contain no nonempty subsets in $\mathcal{F}$. Clearly the sets in $Basic(\mathcal{F})$ are disjoint, so that $|Basic(\mathcal{F})| \leq |W|$. If all sets are measurable, then $Basic(\mathcal{F})$ consists of the singleton subsets of W. If W is finite, it is easy to see that all sets in $\mathcal{F}$ are finite unions of the sets in $Basic(\mathcal{F})$.

Proposition 4.3: *If W is finite, then every LPS over $(W, \mathcal{F})$ is equivalent to an LPS of length at most $|Basic(\mathcal{F})|$.*

I can now define the isomorphism that relates NPS's and LPS's. Given $(W, \mathcal{F})$, let $LPS(W, \mathcal{F})/\approx$ be the equivalence classes of $\approx$-equivalent LPS's over $(W, \mathcal{F})$; similarly, let $NPS(W, \mathcal{F})/\approx$ be the equivalence classes of $\approx$-equivalent NPS's over $(W, \mathcal{F})$. Note that in $NPS(W, \mathcal{F})/\approx$, it is possible that different nonstandard probability measures could have different ranges. For this section, without loss of generality, I could also fix the range of all NPS's to be fixed nonstandard model $\mathbb{R}(\epsilon)$ discussed in Section 2.3. However, in the infinite case, it is not possible to restrict to a single nonstandard model, so I do not do so here either, for uniformity.

Now define the mapping $F_{L\to N}$ from $LPS(W, \mathcal{F})/\approx$ to $NPS(W, \mathcal{F})/\approx$ pretty much as suggested at the beginning of this subsection: If $[\vec{\mu}]$ is an equivalence class of LPS's, then choose a representative $\vec{\mu}' \in [\vec{\mu}]$ with finite length. Suppose that $\vec{\mu}' = (\mu_0, \ldots, \mu_k)$. Let $F_{L\to N}([\vec{\mu}]) = [(1 - \epsilon - \cdots - \epsilon^k)\mu_0 + \epsilon\mu_1 + \cdots + \epsilon^k\mu_k]$.

Theorem 4.4: *If W is finite, then $F_{L\to N}$ is an isomorphism from* $\mathrm{LPS}(W, \mathcal{F})/\approx$ *to* $\mathrm{NPS}(W, \mathcal{F})/\approx$ *that preserves equivalence (that is, each NPS in $F_{L\to N}([\vec{\mu}])$ is equivalent to $\vec{\mu}$).*

BBD [1991a] also relate nonstandard probability measures and LPS's under the assumption that the state space is finite. However, the way they relate them is somewhat different in spirit from the notion of equivalence introduced here. They prove representation theorems essentially showing that a preference orders on lotteries can be represented by a standard utility function on lotteries and an LPS iff it can be represented by a standard utility function on lotteries and an NPS. Thus, they show that NPS's and LPS's are equiexpressive in terms of representing preference orders on lotteries.

The difference between BBD's result and Theorem 4.4 is essentially a matter of quantification. BBD's result can be viewed as showing that, given an LPS, for each utility function on lotteries, there is an NPS that generates the same preference order on lotteries for that particular utility function. In principle, the NPS might depend on the utility function. More precisely, for a fixed LPS $\vec{\mu}$, all that follows from their result is that for each utility function u, there is an NPS ν such that $(\vec{\mu}, u)$ and (ν, u) generate the same preference order on lotteries. Theorem 4.4 says that, given $\vec{\mu}$, there is an NPS ν such that $(\vec{\mu}, u)$ and (ν, u) generate the same preference on lotteries for *all* utility functions u.

4.2 The infinite case

An LPS over an infinite state space W may not be equivalent to any finite LPS. However, ideas analogous to those used to prove Proposition 4.3 can be used to provide a bound on the length of the minimal-length LPS's in an equivalence class.

Proposition 4.5: *Every LPS over $(W, \mathcal{F})$ is equivalent to an LPS over $(W, \mathcal{F})$ of length at most $|\mathcal{F}|$.*

Now, just as in the finite case, given an LPS $(\mu_\beta : \beta < \alpha)$ of length α, we want to show that it is equivalent to some NPS ν. Much like the finite case, the idea will be to take $\nu = \sum_{\beta<\alpha} \epsilon_\beta\mu_\beta$, where $st(\epsilon_{\beta'}/\epsilon_\beta) = 0$ if $\beta < \beta' < \alpha$. There are two issues that must be dealt with in order to get this to work. First, we must ensure that there is a non-Archimedean field where there are infinitesimals ϵ_β, $\beta < \alpha$, such that $st(\epsilon_{\beta'}/\epsilon_\beta) = 0$ if $\beta < \beta' < \alpha$. Note, for example, that this cannot be done in $I\!R(\epsilon)$ if $\alpha > \omega$. Another problem is making sense of the infinite sum. Fields are closed under finite sums; in general, infinite sums may not be defined.

I now construct a family of non-Archimedean fields where these problems are solved. Define a *nonstandard model of the integers* to be a model that contains the integers and satisfies every property of the integers expressible in first-order logic. It follows easily from the compactness theorem of first-order logic [Enderton 1972] that, given an ordinal α, there exists a nonstandard model of the integers that includes elements n_β, $\beta < \alpha$, such that $n_0 = 0$ and $n_\beta < n_{\beta'}$ if $\beta < \beta'$.[2]

Given a nonstandard model I^* of the integers, let $I\!R(I^*)$ be the non-Archimedean model defined as follows: $I\!R(I^*)$ consists of all polynomials of the form $\sum_{\beta<\alpha} r_\beta \epsilon^{n_\beta}$ for some ordinal α, where $n_\beta \in I^*$ for $\beta < \alpha$, $n_\beta < n_{\beta'}$ if $\beta < \beta'$ (so that the set $\{n_\beta : \beta < \alpha\}$ is well founded), and r_β is a standard real for all $\beta < \alpha$. We can identify the standard reals r with a polynomial of the form form $r\epsilon^0$. These polynomials can be added and multiplied using the standard rules for addition and multiplication of polynomials. It is easy to check that the result of adding or multiplying two polynomials is another polynomial in $I\!R(I^*)$. In particular, if p_1 and p_2 are two polynomials, N_1 is the set of coefficients of p_1, and N_2 is the set of coefficients of p_2, then the coefficients of $p_1 + p_2$ lie in $N_1 \cup N_2$, while the coefficients of p_1p_2 lie in the set $N_3 = \{n_1 + n_2 : n \in N_1, n_2 \in N_2\}$. Both $N_1 \cup N_2$ and N_3 are easily seen to be well founded if N_1 and N_2 are. Moreover, for each expression $n_1 + n_2 \in N_3$, it follows from the well-foundedness of N_1 and N_2 that there are only finitely many pairs $(n, n') \in N_1 \times N_2$ such that $n + n' = n_1 + n_2$. Finally, each polynomial (other than 0) has an inverse that can be computed using standard "formal" division of polynomials; I leave the details to the reader. An element of $I\!R(I^*)$ is *positive* if its leading coefficient is positive. Define an order $\leq$ on $I\!R(I^*)$ by taking $a \leq b$ if $b - a$ is positive. With these definitions, $I\!R$ is a non-Archimedean field. Moreover, $st(\epsilon^{n_2}/\epsilon^{n_1}) = 0$ if $n_1 < n_2$.

Given $(W, \mathcal{F})$, let α be the minimal ordinal whose cardinality is greater than $|\mathcal{F}|$. Let $I^*_{(W,\mathcal{F})}$ be a nonstandard model of the integers such that there exist elements n_β in $I^*_{(W,\mathcal{F})}$ for all $\beta < \alpha$ such that $n_0 = 0$ and $n_\beta < n_{\beta'}$ if $\beta < \beta' < \alpha$. We can now define a map $F_{L\to N}$ from $LPS(W,\mathcal{F})/\!\approx$ to $NPS(W,\mathcal{F})/\!\approx$ as follows: Given an equivalence class $[\vec{\mu}] \in LPS(W,\mathcal{F})$, by Proposition 4.5, there exists $\vec{\mu} \in [\vec{\mu}]$ such that $\vec{\mu}$ has length $\alpha' < \alpha$. Let $\nu = \sum_{0<\beta<\alpha'} \epsilon^{n_\beta}\mu_\beta$ and define $F_{L\to N}[\vec{\mu}] = [\nu]$. In the full paper, I show that $\nu \approx \vec{\mu}$. The following result is immediate.

Theorem 4.6: *$F_{L\to N}$ is an injection from* LPS$(W,\mathcal{F})/\!\approx$ *to* NPS$(W,\mathcal{F})/\!\approx$ *that preserves equivalence.*

What about the converse? Is it the case that for every NPS there is an equivalent LPS? As the following example shows, the answer is no.

Example 4.7: As in Example 3.2, let $W = I\!N$, the natural numbers, let $\mathcal{F}$ consist of the finite and cofinite subsets of $I\!N$, and let $\mathcal{F}' = \mathcal{F} - \{\emptyset\}$. Let ν^1 be an NPS with range $I\!R(\epsilon)$, where $\nu^1(U) = |U|\epsilon$ if U is finite and $\nu(U) = 1 - |\overline{U}|\epsilon$ if U is cofinite. This is clearly an NPS, and it corresponds to the cps μ^1 of Example 3.2, in the sense that $st(\nu(V|U)) = \mu^1(V|U)$ for all $V \in \mathcal{F}$, $U \in \mathcal{F}'$. Just as in Example 3.2, it can be shown that there is no LPS $\vec{\mu}$ such that $\nu^1 \approx \vec{\mu}$. ∎

[2] The compactness theorem says that, given a collection for formulas, if each finite subset has a model, then so does the whole set. Consider a language with a function + and constant symbols for each integer, together with constants $\mathbf{n}_\beta$, $\beta < \alpha$. Consider the collection of first-order formulas in this language consisting of all the formulas true of the integers, together with the formulas $\mathbf{n}_0 = 0$ and $\mathbf{n}_\beta < \mathbf{n}_{\beta'}$, for all $\beta < \beta' < \alpha$. Clearly any finite subset of this set has a model—namely, the integers. Thus, by compactness, so does the full set. Clearly the model has the properties we want.

4.3 Countably additive nonstandard probability measures

Do things get any better if we require countable additivity? To answer this question, we must first make precise what countable additivity means in the context of non-Archimedean fields. To understand the issue here, recall that for the standard real numbers, every bounded nondecreasing sequence has a unique least upper bound, which can be taken to be its limit. Given a countable sum each of whose terms is nonnegative, the partial sums form a nondecreasing sequence. If the partial sums are bounded (which they are if the terms in the sums represent the probabilities of a pairwise disjoint collection of sets), then the limit is well defined.

None of the above is true in the case of non-Archimedean fields. For a trivial counterexample, consider the sequence $\epsilon, 2\epsilon, 3\epsilon, \ldots$. Clearly this sequence is bounded (by any positive real number), but it does not have a least upper bound. For a more subtle example, consider the sequence $1/2, 3/4, 7/8, \ldots$ in the field $\mathbb{R}(\epsilon)$. Should its limit be 1? While this does not seem to be an unreasonable choice, note that 1 is not the least upper bound of the sequence. For example, $1 - \epsilon$ is greater than every term in the sequence, and is less than 1. So are $1 - 3\epsilon$ and $1 - \epsilon^2$. Indeed, this sequence has no least upper bound in $\mathbb{R}(\epsilon)$.

Despite these concerns, I define limits in $\mathbb{R}(I^*)$ pointwise. That is, a sequence $a_1, a_2, a_3, \ldots$ in $\mathbb{R}(I^*)$ converges to $b \in \mathbb{R}(I^*)$ if, for every $n \in I^*$, the coefficients of ϵ^n in $a_1, a_2, a_3, \ldots$ converge to coefficient of ϵ^n in b. (Since the coefficients are standard reals, the notion of convergence for the coefficients is just the standard definition of convergence in the reals. Of course, if ϵ^n does not appear explicitly, its coefficient is taken to be 0.) As usual, $\sum_{i=1}^{\infty} a_i$ is taken to be b if the sequence of partial sums $\sum_{i=1}^{n} a_i$ converges to b. Note that, with this notion of convergence, $1/2, 3/4, 7/8, \ldots$ converges to 1 even though 1 is not the least upper bound of the sequence.[3]

With this notion of countable sum, it makes perfect sense to consider countably-additive nonstandard probability measures. If $\mathcal{F}$ is a σ-algebra and $LPS^c(W, \mathcal{F})$ and $NPS^c(W, \mathcal{F})$ denote the countably additive LPS's and NPS's on $(W, \mathcal{F})$, respectively, then Proposition 4.6 can be applied with no change in proof to show the following.

Proposition 4.8: *$F_{L \to N}$ is an injection from* $\mathrm{LPS}^c(W, \mathcal{F})$ *to* $\mathrm{NPS}^c(W, \mathcal{F})$.

However, as the following example shows, even with the requirement of countable additivity, there are nonstandard probability measures that are not equivalent to any LPS.

Example 4.9: Let $W = \{w_1, w_2, w_3, \ldots\}$, and let $\mathcal{F} = 2^W$. Choose any nonstandard I^*. Define an NPS $(W, \mathcal{F}, \nu)$ with range $\mathbb{R}(I^*)$ by taking $\nu(w_j) = a_j + b_j\epsilon$, where $a_j = 1/2^j$, $b_{2j-1} = \epsilon/2^{j-1}$, and $b_{2j} = -\epsilon/2^{j-1}$, for $j = 1, 2, 3, \ldots$. Thus, the probabilities of $w_1, w_2, \ldots$ are characterized by the sequence $1/2 + \epsilon, 1/4 - \epsilon, 1/8 + \epsilon/2, 1/16 - \epsilon/2, 1/32 + \epsilon/4, \ldots$. For $U \subseteq W$, define $\nu(U) = \sum_{\{j: w_j \in U\}} a_j + \epsilon \sum_{\{j: w_j \in U\}} b_j$. It is easy to see that these sums are well-defined. As I show in the full paper, there is no LPS $\vec{\mu}$ over $(W, \mathcal{F})$ such that $\nu \approx \vec{\mu}$. ∎

[3]For those used to thinking of convergence in topological terms, what is going on here is that the topology corresponding to this notion of convergence is not Hausdorff.

5 Relating Popper Spaces to NPS's

Consider the map $F_{N\to P}$ from nonstandard probability spaces to Popper spaces such that $F_{N\to P}(W, \mathcal{F}, \nu)$ $(W, \mathcal{F}, \mathcal{F}', \mu)$, where $\mathcal{F}' = \{U : \nu(U) \neq 0\}$ and $\mu(V|U) = st(\nu(V|U))$ for $V \in \mathcal{F}$, $U \in \mathcal{F}'$. I leave it to the reader to check that $(W, \mathcal{F}, \mathcal{F}', \mu)$ is indeed a Popper space. Define an equivalence relation $\simeq$ on $NPS(W, \mathcal{F})$ (and $NPS^c(W, \mathcal{F})$ by taking $\nu_1 \simeq \nu_2$ if $\{U : \nu_1(U) = 0\} = \{U : \nu_2(U) = 0\}$ and $st(\nu_1(V|U)) = st(\nu_2(V|U))$ for all V, U such that $\nu_1(U) \neq 0$. Let $NPS/\simeq$ (resp., $NPS^c/\simeq$) consist of the $\simeq$ equivalence classes in NPS (resp., NPS^c). Clearly $F_{N\to P}$ is well defined as a map from $NPS/\simeq$ to $Pop(W, \mathcal{F})$ and from $NPS^c/\simeq$ to $Pop^c(W, \mathcal{F})$. As the following result shows, $F_{N\to P}$ is actually a bijection.

Theorem 5.1: *$F_{N\to P}$ is a bijection from* $\mathrm{NPS}(W, \mathcal{F})/\simeq$ *to* $\mathrm{Pop}(W, \mathcal{F})/\simeq$ *and from* $\mathrm{NPS}^c(W, \mathcal{F})/\simeq$ *to* $\mathrm{Pop}^c(W, \mathcal{F})/\sim$.

McGee [1994] proves essentially the same result as Theorem 5.1 in the case that $\mathcal{F}$ is an algebra (and the measures involved are not necessarily countably additive). McGee [1994, p. 181] says that his result shows that "these two approaches amount to the same thing". However, this is far from clear. The $\simeq$ relation is rather coarse. In particular, it is coarser than $\approx$.

Proposition 5.2: *If $\nu_1 \approx \nu_2$ than $\nu_1 \simeq \nu_2$.*

The $\simeq$ relation identifies nonstandard measures that behave quite differently in decision contexts. This difference already arises in finite spaces, as the following example shows.

Example 5.3: Suppose $W = \{w_1, w_2\}$. Consider the nonstandard probability measure ν_1 such that $\nu_1(w_1) = 1/2 + \epsilon$ and $\nu_1(w_2) = 1/2 - \epsilon$. (This is equivalent to the LPS (μ_1, μ_2) where $\mu_1(w_1) = \mu_2(w_2) = 1/2$, $\mu_2(w_1) = 1$, and $\mu_2(w_2) = 0$.) Let ν_2 be the nonstandard probability measure such that $\nu_2(w_1) = \nu_2(w_2) = 1/2$. Clearly $\nu_1 \simeq \nu_2$. However, it is not the case that $\nu_1 \approx \nu_2$. Consider the two random variables $\chi_{\{w_1\}}$ and $\chi_{\{w_2\}}$. (I use the notation χ_U to denote the indicator function for U; that is, $\chi_U(w) = 1$ if $w \in U$ and $\chi_U(w) = 0$ otherwise.) According to ν_1, the expected value of $\chi_{\{w_1\}}$ is (very slightly) higher than that of $\chi_{\{w_2\}}$, however, the expected value of $\chi_{\{w_1\}}$ is less than that of $\alpha\chi_{\{w_2\}}$ for any (standard) real $\alpha > 1$. According to ν_2, $\chi_{\{w_1\}}$ and $\chi_{\{w_2\}}$ have the same expected value. Thus, $\nu_1 \not\approx \nu_2$. Moreover, it is easy to see that there is no Popper measure μ on $\{w_1, w_2\}$ that can make the same distinctions with respect to $\chi_{\{w_1\}}$ and $\chi_{\{w_2\}}$ as ν_1, no matter how we define expected value with respect to a Popper measure. ∎

More generally, Theorem 3.1 shows that, in a precise sense, Popper spaces are equivalent to SLPS's, while Theorem 4.4 shows that LPS's are equivalent to NPS's. Thus, there is a gap in expressive power between Popper spaces and NPS's that essentially amounts to the gap between SLPS's and LPS's.

6 Discussion

As the preceding discussion shows, there is a sense in which NPS's more general than both Popper spaces and LPS's. LPS's are more expressive than Popper measures in finite spaces and in infinite spaces where we assume countable additivity (in the sense discussed at the end of Section 5), but without assuming countable additivity, they are incomparable, as Example 3.2 shows. Although NPS's are equivalent to LPS's in finite state spaces, NPS's have other advantages. For example, as pointed out by Hammond [1994] and BBD, it is easier to define independence in NPS's.

On the other hand, NPS's also have some disadvantages. In particular, working with a nonstandard probability measure requires defining and working with a non-Archimedean field. LPS's have the advantage of using just standard probability measures. Moreover, their lexicographic structure may give useful insights. It seems to be worth considering the extent to which LPS's can be generalized so as to increase their expressive power. I am currently exploring LPS's ordered by an arbitrary (not necessarily well-founded) index set. It seems that such LPS's may be useful in characterizing iterated deletion of weakly dominated strategies. (This is done by Brandenburger and Keisler [2000] using finite LPS's; it seems that results are more cleanly stated using infinite LPS's ordered by the integers.) I hope to report on this in future work.

One final point: defining belief. Brandenburger and Keisler [2000] defined a notion of belief using LPS's and provided an elegant decision-theoretic justification of it. According to their definition, an agent *believes* U *in LPS* $\vec{\mu}$ if there is some $j \leq m$ such that $\mu_i(U) = 1$ for all $i \leq j$ and $\mu_i(U) = 0$ for $i > j$. Independently, van Fraassen [1995] defined a notion of belief using Popper spaces that can be shown to be essentially equivalent to the definition given by Brandenburger and Keisler. (See [Arlo-Costa and Parikh 1999] for a followup to van Fraassen's work.) That there should be equivalent notions of belief in the context of LPS's and Popper spaces is perhaps not that surprising, in light of the close connection between them. The results of this paper suggest that it may also be worth considering notions of belief defined in NPS's.

Acknowledgments: I'd like to thank Adam Brandenburger and Peter Hammond for a number of very enlightening discussions, Bob Stalnaker for pointing out Example 3.2, Brain Skyrms for pointing me to Hammond's work, Bas van Fraassen for pointed me to Spohn's work, and Eddie Dekel and Horacio Arlo-Costa for a number of useful comments on an earlier draft of this paper.

References

Adams, E. (1966). Probability and the logic of conditionals. In J. Hintikka and P. Suppes (Eds.), *Aspects of Inductive Logic*, pp. 265–316. Amsterdam: North Holland.

Arlo-Costa, H. and R. Parikh (1999). Two place probabilities, beliefs, and belief revision: on the foundations of iterative belief kinematics. In P. Dekker (Ed.), *Proc. 12th Amsterdam Colloquium*, pp. 1–6.

Blume, L., A. Brandenburger, and E. Dekel (1991a). Lexicographic probabilities and choice under uncertainty. *Econometrica 59*(1), 61–79.

Blume, L., A. Brandenburger, and E. Dekel (1991b). Lexicographic probabilities and equilibrium refinements. *Econometrica 59*(1), 81–98.

Brandenburger, A. and J. Keisler (2000). Epistemic conditions for iterated admissibility. Unpublished manuscript.

Császár, A. (1955). Sur la structure des espace de probabilité conditionelle. *Acta Mathematica Academiae Scientiarum Hungaricae 6*, 337–361.

Davis, M. (1977). *Applied Nonstandard Analysis*. New York: Wiley.

Enderton, H. B. (1972). *A Mathematical Introduction to Logic*. New York: Academic Press.

de Finetti, B.. (1936). Les probabilités nulles. *Bulletins des Science Mathématiques (première partie) 60*, 275–288.

Hammond, P. J. (1994). Elementary non-Archimedean representations of probability for decision theory and games. In P. Humphreys (Ed.), *Patrick Suppes: Scientific Philosopher; Vol. 1*. Kluwer.

Keynes, J. M. (1921). *A Treatise on Probability*. London: Macmillan.

Kreps, D. M. and R. B. Wilson (1982). Sequential equilibria. *Econometrica 50*, 863–894.

Lehmann, D. and M. Magidor (1992). What does a conditional knowledge base entail? *Artificial Intelligence 55*, 1–60.

McGee, V. (1994). Learning the impossible. In E. Eells and B. Skyrms (Eds.), *Probability and Conditionals*. Cambridge University Press.

Myerson, R. (1986). Multistage games with communication. *Econometrica 54*, 323–358.

Popper, K. R. (1968). *The Logic of Scientific Discovery (revised edition)*. London: Hutchison. The first version of this book appeared as *Logik der Forschung*, 1934.

Rényi, A. (1955). On a new axiomatic theory of probability. *Acta Mathematica Academiae Scientiarum Hungaricae 6*, 285–335.

Rényi, A. (1956). On conditional probability spaces generated by a dimensionally ordered set of measures. *Theory of Probability and its Applications 1*, 61–71. Reprinted as paper 120 in *Selected Papers of Alfred Rényi, I: 1948–1956*, Akadémia Kiadó, 1976, pp. 554–557.

Rényi, A. (1964). Sur les espaces simples de probabilités conditionelles. *Annales de l'Institut Henri Poincaré, Nouvelle série, Section B 1*, 3–21. Reprinted as paper 237 in *Selected Papers of Alfred Rényi, III: 1962–1970*, Akadémia Kiadó, 1976, pp. 284–302.

Robinson, A. (1973). Function theory on some nonarchimedean fields. *American Mathematical Monthly: Papers on the Foundations of Mathematics 80*, S87–S109.

Selten, R. (1965). Spieltheoretische behandlung eines oligopolmodells mit nachfrageträgheit. *Zeitschrift für Gesamte Staatswissenschaft 121*, 301–324 and 667–689.

Spohn, W. (1986). The representation of Popper measures. *Topoi 5*, 69–74.

Spohn, W. (1988). Ordinal conditional functions: a dynamic theory of epistemic states. In W. Harper and B. Skyrms (Eds.), *Causation in Decision, Belief Change, and Statistics*, Volume 2, pp. 105–134. Dordrecht, Netherlands: Reidel.

van Fraassen, B. C. (1976). Representation of conditional probabilities. *Journal of Philosophical Logic 5*, 417–430.

van Fraassen, B. C. (1995). Fine-grained opinion, probability, and the logic of full belief. *Journal of Philosophical Logic 24*, 349–377.

Wilson, N. (1995). An order of magnitude calculus. In *Proc. Eleventh Conference on Uncertainty in Artificial Intelligence (UAI '95)*, pp. 548–555.