International Journal of Approximate Reasoning 53 (2012) 293–315

Contents lists available at SciVerse ScienceDirect

International Journal of Approximate Reasoning

journal homepage: www.elsevier.com/locate/ijar

Belief and probability: A general theory of probability cores ☆

Horacio Arló-Costa, Arthur Paul Pedersen *

Department of Philosophy, Carnegie Mellon University, Pittsburgh, PA 15213, USA

ARTICLE INFO

Article history:
Received 30 May 2011
Received in revised form 31 December 2011
Accepted 7 January 2012
Available online 26 January 2010

Keywords:
Henry E. Kyburg
Probability cores
Acceptance rules
Full belief
Ordinary belief
Expectation

ABSTRACT

This paper considers varieties of probabilism capable of distilling paradox-free qualitative doxastic notions (e.g., full belief, expectation, and plain belief) from a notion of probability taken as a primitive. We show that *core systems*, collections of nested propositions expressible in the underlying algebra, can play a crucial role in these derivations. We demonstrate how the notion of a probability core can be naturally generalized to high probability, giving rise to what we call a *high probability core*, a notion that when formulated in terms of classical *monadic* probability coincides with the notion of *stability* proposed by Hannes Leitgeb [32]. Our work continues by one of us in collaboration with Rohit Parikh [7]. In turn, the latter work was inspired by the seminal work of Bas van Fraassen [46]. We argue that the adoption of *dyadic* probability as a primitive (as articulated by van Fraassen [46]) admits a smoother connection with the standard theory of probability cores as well as a better model in which to situate doxastic notions like *full belief*. We also illustrate how the basic structure underlying a system of cores naturally leads to alternative probabilistic acceptance rules, like the so-called *ratio rule* initially proposed by Isaac Levi [34].

Core systems in their various guises are ubiquitous in many areas of formal epistemology (e.g., belief revision, the semantics of conditionals, modal logic, etc.). We argue that core systems can also play a natural and important role in Bayesian epistemology and decision theory. In fact, the final part of the article shows that probabilistic core systems are naturally derivable from basic decision-theoretic axioms which incorporate only qualitative aspects of core systems; that the qualitative aspects of core systems alone can be naturally integrated in the articulation of coherence of primitive conditional probability; and that the guiding idea behind the primary qualitative features of a core system gives rise to the formulation of lexicographic decision rules.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Personal probability occupies a central role in Bayesian epistemology, delivering rational *quantitative* belief, or degrees of belief, thus understood to observe the laws of probability. These laws furnish standards of consistency for a body of quantitative beliefs, much in the spirit of the laws of logic. Rational quantitative belief itself serves to ground the consistency

☆ We dedicate this paper to the memory of Henry E. Kyburg, Jr., whose ideas inspired some of the main results presented here. One of us (Horacio Arló-Costa) had the good fortune to know Henry personally and to correspond with him, something which always was extremely insightful and helpful. Although he was a man of deep convictions, he was also capable of understanding and articulating in non-trivial ways alternative ideas. His views of the lottery changed during the last years of his life, as we explain below, and we can only speculate how he would have reacted to recent responses to his paradox. What is true is that his ideas have lasting value and the paradoxes he addressed continue to inspire innovative work in formal epistemology. We are happy to contribute to his legacy by offering new responses to his many insightful questions. We still have much to learn from Henry's work, and we look forward to many years of fruitful discussion of his work.
This paper is also dedicated to the memory of Horacio Arló-Costa.

* Corresponding author.
E-mail address: apaulpedersen@cmu.edu
URLs: http://www.hss.cmu.edu/philosophy/faculty-arlocosta.php, http://andrew.cmu.edu/~ppederse

0888-613X/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.ijar.2012.01.002

of an agent's decisions, supporting the explication and application of theories of rational decision making, giving rise to an important and vibrant area of philosophical and scientific research.

In traditional epistemology, the origin of many of the fascinating philosophical problems addressed in Bayesian epistemology, rational *qualitative* belief is the predominant notion of philosophical interest. Itself understood to observe the laws of logic, rational qualitative belief occupies a coarser scale than degrees of belief. In spite of this, one may recognize a variety of related notions of rational qualitative belief, the notion of *certainty*, or *full* belief, being undoubtedly among them, and perhaps a notion of *plain*, or *ordinary*, belief belonging among them as well. Full belief, tied to a strong epistemic commitment, requires probability one, whereas ordinary belief does not. Intermediate among these notions one might also locate other notions such as qualitative *expectations*, as understood among philosophers, logicians, computer scientists, and others.

While degrees of belief and notions of rational qualitative belief are undoubtedly related—something we take for granted in this article—the form of the relationship has remained elusive. Henry Kyburg, amongst other contemporary philosophers, helped to make explicit the difficulties involved in relating rational degrees of belief with rational plain belief or certainty. One natural suggestion to bridge the gap between qualitative and quantitative belief is to adopt an *acceptance rule* according to which an agent believes propositions carrying high probability. Thus, an agent believes a proposition just in case it exceeds a fixed threshold r. Kyburg [30] showed early on that this rule leads to paradoxical conclusions, pointing to what has become known as the *lottery paradox*.

Example 1.1 (*Kyburg*)**.** Consider an agent confronted with a fair lottery with 1 million tickets. For each i, the agent considers highly probable (.999999) that ticket number i will not be the winning ticket. According to the present proposal, where r is less than .999999, he will believe that each lottery ticket is not the winning ticket and that some ticket is a winning lottery ticket.

If we understand rational quantitative belief and qualitative belief to be equipped with normative standards, such a rule, therefore, leads to paradoxical results. In fact, the above example illustrates that a weaker form of the rule, according to which high probability is sufficient for belief, leads to trouble.

Another natural suggestion intended to bridge the gap is to adopt an acceptance rule according to which propositions with probability one are precisely those which are fully believed. While this rule may be too restrictive for ordinary belief, even for full belief such a proposal faces difficulties.

Example 1.2. Consider a fair coin flipped until you see heads. The probability that the coin will lands heads in n flips is $\frac{1}{2^n}$. The agent maintains that it is possible that the coin never lands heads, yet according to the latest proposal, the agent fully believes that the coin will land heads.

Thus, where full belief is understood as an agent's standard for serious possibility, this example conflicts with the requirement that propositions which are seriously possible must be compatible with an agent's full beliefs [33]. While replacing full belief with qualitative *expectation* certainly appears to get things right—indeed, the agent may reasonably be willing to place any amount of money on a bet that the coin lands heads—a rule which *requires* that an agent fully believe that the coin land will land heads is itself too restrictive for an adequate account of full belief.

Moreover, such a proposal is met with a transfinite version of the lottery paradox, as Patrick Maher [35], among others, has remarked.

Example 1.3 (*Maher*)**.** Suppose that an agent assumes that the weight of a stone can be represented by a real number in some interval, say, between .5 and 1 pounds. The agent fully believes that the weight of the stone lies somewhere in the interval [.5, 1]. To be sure, the agent assigns probability 1 to the proposition that the weight of the stone lies in the interval [.5, 1]. In addition, for each x in the interval [.5,1], the agent judges that the weight of the stone is exactly x pounds carries probability zero, as the agent regards it equiprobable that for any x in [.5, 1], the stone has weight x. Thus, that the stone is exactly x pounds will carry probability one. But then if propositions of probability one are fully believed, the agent is certain that the weight of the stone does not lie in [.5, 1].

Here, the notion of serious possibility need not be invoked to demonstrate that the present proposal attempting to relate degrees of belief to full belief is faced with difficulties. One may squirm over the infinitary nature of such a paradox, but we will not allow ourselves to become distracted by such dismissive finitists.

In fact, there are many possible principled reactions to the lottery in its finite and infinite versions. One solution that Kyburg endorsed restores consistency by weakening the underlying logic. Thus, one adopts a drastic solution, abandoning the so-called rule of *adjunction* of logic [31]. One of us [4] has suggested a second, less draconian solution, according to which one may retain the laws of logic but articulate the notion of belief in a suitable epistemic logic in which the axiom $(\Box\phi \wedge \Box\psi) \rightarrow \Box(\phi \wedge \psi)$ is permitted to fail, the modality being interpreted as a high probability operator.

If the underlying logic is stronger—for example, if it has the power of at least a first-order epistemic logic—one must abandon the quantificational version of the aforementioned axiom, the so-called *Barcan schema*, $\forall x\Box\phi \rightarrow \Box\forall x\phi$, resulting in classical non-normal modal logics. As such, the familiar Kripkean semantics cannot be used to study them, but a semantics deriving from the pioneering work of Dana Scott [41] and Richard Montague [36] meets the task.

Henry Kyburg and Choh Man Teng creatively endorse the second approach in a conference paper [28]. An extension of this paper is being published now in the present issue of this journal [29].

Of course, these two strategies (weakening classical logic by adopting a non-adjunctive logic; appealing to non-normal epistemic operators) are related. It is easily seen that some of the most salient non-adjunctive logics can be mapped to non-normal epistemic operators [5]. We can call this unified strategy the *non-adjunctive strategy*.

These solutions, unfortunately, will fail to satisfy those who are convinced that plain (or full) belief is closed under conjunction, whether these solutions appear in their modal or classical varieties. The non-adjunctive strategy to the paradox of the lottery consists in arguing that a *different* notion of rational qualitative belief—one that does not observe classical logical properties—can be smoothly connected with rational degrees of belief. Kyburg argued in various writings that this non-adjunctive notion of belief is epistemologically self-sufficient. As a non-adjunctive *cousin*, the notion of belief, it is claimed, gives us a workable notion of serious possibility and can be used to guide rational action.

Yet, of course, we ask: What about the standard notion of rational belief? If one thinks that it obeys the standard laws of logic, ought one conclude that it is incompatible with or underivable from probability? If this indeed were the case, perhaps a Bayesian epistemology ought to adopt more than one primitive. Belief and probability would then each assume a primitive status, now being mutually irreducible. Such is the view of, for example, Issac Levi [33]. Unfortunately, many probabilists have found this strategy unpalatable. In fact, if one does endorse a *probabilistic* stance recognizing qualitative notions of rational belief, one would like to explicate non-probabilistic notions in terms of probability. Is this impossible?

Bas van Fraassen [46], concerned with some of the foregoing problems and related issues, has elegantly expresses a viewpoint we share concerning attempts to relate belief to degrees of belief:

> Personal or subjective probability entered epistemology as a cure for certain perceived inadequacies in the traditional notion of belief. But there are severe strains in the relationship between probability and belief. They seem too intimately related to exist as separate but equal; yet if either is taken as the more basic, the other may suffer [46, p. 349].

Van Fraassen thereupon proposed an unified probabilistic explication of the notion of full belief, taking *conditional probability* as basic. In addition, van Fraassen, who contends that unconditional beliefs are inadequate to account for the variety of an agent's epistemic attitudes, appealed to the notion of *supposition*:

> There is a third aspect of opinion, besides belief and subjective grading, namely *supposition*. Much of our opinion can be elicited only by asking us to suppose something, which we may or may not believe. The respondent imaginatively puts himself in the position of someone for whom the supposition has some privileged status. But if his answer is to express his present opinion—which is surely what is requested—then this "momentary" shift in status must be guided by what his present opinion is [46, p. 351].

Subsequently, van Fraassen adopted a form of probabilism that takes conditional probability as primitive function as a primitive and suggests how to derive full belief from it by appealing to an interpretation in terms of the notion of supposition. One of us undertook to extend, apply, and revise van Fraassen's proposal (see [1–3,6,7], and the references therein).

The central idea of this modified proposal consists in showing that a primitive conditional probability function induces a *core system* of nested propositions expressible in the underlying algebra under consideration. A core system encodes an agent's judgements of the relative plausibilities of events of zero probability. In the presence of countable additivity, a core system possesses a smallest core, and provided that the algebra is closed under arbitrary unions, it also contains a largest core (the union of all cores). The view that one of us has endorsed is that the smallest core encodes a notion of 'near certainty', or qualitative expectation, while the largest core encodes that of full belief or certainty. The largest core carries measure one and so all full beliefs—that is, those propositions entailed by the largest core—have probability one, but there are probability one propositions which are not fully believed. As argued in a series of articles [1–3,6,7], this strategy offers a principled way to address the transfinite lottery paradox. Probability one is not enough to ground full belief.

This strategy, which we review in Section 2, represents a promising way of proceeding to arrive at a fuller explication of belief in terms of degrees of belief. Furthermore, among other fruits of the resulting account is an elegant probabilistic semantics for conditionals and non-monotonic notions of consequence. Still, this strategy does not say anything about the weaker notion of plain belief, for a proposition could be plainly believed even when it does not carry probability one.

Our goals in this article are rather simple. Among other things, we will endeavor to show that the main ideas proposed by van Fraassen can be further extended and modified to explicate the notion of plain belief and to address the standard version of the lottery paradox. In Section 3, we will propose a natural extension of the notion of a probability core, which we call a *high probability core*. We will see that high probability cores, when derived from classical *monadic* probability, are logically equivalent to the notion of *stability* as proposed by Hannes Leitgeb as part of an explication of belief in terms of degrees of belief [32]. Stable sets enjoy various interesting properties, but they are tied to classical monadic probability, preventing a smooth connection with the standard theory of probability cores, as primitive conditional probability is an essential ingredient in this theory. Nevertheless, as we show in Section 4, it is easy to formulate a notion of a high probability core derived from *dyadic* probability (using the definition thereof presented in [46]). This account meshes rather nicely with the standard theory of cores, naturally extending the theory to high probability. As we will show, the central idea behind the core construction finds application beyond its initial formulation.

High probability acceptance rules are not the only type of rules which have been proposed to connect belief and degrees of belief. For example, Isaac Levi [34] has proposed an alternative acceptance rule he calls the *ratio rule*. Although this rule has a pedigree connected with cognitive decision theory, the ratio rule is formulated purely in probabilistic terms. In Section 4 we briefly discuss how a version of the ratio rule can also be used to define a variant of probabilty cores we call *ratio cores*. Ratio cores enjoy many of the formal properties of standard probability cores and high probability cores. But one consequence of the rule is that the innermost ratio core need not carry probability one or even high probability. Philosophers who think that high probability is a necessary condition for belief can use ratio cores which qualify as high probability cores and possess nice dynamic properties (not necessarily satisfied by the simple version of high probability cores).

After revisiting the lottery paradox in Section 5, we continue with a discussion of rational decision making. Thus, in Section 6 we offer a representation result showing that probability core systems are naturally derivable from basic, compelling qualitative decision-theoretic axioms which incorporate only qualitative aspects of core systems. In fact, we will see that the qualitative features underlying probability core system can play a natural role in the formulation of a Dutch Book argument for primitive conditional probability and more generally conditional expectation. These features also give rise to a lexicographic decision rule.

The unified picture emerging from this article is that core systems in their different guises can play a crucial role in the formulation of a variety of paradox-free acceptance rules linking rational belief and degrees of belief and decision rules linking preference and value. While we do not claim to have provided *the* account of qualitative belief or any other expressible epistemic notion, we believe that our account represents an attractive and principled candidate to serve as the foundation for a unified epistemology, delivering sophisticated versions of probabilism and more complex epistemological accounts in which cognitive utility grounds ampliative rules. Moreover, we believe that our account will bear fruit in areas well-beyond the scope of this article. In Section 7 we close with concluding remarks and a discussion about ongoing and future work.

2. Dyadic probability and probability cores

Let us begin with the basic idea that propositions are sets of possibilities from a space W. Thus, propositions, denoted by the letters A, B, C, etc., are subsets of W.

What basic structural features should we require a collection of propositions to satisfy? A mild requirement is that the set of propositions in question be closed under logical operations—that it forms an algebra. We will use the notation $\overline{A}$ to denote the absolute complement of a proposition A; $\subseteq$ to denote subset inclusion; and $\subset$ to denote proper subset inclusion. We appeal to the usual symbols for intersection and union. We can now make some of the foregoing ideas more precise.

Definition 2.1. A collection $\mathscr{A}$ of subsets of a set W is called an *algebra* of sets (or *field* of sets) over W if it contains W itself and is closed under the formation of complements and finite unions:

(i) $W \in \mathscr{A}$;
(ii) If $A \in \mathscr{A}$, then $\overline{A} \in \mathscr{A}$;
(iii) If $A, B \in \mathscr{A}$, then $A \cup B \in \mathscr{A}$.

The collection $\mathscr{A}$ is called a σ*-algebra* of sets (or a σ*-field* of sets) over W if it is an algebra and it is also closed under countable unions:

(iv) For every collection $\{A_n\}_{n=1}^{\infty} \subseteq \mathscr{A}$, $\bigcup_{n=1}^{\infty} A_n \in \mathscr{A}$.

We call an element A of $\mathscr{A}$ a *proposition* (or an *event*) from $\mathscr{A}$.

The distinction between an algebra and σ-algebra is relevant when W is infinite, collapsing otherwise. Now, since primitive conditional probability plays a central role in the theory of probability cores, we ought to clarify what we mean by it. Accordingly, we offer a definition (cf. [46]).

Definition 2.2. Let $\mathscr{A}$ be an σ-algebra on W. A two-place *probability measure* on $\mathscr{A}$ is a mapping $P : \mathscr{A} \times \mathscr{A} \to \mathbb{R}$ such that for every $A \in \mathscr{A}$:

(I) Either:
  (a) $P(\cdot|A)$ has constant value 1;
  or
  (b) $P(\cdot|A)$ is a countably additive probability measure, i.e.,
    (1) $P(B|A) \geq 0$ for every $B \in \mathscr{A}$
    (2) $P(W|A) = 1$
    (3) For every pairwise disjoint collection $(B_n)_{n<\omega} \subseteq \mathscr{A}$:

$$P\left(\bigcup_{n<\omega} B_n|A\right) = \sum_{n<\omega} P(B_n|A);$$

(II) $P(A|A) = 1$;
(III) $P(B \cap C|A) = P(B|A)P(C|B \cap A)$ for all $B, C \in \mathscr{A}$.

Definition 2.2 allows for conditioning on the absurd event $\emptyset$, but in such a case the definition demands that for every event A, $P(A|\emptyset) = 1$. In fact, part I.a is intended to represent the result of supposing a proposition the agent regards as impossible. This feature deviates from standard presentations of probabilistic belief states, even for primitive conditional probability (cf. [16,18,20,27,39,40]), although it aligns with the presentations of Harper [23], Popper [37] and van Fraassen [46]. For those who find such a deviation difficult to digest, we point out that one may use what could be thought of as a more conservative, and yet essentially equivalent, definition of a two-probability measure, according to which the requirement that P be defined on $\mathscr{A} \times \mathscr{A}$ is replaced by the requirement that P be defined on $\mathscr{A} \times \mathscr{B}$, where $\mathscr{B}$ is a subcollection of $\mathscr{A}$ excluding the absurd event such that $P(A|B) > 0$ implies that $B \in \mathscr{B}$, and the option that $P(\cdot|B)$ have constant value 1 is no longer available. Such a reformulation forbids inconsistent suppositional states, and the theory presented here could be easily articulated using this reformulation instead.

We take property II to be partly constitutive of any notion of conditional probability, although other accounts—notably the standard Kolmogrov account of regular conditional distributions—do not require this property ([11,10,42] contains a valuable discussion of the extent of *impropriety* of regular conditional distributions—those that, roughly speaking, violate property II). Property III has a long history going back at least to Jeffreys and to Keynes. It captures the idea that conditioning preserves ratios, even when conditioning on events of unconditional probability zero. We will follow established terminology by referring to III as the *Multiplication Axiom*. [1]

In this article we require that P be countably additive when it is an additive measure at all. It is well-known that the requirement of countable additivity faces difficulties with existence. A more liberal approach would demand only that P be finitely additive. Such an approach does not face the same difficulties with existence. However, as we will make an effort to emphasize, some of the results for probability cores no longer obtain when the requirement of countable additivity is dropped.

Before we continue, we wish to point out that Definition 2.2 requires that probability be precise. Although in this article we assume that probabilistic assessments are precise, we recognize that a more satisfactory theory would relax this assumption. Indeed, while rational quantitative belief is understood to observe the laws of probability, such an understanding leaves room for degrees of belief to be imprecise, reflecting, for example, an agent's uncertainty concerning particular quantities when on the basis of the available evidence, making a numerically precise probability judgment would be unwarranted and arbitrary. Nonetheless, the standard theory of cores assumes that probabilities are precise, and for our present purposes, we find it better to retain this assumption.

When needed, we will refer to the probability (*simpliciter*) of a proposition $A \in \mathscr{A}$, $p(A)$, which is simply $P(A|W)$. Given $A \in \mathscr{A}$, we call A normal if $P(\cdot|A)$ is a probability measure and *abnormal* otherwise, i.e., if $P(\cdot|A)$ has constant value 1, so in particular, $P(\emptyset|A) = 1$. Thus A is normal just in case $P(\emptyset|A) = 0$. We emphasize that normal propositions *can* have probability 0. For example, under Lebesgue measure, the rationals as a subset of the reals comprise a normal set of probability 0. It follows that the absurd proposition $\emptyset$ has probability 0 conditional on the set of rationals in the interval [0,1]. By contrast, an abnormal proposition not only is assigned the value 0 but also leads to assigning *any* proposition the value 1 if it is conditioned upon. In addition, abnormal propositions contained in normal propositions have probability 0, so if the W is normal (which is the case if P is not the constant function 1), then all abnormal propositions have probability 0. To see this, observe that if $A \subseteq B$, B is normal, and A is abnormal, then we have $P(\emptyset|B) = P(\emptyset|A)P(A|B)$, so since $P(\emptyset|B) = 0$ and $P(\emptyset|A) = 1$, it must be the case that $P(A|B) = 0$. Van Fraassen [46] establishes that supersets of normal propositions are normal and that subsets of abnormal propositions are abnormal. Some useful facts about normal and abnormal propositions can be found in [7].

Modifying van Fraassen's presentation, we now introduce the notion of *superiority* or *domination*. Given propositions A and B, say that A *dominates* B, written $A >_P B$, just in case $P(B|A \cup B) = 0$. The idea is that event A is infinitely more "expected" than B. In terms of gambling, any bet on A for a small dollar amount is strictly preferred to having a bet on B. This notion goes as far back as [14], with different formulations in Rényi [39], Krauss [27], and van Fraassen [46]. We consider variations of this notion and the different notions of cores to which they give rise.

Finally, with the notion of superiority in place, we introduce the standard concept of a *core*.

Definition 2.3 (*P-core*). Let P be a two-place probability measure on $\mathscr{A}$, and let $K \in \mathscr{A}$. We call a K a *core* if it satisfies the *strong superiority condition*:

For every $A, B \in \mathscr{A}$,
If $A \subseteq K$ is nonempty and $K \cap B = \emptyset$, then $A >_P B$.

[1] This axiom appears under the name 'W. E. Johnson's product rule' in [25].

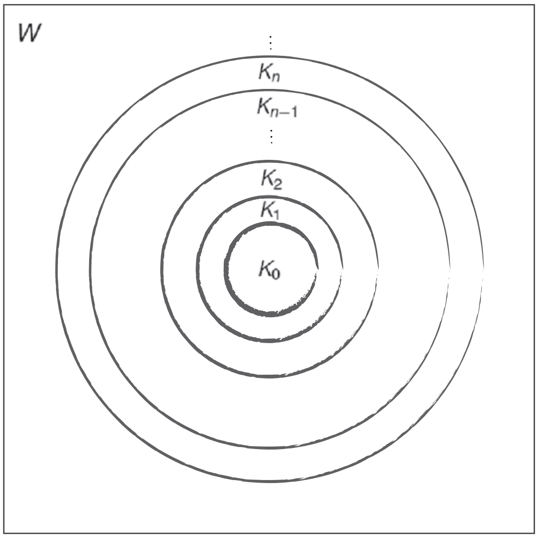

Fig. 1. Probability core system.

Thus *any* non-empty subset of K dominates any proposition outside of K. In a sense, any such proposition within K is infinitely more plausible than any proposition incompatible with K. In terms of supposition, a core K enjoys the property that under the supposition that either a consistent proposition entailing K or a proposition entailing $\overline{K}$ is the case, the degree of belief in the proposition entailing that K is false is infinitely small and so the degree of belief in the proposition entailing K is maximal. Hence, a core places propositions which entail it is false far below those consistent propositions which entail it is true. Moreover, all consistent cores carry probability 1.

While all supersets of normal sets are normal, all non-empty subsets of a core are normal. Let $\mathscr{C}_P$ denote the collection of all consistent cores associated with P. Of course, we drop the subscript when the context is clear. The following property is an elementary consequence of the definitions (cf. [46]).

Proposition 2.4 (Finesse)**.** *All non-empty subsets of a core K for P are normal.*

The next property is an essential feature of cores (cf. [46]).

Proposition 2.5. *The family of cores $\mathscr{C}_P$ is nested, i.e. given any two cores K_1, K_2 either $K_1 \subseteq K_2$ or $K_2 \subseteq K_1$.*

A central result concerning cores was proved in [1].

Theorem 2.6 (Descending chains)**.** *The system of cores $\mathscr{C}_P$ does not contain an infinitely descending chain of cores.*

Thus, probability cores are well-ordered with respect to inclusion, closely resembling Grove spheres [21] and Spohn's ordinal conditional functions [44]. Figure 1 depicts a core system. The smallest core is K_0, and $K_n \subseteq K_{n+1}$ for all. Core systems are abundant and available in any order type. In fact, for any nonzero ordinal α, there is a two-place probability measure P_α on $\mathscr{P}(\alpha)$ such that the family of its cores has ordinal length α (ordered by $\subseteq$, of course). To see this, define P_α on α by setting for every $A, B \subseteq \alpha$, $P_\alpha(A|B) := 1$ if $\min(A \cap B) = \min(B)$ and $A \cap B \neq \emptyset$ or $B = \emptyset$, and $P(A|B) := 0$ otherwise. It is easy to verify that P_α is a two-place probability function on $\mathscr{P}(\alpha)$. Every ordinal $\beta \leq \alpha$ is a core, for if $A \subseteq \beta$ is nonempty and $B \cap \beta = \emptyset$, then B is either 0 or B only contains ordinals following β, whence $P_\alpha(B|A \cup B) = 0$. In addition, no other subset of $\mathscr{P}(\alpha)$ is a core, for if $K \subseteq \alpha$ is a nonempty set which is not an ordinal, then there are ordinals β, γ such that $\beta \in K$ and $\gamma \in \beta \backslash K$, so $P_\alpha(\{\gamma\}|\{\beta\} \cup \{\gamma\}) = 1$ and therefore K is not a core.

As indicated above, some results for high probability cores no longer hold when countable additivity is dropped. To take a simple example in the present context, if P is not countably additive, $\mathscr{C}_P$ may have an infinite descending chain of cores. To see this, take $W = \omega$ (the first infinite ordinal) and $\mathscr{A}$ to consist of all finite and co-finite subsets of ω, and define a two-place probability *function* P by setting for every $A, B \in \mathscr{A}$:

$$P(A|B) := \begin{cases} 1 & \text{if } A, B \text{ are co-finite;} \\ 1 & \text{if } \max(A \cap B) = \max(B) \text{ and } B \text{ is finite;} \\ 0 & \text{otherwise.} \end{cases}$$

Clearly P is not countably additive, and if we set $C_n := \omega \backslash n$, then for every $n < \omega$, C_n is a core and $C_n \supset C_{n+1}$, so $\mathscr{C}_P$ is not well-ordered with respect to inclusion and in particular does not contain a smallest core.

When W is countable, probability cores as studied in [7] have nice properties, such as the following: [2]

Proposition 2.7. *If the space W is countable, then there is a largest core.*

Again, the smallest core (and hence every core) has probability 1, being composed of propositions carrying positive probability, so in particular the largest core, which is the union of all probability cores, has probability 1.

If W is uncountable, however, there may not be a largest core. Consider $W := \omega_1 \cdot 2$, where ω_1 is the first uncountable ordinal, and let $\mathscr{A}$ be the collection of countable and co-countable subsets of W. Define a two-place probability measure P on $\mathscr{A}$ by setting for every $A, B \in \mathscr{A}$, $P(A|B) := 1$ if $\min(A \cap B) = \min(B)$ and $A \cap B \cap \omega_1 \neq \emptyset$ or $B \cap \omega_1 = \emptyset$, and $P(A|B) := 0$ otherwise. Then all and only countable ordinals are cores, but $\bigcup_{\beta<\omega_1} \beta = \omega_1$ is not a core.

When there is a largest core, as when the space W is countable, the proposal advanced in [2] is to explicate full belief in terms of the largest core. Thus, a proposition is believed just in case it is entailed by the largest core. To accommodate cases in which there is no largest core, we adopt the following definition.

Definition 2.8 (*Full belief*)**.** Let P be a two-place probability measure on $\mathscr{A}$, and let $B \in \mathscr{A}$. We shall say that B is *fully believed* with respect to P if it is a superset of the union of all cores in $\mathscr{C}_P$.

Hence, a full belief is entailed by *all* cores and has maximal probability, and as such, is most entrenched and certain among an agent's epistemic commitments. To be sure, an agent cannot consistently suppose a proposition entailing that his full beliefs are false without being in an absurd state of opinion in which he judges all propositions maximally likely. An agent's *expectations* are less demanding, yet still require probability 1, as attested by the following explication.

Definition 2.9 (*Expectation*)**.** Let P be a two-place probability measure on $\mathscr{A}$, and let $E \in \mathscr{A}$. We shall say that a proposition E is *expected* (or *almost certain*) with respect to P if it is a superset of the smallest core in $\mathscr{C}_P$.

Unlike full belief, an agent *can* consistently suppose a proposition incompatible with his expectations. An agent's expectations are stronger than his full beliefs but do not conflict with them, expressing what the agent anticipates to be the case—indeed, with probability 1.

This review points to an important feature of the theory under consideration, viz., while the foregoing epistemic notions obey classical closure properties, they are explicated in terms of primitive conditional probability. We refer the reader to [6] and [1] and in particular [2,3] and [7] for a fuller discussion of the theory outlined here and its interesting connections to conditionals and belief dynamics.

3. Monadic probability and high probability cores

In this section we turn to high probability cores. But we begin with a different probabilistic primitive: classical *monadic* probability. While, as we shall see, the theory of probability cores is formulated in terms of dyadic probability, we take a first step towards the articulation of the extended theory using monadic probability to facilitate a clear discussion of its relationship to the theory proposed in [32]. For the sake of clarity, we present the definition of monadic probability due to Kolmogorov [26].

Definition 3.1. Let $\mathscr{A}$ be an algebra over W. A *probability function* on $\mathscr{A}$ is a non-negative, normalized, and finitely-additive real-valued function P on $\mathscr{A}$:

(i) $P(A) \geq 0$ for every $A \in \mathscr{A}$; (Non-Negativity)

(ii) $P(W) = 1$; (Normalization)

(iii) For every $A, B \in \mathscr{A}$ such that $A \cap B = \emptyset$, (Finite Additivity)
$P(A \cup B) = P(A) + P(B)$.

If $\mathscr{A}$ is in addition a σ-algebra over W, then P is a (σ*-additive*) *probability measure* on $\mathscr{A}$ if it is a probability function such that for every pairwise disjoint collection $(A_n)_{n<\omega} \subseteq \mathscr{A}$:

(iv) $P(\bigcup_{n<\omega} A_n) = \sum_{n<\omega} P(A_n)$. ($\sigma$-additivity)

While the properties of primitive conditional probability *entail* the well-known formula below, in the present setting conditional probability is *defined* in terms of monadic probability:

[2] Observe that there are two-place probability measures which have no cores. To take a simple example, let $W = [0, 1]$, and let λ be the Borel measure on $[0, 1]$. Define a two-place probability measure P by setting $P(\cdot|B) := \lambda(\cdot|B)$ if $\lambda(B) > 0$ and $P(\cdot|B) := 1$ otherwise. Then P has no probability core, for any normal subset K contains a finite set K_0 such that $P(\emptyset|K_0) = 1$.

Definition 3.2. Let P be a probability measure on $\mathscr{A}$, and let $A, B \in \mathscr{A}$. Then the *conditional probability of B given A, $P(B|A)$*, is defined as

$$P(B|A) := \frac{P(A \cap B)}{P(A)},$$

provided $P(A) > 0$, and is undefined otherwise.

Now we are in good position to extend the notion of a probability core as introduced above. The idea is to generalize the notion of a core by generalizing the notion of dominance or superiority to high probability. Recall that we said of two propositions A, B that A dominates B, written $A >_P B$, if $P(B|A \cup B) = 0$. To generalize, given $t \in (0, \frac{1}{2}]$, let us say that A dominates B if $P(B|A \cup B) < t$ when $P(A \cup B) > 0$. Equivalently and officially, given $r \in [\frac{1}{2}, 1)$, let us say that A *r-dominates* B, written $A >_P^r B$, if $P(B|A \cup B) < 1 - r$ when $P(A \cup B) > 0$.

Definition 3.3 (*r-Core*)**.** Let P be a probability measure on $\mathscr{A}$, let $K \in \mathscr{A}$, and $.5 \leq r < 1$. We call K a *high probability core*^{r}, or *r-core*, if it satisfies the *strong r-superiority condition*:

For every $A, B \in \mathscr{A}$,
If $A \subseteq K$ is nonempty and $K \cap B = \emptyset$, then $A >_P^r B$.

This is a natural generalization of the Strong Superiority Condition used in the previous section. Under the supposition that either a consistent proposition entailing K or a proposition entailing $\overline{K}$ is the case, the degree of belief in the proposition entailing that K is false is small. Like a probability core, an r-core favors propositions entailing that it is true to those which proclaim that it false.

Observe that K is a high probability core^r just in case for all $A, B \in \mathscr{A}$ with $\emptyset \neq A \subseteq K$ and $K \cap B = \emptyset$, if $P(A \cup B) > 0$, then $P(A|A \cup B) > r$. Thus, rather than requiring that $P(B|A \cup B) = 0$ and so $P(A|A \cup B) = 1$, it is required that $P(B|A \cup B) < 1 - r$ and so $P(A|A \cup B) > r$. It follows from the definition that every consistent high probability core has a probability exceeding the threshold r. When it is understood that we are operating with a fixed threshold r, we will talk about a high probability core (HPC) rather than a high probability core^r (HPC^r) or r-core.

Given a probability measure P, we will refer to the nested set of consistent cores of probability less than one induced by P as the *high probability core system for P*, and we will denote the high probability core system for P by $\mathscr{C}_{P<1}^r$, dropping subscripts and superscripts when no confusion will arise.

Like probability cores, high probability cores (HPCs) enjoy nice properties. In fact, it is possible to prove a series of observations paralleling those established for probability cores. In particular, high probability cores nest, in the presence of countable additivity there is an innermost core, and so on. In part because they proceed in a similar fashion, we do not furnish all of the proofs of the properties.

We have further historical reasons for providing only a selection of the proofs. In June of 2010, Hannes Leitgeb gave a stimulating talk on probability and acceptance rules during the inaugural workshop of the Center for Formal Epistemology at Carnegie Mellon University. During the workshop, Leitgeb presented a paper showing how to derive belief from degrees of belief from monadic probability using his notion of stability, to be presented momentarily. The theory had some qualitative resemblance to the traditional theory of probability cores, although the central definitions were not shown to be connected to a natural extension of the theory of probability cores presented here. The goal of Leitgeb's talk was to strike a compromise between logical closure and high probability acceptance rules. The theory has not yet been unveiled in publication, although we were fortunate enough to have an unpublished manuscript slightly before the conference [32].

When the workshop ended, we considered whether the theory of cores could be naturally extended to accomplish something similar to what Leitgeb's theory achieved. We accordingly proposed something close to the above definition for dyadic probability as a candidate for such a natural extension, thereupon exploring its formal properties, conducting our rather straightforward investigation from a point of departure different from Leitgeb's to check whether the resulting extension possessed the central properties of cores. At this point we considered the possibility that in the context of monadic probability the two definitions were logically equivalent, and this was easily seen to be the case. Of course, this was welcome news. We now present Leitgeb's notion of *stability* [32, p. 20, Definition 2]), and we show that for monadic probability stable sets are precisely high probability cores.

Definition 3.4 (*Stability*^{r})**.** Let P be a probability measure on a σ-algebra $\mathscr{A}$ over W, and let $S \in \mathscr{A}$. We say that S is *P-Stable*^{r} if for all $B \in \mathscr{A}$ with $B \cap S \neq \emptyset$, if $P(B) > 0$, then $P(S|B) > r$.

On the one hand, as Leitgeb describes it, a P-Stable^r set is characterized by the property that it has *stably* high probabilities under all suppositions consistent with it where probability is well-defined. On the other hand, a high probability core is charactered by the property that the degree of belief in a proposition entailing its falsity is small under the supposition that either the proposition entailing its falsity or a consistent proposition entailing its truth is the case, again where probability is well-defined. As indicated above, these two notions are easily seen to be logically equivalent.

Proposition 3.5. *Let P be a probability measure on $\mathscr{A}$ over W. Then K is a r-core for P if and only if K is P-*Stable^r.

The proof is straightforward, but we include it for the sake of clarity.

Proof

($\Rightarrow$) Suppose that K is P-stable^r. Let $A, B \in \mathscr{A}$ be such that $A \neq \emptyset, A \subseteq K$, and $B \cap K = \emptyset$, and suppose that $P(A \cup B) > 0$. Then $(A \cup B) \cap K = A \neq \emptyset$, whereby since K is P-stable^r, it follows that $P(K|A \cup B) > r$. Then:

$$P(A|A \cup B) = \frac{P(A \cap (A \cup B))}{P(A \cup B)} = \frac{P(A)}{P(A \cup B)} = \frac{P(K \cap (A \cup B))}{P(A \cup B)} = P(K|A \cup B) > r.$$

Hence, K is a high probability core^r.

($\Leftarrow$) Suppose that K is a high probability core^r. Let $A \in \mathscr{A}$ be such that $A \cap K \neq \emptyset$, and suppose that $P(A) > 0$. Then $(A \backslash K) \cap K = \emptyset, A \cap K \subseteq K$ and $P((A \cap K) \cup (A \backslash K)) = P(A) > 0$, whereby since K is a high probability core^r it follows that $P(A \cap K|A) > r$. But

$$P(K|A) = \frac{P(K \cap A)}{P(A)} = \frac{P((A \cap K) \cap A)}{P(A)} = P(A \cap K|A) > r.$$

Hence, K is P-stable^r. $\square$

In light of this and the remarks above, we need not repeat a number of proofs regarding cores. We list the main theorems, asking the reader to consult [32] or try his or her hand at them on his or her own.

Proposition 3.6. *All non-empty subsets of a high probability core K such that $P(K) < 1$ carry positive probability.*

Furthermore, high probability cores are nested.

Proposition 3.7. *The family $\mathscr{C}^r_{P<1}$ is nested. In fact, for all high probability cores K_1, K_2 such that either $P(K_1) < 1$ or $P(K_2) < 1$, either $K_1 \subseteq K_2$ or $K_2 \subseteq K_1$.*

Finally, we turn to the analogue of Theorem 2.6 (Descending Chains). Since this is an elementary though fundamental result in the theory of probability cores and since our proof is a simple generalization of the well-known proof for standard cores, we present the proof here. A different proof of the same fact can be found in [32].

Theorem 3.8. *There is no infinitely descending chain of cores in $\mathscr{C}^r_{P<1}$.*

Proof. For *reductio ad absurdum*, assume that there is an infinitely descending chain of high probability cores which are all subsets of some core $K_0 \in \mathscr{A}$ such that $P(K_0) < 1$:

$$K_0 \supset K_1 \supset K_2 \supset \cdots$$

Consider the sets $A_n := \bigcup_{i=n}^{\infty} K_i \backslash K_{i+1}$ and $B_n := K_0 \backslash K_n$. Then for each n, $A_n \subseteq K_n$ is nonempty, $B_n \cap K_n = \emptyset$, and $P(A_n \cup B_n) > 0$, so $P(A_n| \bigcup_{i=0}^{\infty} K_i \backslash K_{i+1}) = P(A_n|A_n \cup B_n) > r$. Therefore, since $A_n \supseteq A_{n+1}$ for each n, it follows that $\lim_{n\to\infty} P(A_n| \bigcup_{i=0}^{\infty} K_i \backslash K_{i+1}) = P(\bigcap_{n=0}^{\infty} A_n| \bigcup_{i=0}^{\infty} K_i \backslash K_{i+1}) \geq r$. But $\bigcap_{n=0}^{\infty} A_n = \emptyset$, yielding a contradiction. $\square$

Hence, the system $\mathscr{C}^r_{P<1}$ is well-ordered with respect to the subset relation. There is no guarantee that a probability function P has a high probability core with probability less than one, but if one exists, there must be a *least* non-empty core with probability less than one and it carries a probability exceeding r, as do all high probability cores in $\mathscr{C}^r_{P<1}$.

To make things clear, obviously W is a high probability core and indeed so is *every* proposition $A \in \mathscr{A}$ carrying probability one. We know that any core K with $P(K) < 1$ is a subset of a probability one core, but these cores need not be nested. As such, this theory fails to retain proper command of probability 1 cores and to mesh well with the standard theory of probability cores. Among other things, the largest core cannot be used as a non-trivial representation of full belief. But this should not be surprising. We abandoned primitive conditional probability, and this is one of the prices one must pay for this move. Nevertheless, in the next section we will see that it is possible to construct a version of this account for primitive conditional probability, permitting a smoother connection with the standard theory of probability cores.

We now state the culmination of the previous results.

Theorem 3.9. *If $\mathscr{C}^r_{P<1}$ is nonempty, then $\mathscr{C}^r_{P<1}$ has order type at most ω, has a smallest core, and the union of all cores in $\mathscr{C}^r_{P<1}$ is a high probability core, which carries probability 1 if and only if $\mathscr{C}^r_{P<1}$ is countably infinite.*

Leitgeb [32] verifies essentially the same properties, the exercise being routine, so we omit a proof.

Leitgeb [32, p. 34] proposes to restrict attention to probability measures satisfying a condition that he calls the 'Least Certain Set Restriction,' which demands that there is a set K of $\mathscr{A}$ such that $P(K) = 1$ and for every $A \in \mathscr{A}$ with $P(A) = 1$, it follows that $K \subseteq A$. That is, there is a least set of probability one in $\mathscr{A}$. Still the sets of probability one need not be nested. Of course, the least probability one proposition cannot have non-empty subsets of probability zero.

As Leitgeb correctly explains, there are many examples of countably additive measures obeying the aforementioned restriction, including all probability measures on finite algebras $\mathscr{A}$, all probability measures on the power set algebra of a countably infinite set W, and all *regular* countably additive probability measures. In the case of probability measures on the power set of a countably infinite set, the conjunction of every proposition carrying probability 1 is the least set of probability 1. Regular probability measures observe the requirement that only the empty set has probability 0, so for such measures the least set of probability 1 is the set W itself. Leitgeb claims that these examples cover many, if not most, of the typical philosophical toy examples of subjective probabilities [32]. For measures satisfying this restriction, a natural suggestion is to define a high probability core system in the following way:

Definition 3.10. Let P be a probability measure on an algebra $\mathscr{A}$ such that P satisfies the Least Certain Set Restriction. An (*extended*) *high probability core system* for P, $\mathscr{C}_P^r$, is defined by $\mathscr{C}_P^r := \mathscr{C}_{P<1}^r \cup \{K\}$, where K is the least proposition of probability one.

As suggested by the definition, we call both $\mathscr{C}_{P<1}^r$ and $\mathscr{C}_P^r$ high probability core systems when the context is clear. Of course, probability measures satisfying the Least Certain Set Restriction allow for $\mathscr{C}_{P<1}^r$ to be empty, but K, the least proposition of probability 1, is guaranteed to exist. We thereby define ordinary or plain belief in terms of probability as follows.

Definition 3.11 (*Ordinary belief*)**.** Let P be a probability measure on $\mathscr{A}$ satisfying the Least Certain Set Restriction, and let $O \in \mathscr{A}$. We shall say that O is *plainly believed* (or is an *ordinary belief*) if it is a superset of the smallest high probability core in $\mathscr{C}_P^r$.

Hence, all ordinary beliefs have a degree of belief exceeding r.

Leitgeb compelling argues that such a definition is materially adequate because it follows from plausible postulates governing ordinary belief, probability, and the relationships amongst belief and probability. Loosely speaking, these postulates consist of logical norms for belief, the aforementioned axioms of probability, well-known AGM-style axioms, mixed postulates requiring that ordinary belief have high probability and that supposing a proposition with zero probability must result in an absurd state of opinion, and finally an axiom requiring an agent's corpus of beliefs to be in a certain sense maximal. We find Leitgeb's arguments to be headed in the right direction, much in the spirt of the arguments presented by Arló-Costa and Parikh [6,7] and Arló-Costa [1–3] for probability cores. Of course, we take issue with the postulate demanding that suppositions upon zero probability propositions lead to absurdity, but Leitgeb seems to acknowledge that this is a simplifying assumption. Furthermore, it would be desirable to drop the Least Certain Set Restriction, for although a substantial class of probability measures do satisfy this requirement, many probability measures of philosophical interest do not—for example, those used in economics, statistical inference and decision theory, and scientific practice more generally. We therefore advance the theory presented here with cautious enthusiasm, recognizing that the theory thus far developed—and to be more adequately developed in the next section—represents an important stepping stone.

That Leitgeb's theory can be formulated as a generalization of the standard theory of probability cores lends additional credibility to the theory. The theory of probability cores has some pedigree at this point, even being based on ideas that go back to de Finetti's notion of probabilistic *superiority*. In any case, as the reader can see, both theories have identical logical scope. Of course, Leitgeb should be credited for discovering the theory, and we hope that he publishes his paper soon. We wish to put his theory in a different perspective, showing that the central ideas supporting the theory of probability cores can offer a unified account of acceptance transcending the case of probability one.

4. Dyadic probability and high probability cores

The previous section set aside primitive conditional probability as used in the second section to introduce the theory of probability cores. Instead, we focused on monadic probability, the primary purpose being to show that a natural extension of the standard theory of probability cores coincides with Leitgeb's recent proposal. In this section we present the theory of high probability cores for primitive conditional probability. We will see that this move has various advantages. High probability cores for primitive conditional probability are better behaved, affording a smooth connection with the previous work on probability cores. The good news is that since we have already presented many of the main ideas behind high probability cores, the presentation here will be abbreviated.

To be clear, we presuppose the definition of primitive conditional probability of Section 2. Again invoking a variation of de Finetti's notion of superiority, given $r \in [\frac{1}{2}, 1)$, let us say that A *r-dominates* B, written $A >_P^r B$, if $P(B|A \cup B) < 1 - r$.

Definition 4.1 (r-*Core*)**.** Let P be a two-place probability measure on $\mathscr{A}$, let $K \in \mathscr{A}$, and let $.5 \leq r < 1$. We call K a (*two-place*) *high probability core*^{r}, or *r-core*, if it satisfies the *strong r-superiority condition*:

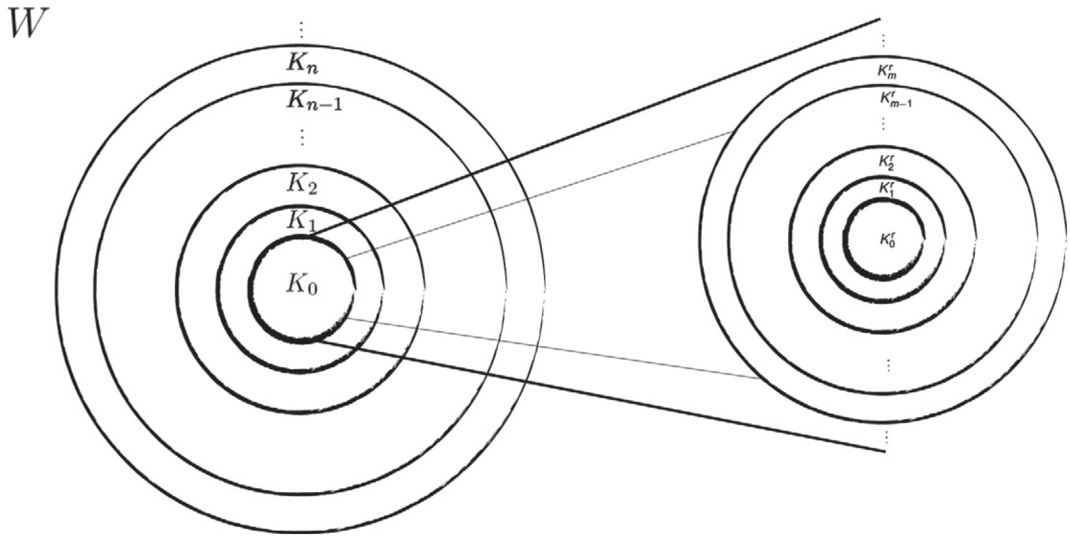

Fig. 2. High Probability Core System.

For every $A, B \in \mathscr{A}$,
If $A \subseteq K$ is nonempty and $K \cap B = \emptyset$, then $A >^r_P B$.

Again, the idea behind probability cores may be expressed in terms of supposition: Under the supposition that either a consistent proposition entailing K or a proposition entailing $\overline{K}$ is the case, the degree of belief in the proposition entailing that K is false is small. An important feature of the above definition is that we have dropped the restriction that $A \cup B$ must have positive probability. As with monadic high probability cores, when it is understood that we are operating with a fixed threshold r, we will talk about a (two-place) high probability core (2-HPC) rather than a (two-place) high probability core^r.

We use the notation from the previous sections, denoting the collection of consistent r-cores for P by $\mathscr{C}^r_P$ and the subcollection of r-cores with probability less than one by $\mathscr{C}^r_{P<1}$. Observe that $\mathscr{C}^r_P$ contains *all* r-cores carrying maximal probability. As usual, we drop the subscript and superscript where no confusion will arise. We now summarize the main properties of high probability cores in the following theorem, the proof of which is straightforward and so omitted.

Theorem 4.2. *Let P be a two-place probability measure on* $\mathscr{A}$.

(i) *All non-empty subsets of any core in* $\mathscr{C}^r_{P<1}$ *carry positive probability;*
(ii) *The system* $\mathscr{C}^r_P$ *is well-ordered with respect to inclusion;*
(iii) *The subsystem* $\mathscr{C}^r_{P<1}$ *has order type at most* ω*;*
(iv) *The union of all cores in* $\mathscr{C}^r_{P<1}$ *is a high probability core, which carries probability* 1 *if and only if* $\mathscr{C}^r_{P<1}$ *is countably infinite.*

The main difference with the previous section is that cores carrying probability one will also be nested. In particular, all probability cores introduced in Section 2 are high probability cores. Moreover, we will have a least such probability one set representing the notion of expectation or 'almost certainty.' And if the probability function has cores with probability less than one, we will also have a least high probability core representing the notion of plain or ordinary belief. So we can accommodate all the attitudes we considered in the introduction of this paper and more. Figure 2 depicts a high probability core system in which each K_m is a probability one core, K_0 is the smallest probability one core, each K^r_m is a member of $\mathscr{C}^r_{P<1}$, and K^r_0 is the least r-core.

Now observe that the largest core—or indeed the union of all cores if a largest does not exist—need not coincide with the universe W. To see this, let us return to Example 1.2—the one about flipping a coin until the agent sees heads. Let $W := \{1, 2, 3, \ldots\} \cup \{\infty\}$, let $\mathscr{A} := \mathscr{P}(W)$, and for each $A \subseteq W$, let $m(A) := \sum_{n \in A \setminus \{\infty\}} \frac{1}{2^n}$. The agent in question may have degrees of belief given by the two-place probability measure P_1 for which $P_1(A|B) = \frac{m(A\cap B)}{m(B)}$, where $m(B) > 0$, $P_1(A|\{\infty\}) = 0$ if $\infty \notin A$, and $P_1(A|B) = 1$ otherwise. In such a case, the agent clearly has two probability 1 cores, $W_0 := \{1, 2, 3, \ldots\}$ and W itself. Yet the agent may instead have degrees of belief P_2 given by $P_2(A|B) = \frac{m(A\cap B)}{m(B)}$, where $m(B) > 0$, and $P_2(A|B) = 1$ otherwise. In this case, the agent has one probability 1 core, W_0. Generally speaking, as in the standard theory of cores, probability measures with abnormal events will determine non-trivial (or non-tautological) full beliefs. The foregoing example also illustrates that the theory is suitably flexible for representing an agent's beliefs.

Let us now consider briefly an appropriate generalization of the notion of P-stability^r. A natural idea is to do what we did for Definition 4.1—simply drop the requirement that B, the conditioning event in the definition of stability, must have positive probability. Thus, we might say that $S \in \mathscr{A}$ is P-Stable^r if for every $B \in \mathscr{A}$ with $B \cap S \neq \emptyset$, $P(S|B) > r$. Yet a consequence of this definition is that high probability cores and stable sets no longer coincide, as can be seen by considering the measure P_2 from above. In fact, a definition for which high probability cores and stable sets coincide takes the following form.

Definition 4.3 (*Stability^r*)**.** Let P be a two-place probability measure on $\mathscr{A}$. We say that $S \in \mathscr{A}$ is *P-Stable^r* if for all $B \in \mathscr{A}$ with $B \cap S \neq \emptyset$, $P(\overline{S}|B) < 1 - r$.

While the basic idea is to drop the requirement that $P(B) > 0$, one must take care of abnormal sets, demanding $P(\overline{S}|B) < 1-r$ rather than just $P(S|B) > r$.

4.1. Core dynamics

In this brief subsection, we point to some properties of the dynamics of core systems. This is the problem that we want to analyze: let P be a two-place probability measure inducing a high probability core system $\mathscr{C}_P^r$. P can then be updated with new information A, so we have:

$$P_A(\cdot|-) = P(\cdot| - \cap A) \qquad (\text{Update})$$

What is the new core system for P_A? For standard probability cores, one of us proved in [3] that:

$$\mathscr{C}_{P_A} = \{K \cap A : K \in \mathscr{C}_P \text{ and } K \cap A \neq \emptyset\}$$

For high probability cores, the situation is slightly different. It is straightforwardly shown that:

$$\{K \cap A : K \in \mathscr{C}_P^r \text{ and } K \cap A \neq \emptyset\} \subseteq \mathscr{C}_{P_A}^r$$

But new cores can also emerge after updating, as illustrated in the following example.

Example 4.4. Let $W = \{\omega_0, \omega_1, \omega_2\}$. Let P be the two-place probability measure on the power set of W such that $P(\{\omega_0\}|W) = \frac{4}{5}$, $P(\{\omega_1\}|W) = \frac{1}{10}$, $P(\{\omega_2\}|W) = \frac{1}{10}$. Then for $A = \{\omega_0, \omega_1\}$ and $r = \frac{17}{20}$, we have $\mathscr{C}_P^{\frac{17}{20}} = \{W\}$, while $\mathscr{C}_{P_A}^{\frac{17}{20}} = \{\{\omega_0\}\}$.

While such a scenario may arise, we find no reason that it should *not*, especially when an agent *learns* a proposition, thereby ruling out certain possibilities. Nevertheless, there are conditions under which the dynamics of high probability cores is identical to the dynamics of standard cores.

4.2. The ratio rule and probability-ratio cores

In this subsection we wish to briefly discuss acceptance rules which differ from those based on high probability, focusing on the so-called *ratio rule* proposed by Isaac Levi [34]. In the following we present background for the ratio rule and its connection to core systems. Readers may skip this subsection and proceed directly to the next section without disrupting the flow of the article.

Let Q and N be unary probability measures. Epistemologically, the probability measure Q is introduced to represent concern for truth in inductive expansion, while the probability measure N is an "information-determining" measure. The idea behind these measures is rather straightforward: At the same that we wish to avoid introducing falsehoods in the current corpus of beliefs, we wish to acquire as much information (measured as propositional content) as possible.

Obviously both concerns might clash and one must strike some kind of compromise between them. One way to represent this compromise mathematically is in terms of the quotient of these measures. Levi introduces a *rejection rule* defined in terms of this quotient and a partition π of events.

Definition 4.5. Let X^* be an element of π carrying maximum value of $\frac{Q(X^*)}{N(X^*)}$. Let q be a real number between 0 and 1. Reject an element X of π if and only if $\frac{Q(X)}{N(X)} < q \cdot \frac{Q(X^*)}{N(X^*)}$.

To simplify things mathematically, we will assume that N is *regular*, i.e., that every consistent event E in the underlying space receives positive measure under N. This ratio rule motivates the following alternative formulation of a probability core that we will call a *probability-ratio core*:

Definition 4.6. Let Q and N be probability measures on $\mathscr{A}$, let $K \in \mathscr{A}$, and $0 < q \leq 1$. Then K is a *probability-ratio core* if for all $A, B \in \mathscr{A}$ with $\emptyset \neq A \subseteq K$ and $K \cap B = \emptyset$,

$$\frac{Q(B)}{N(B)} < q.\frac{Q(A)}{N(A)}$$

Ratio cores enjoy many of the properties of probability cores, giving rise to an interesting class of acceptance rules. It is possible to articulate other acceptance rules which also enjoy interesting static and dynamic properties, and we will engage in a comprehensive discussion of such rules in another article.

5. Coda: the lottery paradox revisited

Can we use the machinery offered above to solve the lottery paradox? We can, and quite simply. Suppose that the lottery has n tickets. Then for each i, the statement that ticket i is not the winning lottery ticket carries high probability. Under a uniform distribution, the *only* core is given by the proposition W itself. Such a result clearly carries over to the transfinite case.

Does this result coincide with intuition? Well, it depends on *whose* intuition we are talking about. The result certainly does not capture the intuitions of stauncher defendants of high probability acceptance rules, and in particular, it does not capture Kyburg's intuitions. Kyburg's central idea is that in the lottery scenario one should be *practically certain* that ticket i is not the winning lottery ticket. Using the most recent terminology that he and Teng have introduced, one has *risky knowledge* of each such proposition ([28]; cf. [29]). Of course, this line of thinking cannot be sustained along with the requirement that risky knowledge be closed under conjunction. By contrast, the solution in terms of high probability cores captures the intuitions of enemies of high probability acceptance rules (e.g., Isaac Levi), the idea articulated in this solution being that in the lottery scenario one only believes that either ticket 1 is the winner, or ticket 2 is the winner, or . . . or ticket n is the winner. The novelty of such a solution is that one arrives at it by way of a probabilistic analysis, while authors like Levi shun such analyses.

In fact, the kind of argument offered here exhibits a compromise between the views of Levi and Kyburg. Although the representation of the epistemic state under consideration is *probabilistic*, the certainties derived from the probability measure in the lottery scenario are those certainties *non-probabilists* think are reasonable in this situation. Some non-probabilists just posit that these certainties are adequate, while others (like Levi) arrive to this conclusion by using non-probabilitic acceptance rules (deploying the notion of epistemic value).

So, it is clear that the model offered here manages to reconcile high probability and logical closure by abandoning some of the basic intuitions that probabilists take for granted in analyses of puzzles such as the lottery paradox. But it is unclear whether probabilists like Kyburg embraced these intuitions just because they are necessary consequences of adopting particular kinds of probabilistic acceptance rules, or whether they pre-systematically enjoyed these intuitions, thereby articulating them in terms of high probability acceptance rules. Of course, we consider the first hypothesis more plausible. The notion of risky knowledge is just the result of embracing a limited form of probabilism intrinsically tied to the use of crude high probability acceptance rules. Once one sees that probabilism can be articulated in alternative ways, such post-systematic intuitions no longer must remain awkward members of the arsenal of probabilism.

6. A decision theory

In this section, we show how qualitative aspects of core systems naturally give rise to a decision theory. Among other things, we will see that the guiding idea behind the primary qualitative features of a core system leads to a lexicographic decision rule which respects compelling principles of rationality. We will also see that the notion of *coherence* due to Finetti [15], when suitably reformulated, provides grounds for observing the principles of two-place probability and indeed finitely additive conditional expectation. To avoid distracting technicalities and unless indicated otherwise, throughout this section we assume that the underlying algebra $\mathscr{A}$ is finite. We begin with our central notion.

Definition 6.1. A *system of cores* over W is a nonempty collection $\mathscr{C}$ of subsets of W well-ordered by $\subseteq$.

We let $\Pi_{\mathscr{C}} := \min_{\subseteq} \mathscr{C}$ and $\Gamma_{\mathscr{C}} := \max_{\subseteq} \mathscr{C}$. According to the intended interpretation, any superset of $\Pi_{\mathscr{C}}$ is *expected* while every superset of $\Gamma_{\mathscr{C}}$ is *fully believed*. The collection $\Gamma_{\mathscr{C}}$ is the space of *serious* or *epistemic possibilities* available to the agent. Thus, any proposition disjoint from $\Gamma_{\mathscr{C}}$ is *epistemically impossible*; let us set $\Lambda_{\mathscr{C}} := W \backslash \Gamma_{\mathscr{C}}$. We drop the superscript $\mathscr{C}$ when the context is clear.

Now given a system of cores $\mathscr{C}$ and a proposition $A \in \mathscr{A}$, define $\mathscr{C}_A$ by setting:

$$\mathscr{C}_A := \{A \cap K : K \in \mathscr{C}\}.$$

Definition 6.2. Let $\mathscr{C}$ be a system of cores over W. Define a map $\pi_{\mathscr{C}} : \mathscr{A} \to \mathscr{A}$ by setting for every $A \in \mathscr{A}$:

$$\pi_{\mathscr{C}}(A) := \min_{\subseteq}(\mathscr{C}_A).$$

We call $\pi_{\mathscr{C}}$ the (*suppositional*) *expectation function* for $\mathscr{C}$.

Again, we drop the subscript $\mathscr{C}$ when there is no danger of confusion. According to the intended interpretation, $E \supseteq \pi(A)$ just in case E is *expected* under the supposition that A. Thus, $\pi(A)$ is *the strongest proposition expected under the supposition that A*. Let us say that E is *materially expected* under the supposition that A if $\pi(A) \subseteq E \subseteq \Gamma \cap A$. Events disjoint from $\Gamma \cap A$ are regarded as *epistemically impossible* under the supposition that A.

Observe that π satisfies the following properties:

(i) $\pi(A) \subseteq A$. (Reflexivity)

(ii) $\pi(A) \subseteq \Gamma$. (Entertainability)

(iii) If $\pi(A) \cap B \neq \emptyset$, then $\pi(A) \cap B = \pi(A \cap B)$. (Arrow)

(iv) If $\Gamma \cap A \neq \emptyset$, then $\pi(A) \neq \emptyset$. (Consistency Preservation)

Property (i) says that A should be expected under the supposition that A, while property (ii) says that regardless of your hypothetical supposition, any full belief ought to be expected. Condition (iii) owes its namesake to Kenneth Arrow, who introduced it in his [8] in the context of rational choice. It captures two plausible properties:

(iiia) $\pi(A) \cap B \subseteq \pi(A \cap B)$. (Conditionalization)

(iiib) If $\pi(A) \cap B \neq \emptyset$, then $\pi(A \cap B) \subseteq \pi(A) \cap B$. (Rational Monotonicity)

According to the intended interpretation, if a proposition B is compatible with the expectations under the supposition that A, then expecting E on the supposition that A and B ought to be the same as deducing E from the meet of B and the expectations obtained under the supposition that A. Finally, property (iv) says that any supposition consistent with your full beliefs—any supposition unqualifiedly epistemically possible—ought to give rise to a consistent body of expectations.

In fact, given any mapping $\pi : \mathscr{A} \to \mathscr{A}$ satisfying conditions (i) to (iv) for some Γ, there is a unique system of cores $\mathscr{C}$ such that $\Gamma = \max_{\subseteq} \mathscr{C}$ and $\pi = \pi_{\mathscr{C}}$. To see this, define a sequence of sets $(\pi_i)_{i<n}$ by setting $\pi_m := \pi(W \backslash (\bigcup_{i<m} \pi_i))$ for each natural number m, and let k be the least natural number such that $\pi_n = \emptyset$. Then $\Gamma = \bigcup_{i<n} \pi_i$, and setting $K_m := \bigcup_{i \leq m} \pi_i$ for each $m < n$, $\mathscr{C} := (K_i)_{i<n}$ is a system of cores with $\Gamma = K_{n-1} = \max_{\subseteq} \mathscr{C}$ and $\pi(A) = \pi_{\mathscr{C}}(A)$ for every $A \in \mathscr{A}$. The simple construction of $\mathscr{C}$ is depicted in Figure 3, while the relationship between π and $\pi_{\mathscr{C}}$ in terms of $(\pi_i)_{i<n}$ is depicted in Figure 4.

In light of this, we may say that a function $\pi : \mathscr{A} \to \mathscr{A}$ is an *expectation function with respect to* Γ. We will also call $(\pi_i)_{i<n}$ *the system of tiers* for $\mathscr{C}$ (or π), given its tier-like status.

Let us now turn to the data of the decision models with which we will concern ourselves.

Definition 6.3

(i) An *act* is a mapping $f : W \to \mathbb{R}$.
(ii) A *constant act* is an act f such that for some $r \in \mathbb{R}$ and every $\omega \in W, f(\omega) = r$.
(iii) Given acts f, g and an event $A \in \mathscr{A}$, we define $f \diamond_A g : W \to \mathbb{R}$ by setting for every $\omega \in W$:

$$(f \diamond_A g)(\omega) := \begin{cases} f(\omega) & \text{if } \omega \in A; \\ g(\omega) & \text{otherwise.} \end{cases}$$

In other words, $f \diamond_A g = f \cdot I_A + g \cdot I_{\overline{A}}$. Thus, $f \diamond_A g$ takes on the value of f on A and the value of g on $\overline{A}$. As usual, a constant function f taking the value r will be abbreviated by $\mathbf{r}$ itself (in bold). We denote the collection of all acts by $\mathfrak{A}$, and the collection of constant acts by $\mathfrak{C}$.

To simplify our exposition, we take real-valued functions $f : W \to \mathbb{R}$ as acts instead of functions on the state space W to a set of consequences, and we assume that all such mappings are measurable. As we will try to make clear in our exposition, much of what we have done in the following can be suitably reformulated in a more general setting.[3]

Given a binary relation $\succsim$ on $\mathfrak{A}$, we define $\succ$ by setting for every $f, g \in \mathfrak{A}$:

$$f \succ g \text{ :iff } f \succsim g \text{ and } g \not\succsim f.$$

We define $\sim$ in the usual way by setting $f \sim g$:iff $f \succsim g$ and $g \succsim f$.

[3] In this context, a mapping $f : W \to \mathbb{R}$ is measurable if $f^{-1}(B) \in \mathscr{A}$ for every Borel set B in $\mathbb{R}$. If $\mathscr{A} = \mathscr{P}(W)$, then every mapping is measurable.

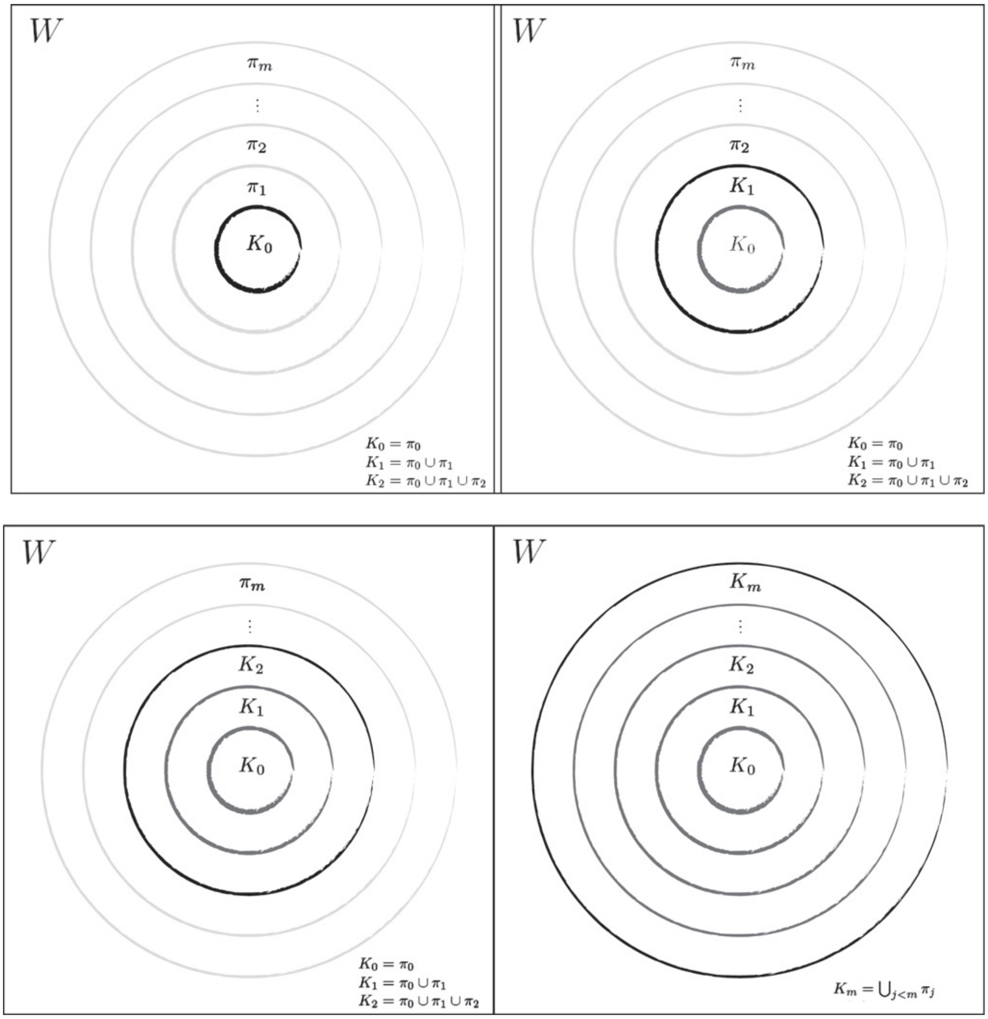

Fig. 3. Core system. Building $(K_i)_{i<n}$ out of $(\pi_i)_{i<n}$.

Recall that the notion of conditional preference is usually defined so that an act f is (weakly) preferred to an act g on the condition that A just in case there is $h \in \mathfrak{A}$ such that $f \diamond_A h \succsim g \diamond_A h$. We modify the definition of conditional preference as follows.

Definition 6.4. Let $\succsim$ be a binary relation on $\mathfrak{A}$, and let π be an expectation function for $\mathscr{C}$. Given an event $A \in \mathscr{A}$, we define a binary relation $\succsim_A$ on $\mathfrak{A}$ by setting for every $f, g \in \mathfrak{A}$:

$$f \succsim_A g \;\text{:iff}\; \text{there is } h \in \mathfrak{A} \text{ such that } f \diamond_{\pi(A)} h \succsim g \diamond_{\pi(A)} h.$$

The intended interpretation of $f \succsim_A g$ is that 'f is (weakly) preferred to g on the supposition that A' or more accurately 'on all expectations under the supposition that A,' the expectations in question belonging to the decision maker. Thus, what is relevant to the decision maker for determining whether or not $f \succsim_A g$ are his comparisons of f and g in light of his body of *expectations* under the supposition that A, ignoring everything inconsistent with these expectations by making f and g the same.

Except for axiom (0) and axiom (4), the following axioms are analogues of those from standard expected utility representations.

Definition 6.5. Let $\pi : \mathscr{A} \to \mathscr{A}$ be a mapping, and let $\succsim$ be a binary relation on $\mathscr{A}$. We call the pair $(\succsim, \pi)$ a *preference order* on $\mathfrak{A}$ if it satisfies the following conditions:

(0) π is an expectation function with respect to Γ;
(1) $\succsim$ is a weak order on $\mathfrak{A}$;

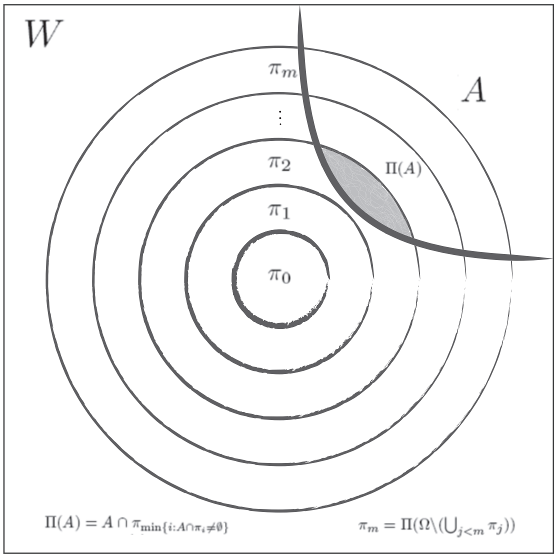

Fig. 4. Expectation function in terms of $(\pi_i)_{i<n}$.

(2) For every $f, g, h \in \mathfrak{A}$:

$$\text{If} \quad f \succsim g, \text{ then } f + h \succsim g + h.$$

(3) For every $f \in \mathfrak{A}$ and $A \in \mathscr{A}$, there is $\mathbf{c} \in \mathfrak{C}$ such that $f \sim_A \mathbf{c}$.
(4) For every $f, g \in \mathscr{A}$ and $A \in \mathscr{A}$ such that $A \subseteq \Lambda, f \cdot I_A \sim g \cdot I_A$.

We call conditions (1)–(4) *Order, Additivity*, *Suppositional Price Equivalence*, and *Absurdity Equivalence*. These axioms capture basic rationality commitments of a decision maker. Additivity corresponds to the axiom of independence, and if we took acts more generally as mappings on the state space to a set of consequences (in particular, horse race lotteries), we would replace additivity with a properly formulated axiom of independence. Suppositional price equivalence corresponds to the property in standard expected utility representations that every act has a certainty equivalent, or to the property of the betting interpretation of probability and prevision (i.e., expected value) according to which the prevision of a gamble is the price at which the decision maker is willing to exchange the gamble (the "fair" price). In our context, however, the name "certainty equivalent" would be inappropriate, for, among other things, $f \sim_W \mathbf{c}$ says that the decision maker is indifferent to the amount c with respect to his *expectations*. Indeed, it would be unobjectionable to call axiom (3) something like *Expected Suppositional Price Equivalence*. As before, in a more general setting where acts map states to consequences, we could replace (3) with an Archimedean axiom which would be assumed to hold under every supposition. Absurdity equivalence requires that an agent be indifferent between acts contingent on an event he regards as impossible. In the context of standard expected utility representations, an event with this property is often called *Savage-null*.

Definition 6.6. Let Γ be the space of serious possibilities. and let $f, g : W \to \mathbb{R}$ be two acts. We say that f *weakly dominates* g with respect to $\Gamma, f \gg_\Gamma g$, if (i) for every $\omega \in \Gamma, f(\omega) \geq g(\omega)$, and (ii) for some $\omega \in \Gamma, f(\omega) > g(\omega)$.

When there is no danger of confusion, we will drop the subscript Γ.

Definition 6.7. We call the pair $(\succsim, \pi)$ a *dominance-sensitive preference order* on $\mathfrak{A}$ if it is a preference ordering satisfying the following condition:

(5) For every $f, g \in \mathfrak{A}$:

$$\text{If} \quad f \gg g, \text{ then } f \succ g.$$

We call (5) the principle of *Weak Dominance*. This is an eminently plausible principle indeed. Assuming that the agent's decision is independent of the state to obtain, when faced with a decision between two acts such that in every state one act is at least good as the other act and in some states better, the agent in question ought to regard the former act as superior to the latter act and regard the latter act as inferior to the former act. The following table illustrates the principle of weak dominance with a state space $W = \{\omega_0, \omega_1, \omega_2\}$ and full beliefs $\Gamma = W$. According to this table, f weakly dominates g and so by the principle of weak dominance f ought to be preferred to g.

	ω_0	ω_1	ω_2
f	\$2	\$1	\$3
g	\$2	\$0	\$2

A standard way to justify the laws of probability and more generally prevision (as de Finetti called it; the more familiar term is "expected value") is by way of a Dutch Book argument. Such an argument, tracing back to Bruno de Finetti [15] and Frank Ramsey [38], standardly presume that an individual's degrees of beliefs can be identified with her fair betting quotients on bets. Such arguments proceed as follows. Acting as bookie, the individual posts her betting quotients for a collection of events subject to the condition that she is willing to accept a gambler's offer to exchange bets on or against the events for sums of money determined by the betting quotients. The individual's degrees of belief are then called *incoherent* if a gambler can make Dutch Book against her with a combination bets, subjecting her to an (almost) sure loss; otherwise, her degrees of belief are *coherent*. The prized mathematical result is that an agent's degrees of belief are coherent just in case they are probabilities. Accordingly, the Dutch Book argument offers prudential grounds for acting in conformity with probabilities.

De Finetti himself considered a more general case in which the individual in question posts her *fair prices* for a collection of *gambles*, or random quantities, subject to the condition that she is willing to buy or sell these gambles for sums of money determined by the prices [17,18]. We also take this more general approach. We may express coherence in this context by requiring that a decision maker should not find himself entangled in a situation in which for some collection of pairs of acts $(f_0, g_0), \ldots, (f_{n-1}, g_{n-1})$ he prefers each f_i to g_i yet when taken together the f_i are never better and sometimes even worse than the g_i taken together.

Definition 6.8. We call the pair $(\succsim, \pi)$ a *coherent preference order* on $\mathfrak{A}$ if it is preference order satisfying the following condition:

(6) There is no collection of acts $(f_i, g_i)_{i<n} \subseteq \mathfrak{A}$ such that for every $i < n$,

$$f_i \succsim g_i,$$

and

$$\sum_{i<n} f_i \ll \sum_{i<n} g_i.$$

We call (6) *Coherence*. As a special case, where $P(A|B)$ is such that $I_A \sim_B P(A|B)$, coherence requires that there is no collection of pairs of events $(A_i, B_i)_{i<n}$ such that for every $i < n$:

$$\lambda_i I_{\pi(A_i)}(I_{B_i} - P(A_i|B_i)) \sim \mathbf{0},$$

yet

$$\sum_{i<n} \lambda_i I_{\pi(A_i)}(I_{B_i} - P(A_i|B_i)) \ll \mathbf{0}.$$

That is, there is no collection of pairs of events such that each called-off bet corresponding to a pair is considered fair yet the sum of the bets pays off no better than the constant gamble with payoff zero and in some states suffers a loss. Hence, coherence resembles de Finetti's related requirement of coherence for called-off bets, i.e., bets rendered void if the conditioning event fails to obtain. The above special case of coherence differs in at least two important respects from de Finetti's notion of coherence for events. First, it invokes the expectation function π, whereas de Finetti's notion does not. De Finetti's called-off bets are a special case of our called-off bets where π is the identity map. Second, our notion invokes weak dominance, whereas de Finetti's notion invokes strict dominance, according to which an act f *strictly dominates* an act g if $f(\omega) > g(\omega)$ for each $\omega \in \Gamma$. Like weak dominance, strict dominance has a corresponding principle which demands that an act f that strictly dominates an act g ought to be preferred to g. Strict dominance is also a plausible principle of rationality and clearly follows from weak dominance. A variant of condition (6) formulated in terms of strict dominance for monadic expectation (and so probability) appears in [19].

Interestingly, in the context of the other axioms the principle of weak dominance and coherence are *equivalent*. In particular, coherence is a *consequence* of weak dominance. Coherence is a commitment a decision maker must undertake if he is also committed to observing weak dominance in the presence of order, additivity, and suppositional price equivalence.

Proposition 6.9. *Let $\mathscr{D} = (\succsim, \pi)$ be a decision model. Then $\mathscr{D}$ is a dominance-sensitive preference order if and only if $\mathscr{D}$ is a coherent preference order.*

Proof. The proof is straightforward and so omitted. Observe that the proof does not utilize axiom (0) or axiom (3). □

Now recall that the idea behind the strong superiority condition is that any non-empty subset of K strongly dominates, or is infinitely more expected than, any proposition disjoint from K. Taking this idea seriously, we can express the idea in terms of preference over acts as follows: any bet paying an arbitrary positive amount on the occurrence of a proposition entailing that K is true ought to be preferred to any gamble paying a unit on the occurrence of a proposition entailing that K is false. More precisely, for every nonempty $A, B \in \mathscr{A}$, $K \in \mathscr{C}$, and $\alpha \in \mathbb{R}$ such that $A \subseteq K$ and $K \cap B = \emptyset$:

$$\text{If } \alpha > 0, \text{ then } \alpha I_A \succ I_B.$$

Thus, no matter how small α may be—indeed even below a unit—a bet paying α on A is strictly preferred to a bet paying a unit on B. This coincides with the idea that everything within K is superior to everything outside of K. The above condition can also be expressed in terms of the expectation function π associated with the system of cores.

Definition 6.10. We call the pair $(\succsim, \pi)$ a *core preference order* if it is a dominance-sensitive preference order satisfying the following condition:

(7) For every $A \in \mathscr{A}$ with $A \cap \Gamma \neq \emptyset$, and $\alpha \in \mathbb{R}$:

$$\text{If } \alpha > 0, \text{ then } \alpha I_{\pi(A)} \succ I_{A \setminus \pi(A)}.$$

In other words, no rate α is small enough to render a simple bet on the strongest proposition expected under the supposition that A dispreferred to *any* simple bet on the weakest proposition inconsistent with $\overline{A}$ and the decision maker's expectations under the supposition that A. We call (7) *Strong Dominance*.

It is easy to verify that in the context of the other axioms, strong dominance is equivalent to the condition preceding it if expressed in terms of the associated system of cores, and strong dominance entails the following condition: For every $f \in \mathfrak{A}$, $A \in \mathscr{A}$ with $A \cap \Gamma \neq \emptyset$, and $\alpha \in \mathbb{R}$:

$$\text{If } \alpha > 0, \text{ then } \alpha I_{\pi(A)} \succ f \cdot I_{A \setminus \pi(A)}.$$

In fact, it is possible to combine weak dominance and strong dominance into a compact and elegant axiom resembling the condition of coherence of de Finetti [15]. To see this, let us introduce a new notion of dominance.

Definition 6.11. Let $\mathscr{C}$ be a system of cores. and let $f, g : W \to \mathbb{R}$ be two acts. We say that f *weakly core dominates* g with respect to $\mathscr{C}$, written $f \ggg_{\mathscr{C}} g$, if there is $K \in \mathscr{C}$ such that $f \cdot I_K \gg g \cdot I_K$.

Clearly weak dominance implies weak core dominance. It is easy to check that the converse does not hold in general. Weak core dominance naturally exploits the ordering of plausibility given by the system of cores $\mathscr{C}$, applying weak dominance successively through the tiers for the core system. Of course, we drop the subscript $\mathscr{C}$ when the context is clear.

Definition 6.12. We call the pair $(\succsim, \pi)$ a *core-coherent preference order* on $\mathfrak{A}$ if it is a preference ordering satisfying the following condition:

(6′) There is no collection of acts $(f_i, g_i)_{i<n} \subseteq \mathfrak{A}$ such that for every $i < n$,

$$f_i \succsim g_i,$$

and

$$\sum_{i<n} f_i \lll \sum_{i<n} g_i.$$

We call (6′) *Core Coherence*. Core coherence presupposes yet another plausible principle of rationality, what we will call the principle of *Weak Core Dominance*, demanding that if $f \ggg_{\mathscr{C}} g$, then $f \succ g$. Indeed, as with Proposition 6.9, it can be shown that weak core dominance is equivalent to core coherence. Let us make this official.

Definition 6.13. We call the pair $(\succsim, \pi)$ a *core dominance-sensitive preference order* on $\mathfrak{A}$ if it is a preference order satisfying the following condition:

(5′) For every $f, g \in \mathfrak{A}$:

$$\text{If } \quad f \ggg g, \text{ then } f \succ g.$$

Of course, we call (5′) *weak core dominance*. We thereby have the following proposition whose proof we omit for the sake of brevity.

Proposition 6.14. *Let $\mathscr{D} = (\succsim, \pi)$ be a decision model. Then the following are equivalent:*

(i) *$\mathscr{D}$ is a core preference order;*
(ii) *$\mathscr{D}$ is a core coherent preference order;*
(iii) *$\mathscr{D}$ is a core dominance-sensitive preference order.*

Now let $\mathscr{D} = (\succsim, \pi)$ be a decision model, and let $(\pi_i)_{i<n}$ be the system of tiers for π. Let us say that suppositional preference is represented by *conditional expected value* if there is a finitely additive conditional expectation $\mathbb{E}[\cdot|\cdot]$ such that for every $f, g \in \mathfrak{A}$ and $A \in \mathscr{A}$:

$$f \succsim_A g \text{ if and only if } \mathbb{E}[f|A] \geq \mathbb{E}[g|A].$$

In addition, let us say that $\mathscr{D}$ is represented by *lexicographic expected value* if for every $f, g \in \mathfrak{A}$:

$$\begin{aligned} f \succsim g \text{ if and only if } & \mathbb{E}[f|\pi_i] = \mathbb{E}[g|\pi_i] \text{ for all } i < n, \text{ or} \\ & \text{there is } m < n \text{ such that for each } i < m, \\ & \mathbb{E}[f|\pi_i] = \mathbb{E}[g|\pi_i] \text{ and } \mathbb{E}[f|\pi_m] > \mathbb{E}[g|\pi_m]. \end{aligned}$$

More compactly, $\succsim$ is represented by lexicographic expected utility if for every $f, g \in \mathfrak{A}$:

$$\begin{aligned} f \succsim g \text{ if and only if } & \text{for every } k < n, \text{ if } \mathbb{E}[g|\pi_i] \geq \mathbb{E}[f|\pi_i] \text{ for each } i < k, \\ & \text{then } \mathbb{E}[f|\pi_k] \geq \mathbb{E}[g|\pi_k]. \end{aligned}$$

Since we were working in a more general framework, let us define precisely what we mean by conditional expected value. In the following definition, let $\mathscr{A}$ be an arbitrary algebra of sets over W.

Definition 6.15. We say that a mapping $\mathbb{E} : \mathfrak{A} \times \mathscr{A} \to \mathbb{R}$ is a *finitely additive conditional expectation* if it satisfies the following properties:

(I) For every $A \in \mathscr{A}$, either:
  (a) $\mathbb{E}[\cdot|A]$ has constant value 1, **or**
  (b) $\mathbb{E}[\cdot|A]$ satisfies the following conditions:
    (b.1) For every $f, g \in \mathfrak{A}$:

$$\mathbb{E}[f + g|A] = \mathbb{E}[f|A] + \mathbb{E}[g|A];$$

    (b.2) For every $f \in \mathfrak{A}$ and $\alpha \in \mathbb{R}$:

$$\mathbb{E}[\alpha \cdot f|A] = \alpha \cdot \mathbb{E}[f|A];$$

    (b.3) For every $f \in \mathfrak{A}$:

$$\inf_{\omega \in A} f(\omega) \leq \mathbb{E}[f|A] \leq \sup_{\omega \in A} f(\omega);$$

(II) For every $f \in \mathfrak{A}$ and $B, C \in \mathscr{A}$:

$$\mathbb{E}[f \cdot I_B|C] = \mathbb{E}[I_B|C] \cdot \mathbb{E}[f|B \cap C].$$

Thus, the conception of a finitely additive conditional expectation is an extension of the concept of a two-place conditional probability, which in turn is a generalization of the standard notion of primitive conditional probability (see, e.g., [20,24]).

Theorem 6.16. *Let $\mathscr{D} = (\succsim, \pi)$ be a decision model. Then the following are equivalent:*

(i) *$\mathscr{D}$ is a core coherent preference order;*
(ii) *There is a finitely additive conditional expectation $\mathbb{E}[\cdot|\cdot]$ such that:*
  (a) *$\pi(A)$ is the support of $\mathbb{E}[\cdot|A]$ for each $A \in \mathscr{A}^\circ$;*
  (b) *$\mathscr{D}$ is represented by conditional expected utility;*
  (c) *$\mathscr{D}$ is represented by lexicographic expected utility.*

Proof

(i) Suppose that $\mathscr{D}$ is a core coherent preference order. First observe that by (1), (2), (3), (4), and (6′), for every $f \in \mathfrak{A}$ and $A \in \mathscr{A}$ such that $A \cap \Gamma \neq \emptyset$, there is an *unique* $\mathbf{c}_{f|A} \in \mathbb{R}$ such that $f \sim_A \mathbf{c}_{f|A}$. Let us therefore define $\mathbb{E} : \mathfrak{A} \times \mathscr{A} \to \mathbb{R}$ by setting for every $f \in \mathfrak{A}$ and $A \in \mathscr{A}$:

$$\mathbb{E}[f|A] := \begin{cases} \mathbf{c}_{f|A} & \text{if } A \cap \Gamma \neq \emptyset \\ 1 & \text{otherwise.} \end{cases}$$

We now show that $\mathbb{E}[\cdot|\cdot]$ is a finitely additive conditional expectation. First we establish that $\mathbb{E}[\cdot|A]$ satisfies properties (I.b) for all $A \in \mathscr{A}$ such that $A \cap \Gamma \neq \emptyset$. We only verify (I.b.2), leaving the proofs for (I.b.1) and (I.b.3) to the reader.

(I.b.2) If α is a natural number, then on the one hand, if $\alpha = 0$, then $\mathbb{E}[0 \cdot f|A] = \mathbb{E}[\mathbf{0}|A] = 0$, and on the other hand, if $\alpha > 0$, then by part (I.b.1) it follows that $\mathbb{E}[\alpha \cdot f|A] = \alpha \cdot \mathbb{E}[f|A]$. If $\alpha > 0$ is a rational number of the form $\frac{m}{n}$, where m, n are natural numbers, then $n \cdot \mathbb{E}[\frac{m}{n} \cdot f|A] = m \cdot \mathbb{E}[f|A]$ and therefore $\mathbb{E}[\frac{m}{n} \cdot f|A] = \frac{m}{n} \cdot \mathbb{E}[f|A]$. Now since $\mathbb{E}[f|A] + \mathbb{E}[-f|A] = \mathbb{E}[f - f|A] = 0$ and so $\mathbb{E}[-f|A] = -\mathbb{E}[f|A]$, it follows that if $\alpha < 0$ is a rational number, then $\mathbb{E}[\alpha \cdot f|A] = \alpha \cdot \mathbb{E}[f|A]$. Finally, let $\alpha \in \mathbb{R}$. First consider $E \in \mathscr{A}$. Then by (6′) it follows that:

$$\begin{aligned} \sup\{\alpha_* \cdot \mathbb{E}[I_E|A] : \alpha_* \in \mathbb{Q} \text{ and } \alpha_* < \alpha\} &\leq \mathbb{E}[\alpha \cdot I_E|A] \\ &\leq \inf\{\alpha^* \cdot \mathbb{E}[I_E|A] : \alpha^* \in \mathbb{Q} \text{ and } \alpha < \alpha^*\}. \end{aligned}$$

Hence, $\mathbb{E}[\alpha \cdot I_E|A] = \alpha \cdot \mathbb{E}[I_E|A]$. It follows straightforwardly from the foregoing that for every $f \in \mathfrak{A}$ and $\alpha \in \mathbb{R}$, $\mathbb{E}[\alpha \cdot f|A] = \alpha \cdot \mathbb{E}[f|A]$.

Clearly by definition $\mathbb{E}[\cdot|A]$ satisfies property (I.a) for all $A \in \mathscr{A}$ such that $A \cap \Gamma = \emptyset$.

(II) The argument here proceeds along familiar lines, but we provide it for the sake of completeness. First let $B, C \in \mathscr{A}$ and $f \in \mathfrak{A}$ be such that $C \cap \Gamma \neq \emptyset$. For *reductio ad absurdum*, assume that $\mathbb{E}[f \cdot I_B|C] \neq \mathbb{E}[I_B|C] \cdot \mathbb{E}[f|B \cap C]$. Define:

$$\lambda_{f \cdot I_B|C} := \begin{cases} 1 & \text{if } \mathbb{E}[f \cdot I_B|C] > \mathbb{E}[I_B|C] \cdot \mathbb{E}[f|B \cap C]; \\ -1 & \text{if } \mathbb{E}[f \cdot I_B|C] < \mathbb{E}[I_B|C] \cdot \mathbb{E}[f|B \cap C]; \\ 0 & \text{otherwise.} \end{cases}$$

Also, define $\lambda_{I_B|C} := -\mathbb{E}[f|B \cap C]\lambda_{f \cdot I_B|C}$ and $\lambda_{f|B\cap C} := -\lambda_{f \cdot I_B|C}$. Now since $\mathbb{E}[\lambda_{f \cdot I_B|C} \cdot f \cdot I_B|C] = \lambda_{f \cdot I_B|C} \cdot \mathbb{E}[f \cdot I_B|C]$ and $\mathbb{E}[\lambda_{I_B|C} \cdot I_B|C] = \lambda_{I_B|C} \cdot \mathbb{E}[I_B|C]$, it follows by (2) and (3) that:

$$\lambda_{f \cdot I_B|C} I_{\pi(C)} f \cdot I_B \sim \lambda_{f \cdot I_B|C} I_{\pi(C)} \mathbb{E}[f \cdot I_B|C]$$
$$-\mathbb{E}[f|B \cap C]\lambda_{f \cdot I_B|C} I_{\pi(C)} I_B \sim -\mathbb{E}[f|B \cap C]\lambda_{f \cdot I_B|C} I_{\pi(C)} \mathbb{E}[I_B|C]$$

If $B \cap \pi(C) = \emptyset$, then the sum of the two gambles on the left is $\lll$ the sum of the gambles on the right. If $B \cap \pi(C) \neq \emptyset$, then by (2) and (3) it follows that $\mathbb{E}[\lambda_{f|B\cap C} \cdot f|B \cap C] = \lambda_{f|B\cap C} \cdot \mathbb{E}[f|B \cap C]$ and:

$$-\lambda_{f \cdot I_B|C} I_{B\cap\pi(C)} f \sim -\lambda_{f \cdot I_B|C} I_{B\cap\pi(C)} \mathbb{E}[f|B \cap C]$$

Again, the sum of the three gambles on the left is $\lll$ the sum of the gambles on the right. If $C \cap \Gamma = \emptyset$, then we immediately have that $\mathbb{E}[f \cdot I_B|C] = \mathbb{E}[I_B|C] \cdot \mathbb{E}[f|B \cap C]$.

In light of the foregoing considerations, it follows that $\mathbb{E}$ is a finitely additive conditional expectation. Moreover, for each $A \in \mathscr{A}$, $\pi(A)$ is the support of $\mathbb{E}[\cdot|A]$ (i.e., the strongest proposition S such that $\mathbb{E}[I_S|A] = 1$), for otherwise (6′) is violated, as may be verified. We have therefore establish parts (ii.a) and (ii.b) of the theorem.[4] We now establish part (ii.c).

(ii.c) Observe that the system of tiers $(\pi_i)_{i<n}$ partitions Γ and is such that for every $m < n$, π_m is the support of $\mathbb{E}[\cdot|W \backslash (\bigcup_{i<m} \pi_i)]$. In the following, let $\Lambda := W \backslash \Gamma$. Now let $f, g \in \mathfrak{A}$. Then $f = (\sum_{i<n} I_{\pi_i} \cdot f) + I_\Lambda \cdot f$ and $g = (\sum_{i<n} I_{\pi_i} \cdot g) + I_\Lambda \cdot g$. Since $\mathbb{E}[f|\pi_i] \cdot I_{\pi_i} \sim f \cdot I_{\pi_i}$ for each $i < n$, we have $f \sim (\sum_{i<n} \mathbb{E}[f|\pi_i] \cdot I_{\pi_i}) + I_\Lambda \cdot f$. Similarly, $g \sim (\sum_{i<n} \mathbb{E}[g|\pi_i] \cdot I_{\pi_i}) + I_\Lambda \cdot g$.

($\Rightarrow$) Now suppose that $f \sim g$. For *reductio ad absurdum*, assume that $\mathbb{E}[f|\pi_i] \neq \mathbb{E}[g|\pi_i]$ for some $i < n$, and let k be the least such i. First consider the case in which $\mathbb{E}[g|\pi_k] > \mathbb{E}[f|\pi_k]$. Then by (7) and the other axioms:

[4] The astute reader will have observed that we have not used anything beyond axioms (1) - (5). In particular, the full force of weak core dominance has not been required thus far.

$$(\mathbb{E}[g|\pi_k] - \mathbb{E}[f|\pi_k])I_{\pi_k} \succ \left[\sum_{j<n-(k+1)} (\mathbb{E}[f|\pi_{k+j+1}] - \mathbb{E}[g|\pi_{k+j+1}])I_{\pi_{k+j+1}}\right] \cdot I_{(\bigcup_{j<n-k}\pi_{k+j})\setminus\pi_k}.$$
$$= \sum_{j<n-(k+1)} (\mathbb{E}[f|\pi_{k+j+1}] - \mathbb{E}[g|\pi_{k+j+1}]) \cdot I_{\pi_{k+j+1}}.$$

It follows that:

$$\sum_{j<n-k} \mathbb{E}[g|\pi_{j+k}] \cdot I_{\pi_{j+k}} \succ \sum_{j<n-k} \mathbb{E}[f|\pi_{j+k}] \cdot I_{\pi_{j+k}}.$$

Therefore, since $\mathbb{E}[f|\pi_i] = \mathbb{E}[g|\pi_i]$ for every $i < k$, it follows that:

$$f \sim (\sum_{i<n} \mathbb{E}[f|\pi_i] \cdot I_{\pi_i}) + I_\Lambda \cdot f \succ (\sum_{i<n} \mathbb{E}[g|\pi_i] \cdot I_{\pi_i}) + I_\Lambda \cdot g \sim g.$$

This contradicts our supposition. Similarly, if $\mathbb{E}[g|\pi_k] < \mathbb{E}[f|\pi_k]$, then another contradiction unravels. Hence, $\mathbb{E}[f|\pi_k] = \mathbb{E}[g|\pi_k]$ and so $\mathbb{E}[f|\pi_i] = \mathbb{E}[g|\pi_i]$ for every $i < n$, as desired.

($\Leftarrow$) Suppose that $f \succ g$. Of course, if $\mathbb{E}[f|\pi_i] = \mathbb{E}[g|\pi_i]$ for each $i < n$, then $f \sim g$. So let k be the least natural number such that $\mathbb{E}[f|\pi_k] \neq \mathbb{E}[g|\pi_k]$. Now argue as in the 'only if' case that if $\mathbb{E}[g|\pi_k] > \mathbb{E}[f|\pi_k]$ then $g \succ f$, yielding a contradiction. $\square$

Among other things, Theorem 6.16 establishes that the qualitative ideas behind a system of cores naturally gives rise to a lexicographic decision rule with respect to a system of cores. Such a rule respects the eminently plausible principle of weak dominance. Shimony [43] has demonstrated that when de Finetti's definition of conditional probability in terms of betting odds is subject to the principle of weak dominance, conditional probability must be *regular* in the sense that all epistemically possible events must receive positive probability (and in fact, for all events A, B, if $P(A|B) = 1$ then $A \subseteq B$). Yet if one offers the modified definition of conditional expectation and so conditional probability as exhibited in part (3) of Definition 6.5 (i.e., Suppositional Price Equivalence), one arrives at a plausible notion of coherence underwritten by the principle of weak dominance, resulting in a notion of conditional expectation and in particular conditional probability admitting epistemically possible events receiving probability zero. While one may regard such a modified operational definition as a departure from the betting framework of de Finetti, it is easily seen that one can reformulate the above definitions and results with respect to a betting interpretation for *vector-valued probability*, an interpretation which one may regard as more faithful and less of a departure. In fact, we regard the vector-valued approach as a promising reorientation of research, offering what appears to be a more general, and fruitful framework (see [22] for a discussion of the relationships amongst lexicographic probability, conditional probability, and nonstandard probability).

7. Conclusion and future work

There are a variety of acceptance rules linking qualitative epistemic notions and degrees of belief. Acceptance rules based on high probability have a special important history in epistemology, and Henry Kyburg focused on such rules to build his well known version of probabilism. Despite intuitions supporting acceptance rules, it is well known that they engender paradoxes. Kyburg was a pioneer in the study of these paradoxes, his version of the lottery paradox clarifying that the price one must pay to adopt high probability rules is the abandonment of classical closure rules such as the rule of Adjunction. Many versions of acceptance rules abound, some modal and some non-modal, but in either case one has the feeling that the use of non-Adjuctive epistemic operators changes the main object of study. Of course, this feeling was not shared by Kyburg. He thought that we ought to bite the bullet and abandon Adjunction, thereby conceptualizing rational qualitative belief in a novel way compatible with probabilism. The lottery paradox also appears as a limit case for full belief, which we believe ought not be conflated with (qualitative) *expectation* or *almost certainty*. Infinitary versions of the lottery make this point clearly.

The seminal work of van Fraassen and the later work of one of us (in collaboration; see [7]) shows that one can indeed distinguish between the strongest proposition that is almost certain and the strongest proposition that is fully believed. The notion of a probability core plays a central role in this argument. Moreover, the notion of probability core can be naturally extended to what we have called a *high probability core*, which is connected with notions employed in other reductions of belief in terms of degrees of full belief.

This would suffice to show the importance of the notion of probability core in distilling paradox-free doxastic notions from probability. But we can go beyond rules based on high probability (or extremely high probability or even probability one). There are other types of probabilistic rules whose underlying ideas are different in nature, like the ratio rule briefly discussed above. The core structure emerges again in the study of such rules, providing a sort of unification of a multiplicity of acceptance rules used in Bayesian epistemology.

In addition, it is possible to show how the qualitative ideas underlying a system of cores naturally gives rise to a decision theory. Among other things, we saw that the guiding idea behind a system of cores leads to a lexicographic decision rule which respects compelling principles of rationality. We also saw that the notion of *coherence* due to de Finetti [15], when suitably reformulated, provides grounds for observing the principles of two-place probability and indeed finitely additive conditional expectation. Thus, core systems can play a crucial role in the articulation of an unified form of Bayesian epistemology for which we not only have probability but also the notion of preference, affording an account of rational action.

Our study in the last section focused on standard probability cores rather high probability cores or even ratio cores. These other notions also naturally give rise to suitable decision theories and decision rules. Precisely formulating these ideas and results is the subject of ongoing work.

Core systems in their various guises have been used extensively in various branches of conditional non-monotonic logic, belief revision, as well as other areas. One natural project is to derive a unified theory of belief and conditional belief which has probabilistic or decision theoretic roots. We have engaged in preliminary work in this direction in a paper published electronically in FEW '10. In addition, probabilistic core systems have interesting connections to work in game theory, dynamic epistemic logic, and other areas (see, for example, [9,12,13,45], and related literature) which require further investigation, representing exciting lines of future research.

Finally, as we have indicated, our understanding of rational quantitative belief leaves room for degrees of belief to be imprecise or indeterminate. Future work shall focus on the notions of belief and conditional belief arising when the underlying probabilities are permitted to be imprecise or indeterminate. In this way, one can retrieve conditional systems weaker than Lewis' system **V**.

Acknowledgments

We are grateful for comments received from Hannes Leitgeb at different stages of the preparation of this paper. We had access to a paper he had been preparing on the reduction of belief to degrees of belief. Later on Leitgeb gave two talks at Carnegie Mellon University organized by the Center for Formal Epistemology, where he presented updated versions of his paper. We benefited from interesting discussions with him on these and other occasions (like FEW 2010 in Konstanz). Teddy Seidenfeld contributed detailed and insightful comments about issues presented in the second part of the paper. Bas van Fraassen read the final version of the manuscript and raised various issues which in part were discussed in the section devoted to the lottery paradox. Of course, his work in this area has been a constant source of inspiration for our own work. We are grateful to Rohit Parikh for insightful conversations and exchanges at different stages of preparation of the paper. We also thank Hailin Liu for very helpful remarks on several drafts of the paper. Finally, we wish to express our gratitude to the referees for valuable comments and suggestions.

References

[1] Horacio Arló-Costa, Qualitative and probabilistic models of full belief, in: P. Pudlak, S. Buss, P. Hajek (Eds.), Proceedings of Logic Colloquim '98, Lecture Notes on Logic, vol. 13, ASL, A.K. Peters, 1999.
[2] Horacio Arló-Costa, Bayesian epistemology and epistemic conditionals: on the status of the export-import laws, Journal of Philosophy 98 (11) (2001) 555–598.
[3] Horacio Arló-Costa, Hypothetical revision and matter-of-fact supposition, Journal of Applied Non-Classical Logics 11 (1–2) (2001) 203–229.
[4] Horacio Arló-Costa, First order extensions of classical systems of modal logic: the role of the barcan schemas, Studia Logica 71 (1) (2002) 87–118.
[5] Horacio Arló-Costa, Non-adjunctive inference and classical modalities, Journal of Philosophical Logic 34 (5–6) (2005) 581–605.
[6] Horacio Arló-Costa, Rohit Parikh, Two-place probabilities, beliefs, and belief revision, in: Dekker (Ed.), Proceedings of the Twelfth Amsterdam Colloquium, Amsterdam: ILLC, 1999, pp. 1–6.
[7] Horacio Arló-Costa, Rohit Parikh, Conditional probability and defeasible inference, Journal of Philosophical Logic 34 (1) (2005) 97–119.
[8] Kenneth J. Arrow, The possibility of a universal welfare function. Technical report, Project RAND, RAD(L)-289, October 1948.
[9] Alexandru Baltag, Sonja Smets, Probabilistic dynamic belief revision, Synthese 165 (2) (2008) 179–202.
[10] David Blackwell, Lester E. Dubins, On the existence and non-existence of proper, regular, conditional distributions, Annals of Probability 3 (5) (1975) 741–752.
[11] David Blackwell, C. Ryll-Nardzewski, Non-existence of everywhere proper conditional distributions, The Annals of Mathematical Statistics 34 (1) (1963) 223–225.
[12] Lawrence Blume, Adam Brandenburger, Eddie Dekel, Lexicographic probabilities and choice under uncertainty, Econometrica 59 (1) (1991) 61–79.
[13] Lawrence Blume, Adam Brandenburger, Eddie Dekel, Lexicographic probabilities and equilibrium refinements, Econometrica 59 (1) (1991) 81–98.
[14] Bruno de Finetti, Les probabilités nulles, Bulletin de Sciences Mathématiques 60 (1936) 275–288.
[15] Bruno de Finetti, Foresight: its logical laws, its subjective sources, in: H.E. Kyburg, H.E. Smokler (Eds.), Studies in Subjective Probability, Wiley, 1937
[16] Bruno de Finetti, On the axiomatization of probability theory, in: Giandomenico Majone (Ed.), Probability, Induction, and Statistics, 1979, John Wiley and Sons Ltd., 1949
[17] Bruno de Finetti, Probability, statistics and induction: their relationship according to the various points of view, in: Isotta Cesari, Leonard J. Savage (Eds.), Probability, Induction, and Statistics, 1979, John Wiley and Sons Ltd., 1959
[18] Bruno de Finetti, 1990 ed., Wiley, 1974.
[19] Enrico Deicidue, Peter Wakker, Dutch books: avoiding strategic and dynamic complications, and a comonotonic extension, Mathematical Social Sciences 43 (2002) 135–149.
[20] Lester E. Dubins, Finitely additive conditional probabilities, conglomerability and disintegrations, The Annals of Probability 3 (1) (1975) 89–99.
[21] Adam Grove, Two modellings for theory change, Journal of Philosophical Logic 17 (2) (1988) 157–170.
[22] Joseph Y. Halpern, Lexicographic probability, conditional probability, and nonstandard probability, Games and Economic Behavior 68 (1) (2010) 155–179.
[23] William Harper, Rational belief change, popper functions and counterfactuals, Synthese 30 (1–2) (1975) 221–262.
[24] Janina Hosiasson-Lindenbaum, On confirmation, The Journal of Symbolic Logic 5 (4) (1940) 133–148.
[25] Harold Jeffreys, Theory of Probability, Oxford University Press, 1939.
[26] A.N. Kolmogorov, Foundations of the Theory of Probability, Chelsea Publishing Company, 1950.
[27] P.H. Krauss, Representation of conditional probability measures on boolean algebras, Acta Mathematica Academiae Scientiae Hungaricae 19 (3–4) (1968) 229–241.

[28] H.E. Kyburg, C.M. Teng, The logic of risky knowledge, in: WoLLIC'2002, 9th Workhop on Logic, Language, Information and Computation, October 2002, pp. 254–262.
[29] H.E. Kyburg, C.M. Teng, The logic of risky knowledge, reprised. In Press, 2011.
[30] Henry E. Kyburg, Probability and the Logic of Rational Belief, Wesleyan University Press, 1961.
[31] Henry E. Kyburg, The rule of adjunction and reasonable inference, The Journal of Philosophy 94 (3) (1997) 109–125.
[32] Hannes Leitgeb, Reducing belief simpliciter to degrees of belief, Manuscript, Bristol, 2010.
[33] Isaac Levi, The Enterprise of Knowledge: An Essay on Knowledge, Credal Probability, and Chance, MIT Press, 1980.
[34] Isaac Levi, For the Sake of the Argument, Cambridge University Press, 1996.
[35] Patrick Maher, Acceptance without belief, in: PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, Volume One: Contributed Papers, vol. 1990, 1990, pp. 381–392.
[36] Richard Montague, Universal grammer, Theoria 36 (3) (1970) 373–398.
[37] K.R. Popper, The Logic of Scientific Discovery, Hutchinson, 1975.
[38] Frank Plumpton Ramsey, Truth and probability, in: D.H. Meller (Ed.), Philosophical Papers, 1990, Cambridge University Press, 1926
[39] A. Rényi, On a new axiomatic theory of probability, Acta Mathematica Academiae Scientiae Hungaricae 6 (1955) 268–335.
[40] A. Rényi, On conditional probability spaces generated by a dimensionally ordered set of measures, Theory of Probability and its Applications 1 (1) (1956) 55–64.
[41] Dana Scott, Advice in modal logic, in: Karel Lambert (Ed.), Philosophical Problems in Logic, Reidel, Dordrecht, 1970
[42] Teddy Seidenfeld, Mark J. Schervish, Joseph B. Kadane, Improper regular distrbutions, The Annals of Probability 29 (4) (2001) 1612–1624.
[43] Abner Shimony, Coherence and the axioms of confirmation, The Journal of Symbolic Logic 20 (1) (1955) 1–28.
[44] Wolfgang Spohn, A general non-probabilistic theory of inductive reasoning, vol. 4in: R.D. Schachter, T.S. Levitt, L.N. Kanal, J.F. Lemmer (Eds.), Uncertainty in Artifical Intelligence, , North Holland Publishing Company, 1990
[45] Robert Stalnaker, Knowledge, belief and counterfactual reasoning in games, 12Economics and Philosophy (1996) 133–163.
[46] Bas van Fraassen, Fine-grained opinion, probability, and the logic of full belief, 24Journal of Philosophical Logic (4) (1995) 349–377.