1-1-2001

Iterative Probability Kinematics

Horacio Arlo-Costa
Carnegie Mellon University, hcosta@andrew.cmu.edu

Richmond H. Thomason

Iterative Probability Kinematics

Horacio Arló-Costa and Richmond H. Thomason

Abstract

ABSTRACT. Following the pioneer work of Bruno De Finetti [12], conditional probability spaces (allowing for conditioning with events of measure zero) have been studied since (at least) the 1950's. Perhaps the most salient axiomatizations are Karl Popper's in [30], and Alfred Renyi's in [32]. Non-standard probability spaces [33] are a well know alternative to this approach. Vann McGee proposed in [29] a result relating both approaches by showing that the standard values of infinitesimal probability functions are representable as Popper functions, and that every Popper function is representable in terms of the standard real values of some infinitesimal measure.

Our main goal in this article is to study the constraints on (qualitative and probabilistic) change imposed by an extended version of McGee's result. We focus on an extension capable of allowing for iterated changes of view. Such extension, we argue, seems to be needed in almost all considered applications. Since most of the available axiomatizations stipulate (definitionally) important constraints on iterated change, we propose a non-question-begging framework, Iterative Probability Systems (IPS) and we show that every Popper function can be regarded as a Bayesian IPS. A generalized version of McGee's result is then proved and several of its consequences considered. In particular we note that our proof requires the imposition of Cumulativity, i.e. the principle that a proposition that is accepted at any stage of an iterative process of acceptance will continue to be accepted at any later stage. The plausibility and range of applicability of Cumulativity is then studied. In particular we appeal to a method for defining belief from conditional probability (first proposed in [41] and then slightly modified in [6] and [3]) in order to characterize the notion of qualitative change induced by Cumulative models of probability kinematics. The resulting cumulative notion is then compared with existing axiomatizations of belief change and probabilistic supposition. We also consider applications in the probabilistic accounts of conditionals [1] and [29].

KEY WORDS: Conditional probability, Hypothetical revision, Infinitesimal probability, Popper functions, Probability conditionals, Supposition.

1.   Introduction

Epistemic states undergo rational change. Therefore, probability measures need to represent epistemic transitions, as well as static epistemic states. The same should hold for the representations of conditional probability known as "Popper functions."[1] The point applies with special force in this case, since one popular motivation for Popper functions is the need for reasonable epistemic transitions in cases more general than those that are allowed when conditional probability is defined in terms of absolute probability.

Vann McGee, for instance, argues that events with probability zero (such as an infinite sequence of coin tosses, all coming up heads) are possible, and goes on to say this:

> If a proposition we now assign the probability 0 is discovered to be true, we shall revise our system of beliefs. In seeing how we ought rationally to revise our beliefs, standard Bayesian theory is no help.
> [29, p. 179]

Clearly, epistemic transitions should iterate. It is not much good to have a characterization of revision that only allows you to revise your opinion once. For some reason, however, Popper functions have not been formulated in a way that provides a direct, general way of of dealing with iterated revisions.

Here is the problem with Popper's idea of using a relational probability $C(p, q)$ as a theoretical primitive. Suppose we take such a probability, C, together with a proposition p_0 (the latter representing everything that an agent unconditionally accepts)[2] to be the representation of an ideal agent's epistemic state. The agent's initial state, then, will have the form

$$(1.1)\ \langle \lambda p C(p, p_0), p_0 \rangle.$$

Suppose now that the agent accepts a new proposition p_1. Then the new epistemic state, according to Popper, would be

$$(1.2)\ \langle \lambda p C(p, p_0 \wedge p_1), p_0 \wedge p_1 \rangle.$$

If the agent in state (1.2) now accepts p_2, the new epistemic state will be

$$(1.3)\ \langle \lambda p C(p, p_0 \wedge p_1 \wedge p_2), p_0 \wedge p_1 \wedge p_2 \rangle.$$

Thus, as propositions are accepted they are combined into a conjunction that compiles the acceptance history. Since this cumulative combining of accepted propositions is built into the definition of revision, it holds universally, regardless of eventual conflicts among the accepted propositions. One consequence of the conjunctive policy of dealing with iterated acceptance is *Simple Cumulativity*: the principle that a proposition that is accepted at any stage of an

---

[1] The reference here is to the new appendices *ii to *v of [30] as well as to the work of Alfred Renyi in [32]. The axioms adopted in this paper, nevertheless, do not exactly coincide with either the ones offered in [30] or [32]. We adopt a axiomatization more akin to the one used later by van Fraassen in [41]. See the comments in Section 1.1, below, comparing the two approaches.

[2] A more precise epistemological analysis of the attitude encoded by p_0 will be offered later on. For the moment the reader can think about p_0 as the encoding of the *certainties* or *full beliefs* of the agent.

iterative process of acceptance will continue to be accepted at any later stage. Cumulativity is a rather severe constraint on learning scenarios. In fact, the principle requires a strong form of path-dependence according to which a sufficient condition for the incoherence of the current epistemic state is an epistemic past where incompatible inputs have been accepted.

It is easy to make this constraint seem implausible. For instance, consider the following example.

Example 1.1.

> I assign probability 0 to a certain die's falling edgewise. Now I am told by an informant whom I trust that
>
> > (1.4) The die fell on an edge.
>
> Then I am told by another informant whom I also trust that
>
> > (1.5) The other informant lied.

Simple Cumulativity can be defended against such examples by constraining the interpretation of C so that revisions with mutually incompatible pieces of information become incoherent. One could argue, for instance, that what should be have been accepted in the first place in the above example is not (1.4) but

> (1.4′) Either the die fell on an edge or my informant is lying.

We have a certain amount of sympathy with this idea, if it is applied cautiously. But even the most enthusiastic advocates should reject formalizations which, like Popper's, make it impossible to even formulate rival interpretations of revision. Even if cumulativity assumptions are reasonable and appropriate in some or even all cases, it is going too far to make them hold as a matter of definition.

In the next two sections of this paper, we develop a formal account of conditional probability that generalizes Popper's formalization, and that does not enforce cumulativity[3] as a matter of definition. We believe that this account provides insights into probability kinematics that are not readily obtainable using Popper's formalism. For instance, it provides a better basis for comparison with qualitative theories of belief revision, and semantic theories of conditionals, and it enables us to formally characterize the cases of probability kinematics that are representable using nonarchimedean probability.

1.1. Popper functions and infinitesimal probability: two sides of the same coin?

There are two standard ways to represent conditioning with events carrying zero measure: infinitesimal probabilities and the so-called Popper functions. The former notion appeals to the idea that a proposition can carry *infinitesimal* value greater than zero, but smaller than any standard number. Instead of assigning zero to the proposition that a randomly chosen point is in the Western Hemisphere, we give it an infinitesimal value. So, if we assign an infinitesimal value i to the point being in the southern part of the Western Hemisphere and

---

[3]In a general setting, different, non-equivalent cumulativity principles become formulable. None of these forms of cumulativity hold automatically.

$2i$ to the proposition that it is in the Western Hemisphere; the proposition that it is in the southern part of the Western Hemisphere, conditional on its being in the Western Hemisphere is defined and carries infinitesimal value $\frac{i}{2i}$. Of course, in order to have a workable proposal, a calculus of infinitesimal magnitudes has to be developed. Abraham Robinson's pioneering work in this area is standard in the field today.

Karl Popper [30] and Alfred Renyi [32] proposed a different way of proceeding. The basic idea is to treat conditional probability as primitive. The two-place functions in question receive standard real values, but the probability of $P(p \mid q)$ where q carries unconditional value zero is well defined (i.e., it is defined when $P(q \mid U) = 0$, where U is the universe of the underlying sigma-field F). The central axiom constraining the two-place functions is the so-called Multiplication Axiom, according to which $P(p \cap r | q) = P(p|q)P(r|p \cap q)$ for all propositions q, p, r

These two approaches are differently motivated. They have grown independently and their foundational problems have different roots.[4]

Perhaps non-standard probability is a more adequate (or 'intuitive') representational tool, but two-place functions have the advantage of a simple axiomatization—at least this is so for most applications where the domain of the underlying sigma field is at most countable.[5]

In spite of the aforementioned differences, Vann McGee has recently offered an interesting result showing that it is possible to map statements about non-standard probabilities into corresponding facts about ordinary valued two-place probability functions (and vice-versa). The gist of the result is a back-and-forth proof showing that the standard values of infinitesimal probability functions are representable as two-place functions, and that every two-place function is representable in terms of the standard real values of some infinitesimal measure. McGee's result is presented only for non-iterated probability change. The first five sections of this piece focus on showing that: (1) McGee's original proof does not extend to an iterative setting; (2) that there is nevertheless a modified proof that does extend to this case, and (3) that this extended proof requires cumulativity assumptions.

As we explained above, we will first present two-place iterative kinematics in such a way that cumulativity is not imposed definitionally. This will help to display more effectively the role played by cumulativity in our representation result. This will also have the additional virtue of developing a generalized account of iterated Popper functions, which allows cumu-

---

[4]Popper's main motivation was the development of a probabilistic semantics for classical logic. Therefore in his model the carriers of probability are *sentences*. Renyi followed Kolmogorov's lead, and therefore in his axiomatixation of generalized conditional probability *propositions* (sets of points in the underlying sigma-field) are the carriers of probability. Here we follow the latter path rather than the former. But we do not adopt Renyi's approach in full either. For example, Renyi relativizes conditioning events to what he calls *bunches*. So, the terminology 'Popper functions', or even 'Popper-Renyi functions', in the context of this paper, will only reflect the common attempt to axiomatize (without appealing to non-standard numbers) a generalized notion of conditional probability, rather than particular historical features of the models developed by either Popper of Renyi.

[5]The work done by Alfred Renyi on conditional probability is perhaps the most serious attempt to provide solid foundations for two-place functions. Still it is well known that the general definition of the probability of events conditional on events, rather than on sigma fields, is problematic in the context of Kolmogorovian measure theory (see [7], [34], [4]). Probabilities conditional on arbitrary events (defined with respect to arbitrary spaces) are considerably less problematic if finite additivity is assumed. Most of the present work applies to finite or at most countable spaces. The imposition of countable additivity will be stated when needed, although it is not presupposed for most of the results.

lativity assumptions, but does not require them. Depending on the nature of the intended interpretation, cumulativity conditions might or might not be required.

If the goal of using two-place functions is to provide a more general modeling of the way we actually update our beliefs when we acquire new evidence, Simple Cumulativity is at best doubtful as a universal constraint. It seems perfectly reasonable that an agent whose current view is consistent could have an epistemic past where he updated his view sequentially with inputs that do not form a consistent set. The situation is different when the intended interpretation is suppositional, but even in this case it is unclear whether *all* forms of supposition are cumulative. Nevertheless, as we explained above, we will show that in order to extend McGee's result a cumulativity condition *needs* to be imposed. This seems to constrain the set of feasible interpretations of the mapped extensions of classical probability.

Extended conditional probability can be constructed in the context of a pluralist epistemology where belief is a primitive alongside probability. Extended conditional probability measures can then be constructed as coherent degree of belief conditional on qualitative suppositions. Hájek and Harper offered in [18] the first attempt to axiomatize the resulting notion of qualitative supposition. Although the authors try to stress some commonalties with axiomatizations of contemporary theories of belief change, the resulting notion is actually inconsistent with all the standard notions of belief change—independently of their intended interpretation. The model offered in [18] has the further limitation of not applying to sequential changes.

The final sections of this piece are devoted to studying the qualitative constraints on supposition imposed by cumulativity. In other words, we will offer an exact account of the qualitative constraints on supposition imposed by the generalized mapping between Popper functions and infinitesimal probability. In order to do so we will need to associate in a natural manner a qualitative notion of belief with each Popper function, and we will need to show how this body of belief changes when its associated Popper function is updated. We will follow a proposal recently offered by Bas van Fraassen (and recently extended and slightly modified in [3] and [6]). This framework permits a very direct analysis revealing how the probabilistic and qualitative approaches (of belief kinematics) connect.

The notion of belief change that thus arises (Hypothetical Revision) is axiomatized and compared with some of the standard notions of belief change in the literature. Since the qualitative counterpart of Simple Cumulativity is rejected by standard accounts, we will see that Hypothetical Revision is irreducible to any of those accounts. This might be surprising, given that historically part of the recent work on (qualitative) belief dynamics was motivated by previous work on probability kinematics. Contrary to what untutored intuition might suggest, the epistemology induced by (the standard presentation of) Popper functions[6] is at odds with some of the most popular qualitative accounts of inquiry and supposition developed to date.

The last section of the paper will be devoted to showing that Hájek and Harper's approach can be reconciled with our approach. We will show that a *modified* version of Hájek and Harper's axioms are theorems of Hypothetical Revision. We will also suggest that this notion of hypothetical change manages to articulate the notion of supposition tacitly used in the probabilistic semantics of conditionals developed by Adams and McGee.

---

[6]I.e. a presentation imposing Cumulativity definitionally.

2. An abstract characterization of iterative Popper functions

A semantics for absolute probability is provided by measure theory. But, in itself, a measure-theoretic approach will not provide an interpretation of primitive *conditional* probability. The approach we adopt here retains the classical idea of a probability function, as a function from an algebra of propositions to $[0, 1]$. To this we add a selectional approach to the conditionalization of probabilities, somewhat like Stalnaker's approach to the interpretation of conditionals in [38]. In other words, we postulate a selection function that, given a probability function P and a hypothesis q produces another probability function P. We can then ask about the probability of p on this derived probability function. From this standpoint, a notation like $[U(q, P)](p)$ is preferable to the more familiar notation $P(p, q)$ that derives from Popper [30].[7] This notation has the advantage of providing a neutral representation for iterative updates. The probability of p given an initial probability function P and successive hypotheses $q_0, q_1, \ldots, q_n$ (assumed in this order) is then

$$[U(q_n, U(q_{n-1}, U(\ldots, U(q_0, P))\ldots))](p).$$

$U_{q_0,\ldots,q_n}(P)(p)$ is an alternative, more readable notation for the same thing.

The idea of perturbing a probability function with a hypothesis is similar to van Fraassen's proposals in [39, 40, 41]. In fact, there is an interesting convergence between van Fraassen's ideas and the results presented in this essay, although this coincidence is only clearly revealed by focusing on certain models of a slightly modified version of van Fraassen's original proposal (see [3], [6], and Section 6, below, for details).

In view of cases like Example 1.1, it may be surprising that Simple Cumulativity should be an appropriate constraint to impose on iterative probability kinematics. Tracking the epistemological states of an ideal observer as evidence is gathered is a primary motivation for probability kinematics. (If this is not the entire point of Bayesian epistemology it is a very large part of the point.) But, as Example 1.1 shows, it is not always unreasonable to revise in ways that do not satisfy Simple Cumulativity.[8] We firmly believe that the phenomenon of noncumulativity needs to be taken into account when iterated updates are interpreted in terms of idealized inquiry.

However, there is another interpretation of probabilistic conditionalization, in terms of *supposition*. Here, the ideal agent doesn't undergo a learning experience, but performs a thought experiment involving a hypothetical news item, or supposition.

It would be strange if there were not close connections between hypothetical conditionalization and ideal inquiry conditionalization; in fact, an important use of the former would be to precompile or plan for potential epistemic changes.[9] Nevertheless, most of the researchers who followed Adams' steps in developing a probabilistic semantics of conditionals have adopted Simple Cumulativity in order to explain away the puzzling evidence offered by McGee in [28]. (See [1], [13], [28] for recent literature where Simple Cumulativity is

---

[7]Here, "U" stands for "Update".

[8]One reasonable revision for the observer to make on being told (1.5) is to retract the previous hypothesis (1.4). In this connection, see [15, Chapter 5], where the goal seems to be to develop a theory of noncumulativity in Bayesian inquiry.

[9]Bratman argues for the need for such planning in [9]. Bratman's examples have to do with physical activities, but the same points apply to epistemic actions and plans.

explicitly adopted).[10] A self-contained version of McGee's examples is presented below in Sections 5.2, 6, 7, and 8. Section 3, below, we will give further evidence in support of the view that the notion of supposition embedded in the probabilistic semantics of conditionals *is* cumulative.[11]

In the neutral framework for iterative probabilistic update developed here, Simple Cumulativity has to be added explicitly. The natural way, it seems, to formalize this principle would be as follows:

> *Simple Cumulativity for Probability Update:*
> If $p \in \{q_0, \ldots, q_n\}$, then $U_{q_0,\ldots,q_n}(P)(p) = 1$.

But unless we modify our account of conditional probability, Simple Cumulativity will conflict with the generally accepted constraint that primitive conditional probability should be a conservative extension of Bayesian conditional probability:

> *Favor Bayesian Update:*
> $U_p = P/p$ whenever $P(p) > 0$, where $P/p = \lambda q \, \frac{P(p \wedge q)}{P(p)}$.

In fact, suppose that we have at least one nontrivial Popper revision; that is, assume there is a probability function P and proposition p such that $P(p) = 1$ and $U_{\neg p}(P)(p)$ is defined, so that $U_{\neg p}(P)(p) = 0$. Argue as follows.[12]

> By *Favor Bayesian Update*, $U_p(P) = P$.
> So $U_{p,\neg p}(P) = U_{\neg p}(P)$.
> But by *Simple Cumulativity for Probability Update*, $U_{p,\neg p}(p) = 1$.
> Whereas, by assumption, $U_{\neg p}(P)(p) = 0$.

The upshot is that update can't operate just on a probability function and satisfy both *Simple Cumulativity for Probability Update* and *Favor Bayesian Update*. Therefore, we will think of update as having three inputs: (1) a probability function, (2) the update proposition, and (3) a proposition indicating "necessary information" that is to be assigned probability 1 at all costs for purposes of this update.[13]

With the above intuitions in hand, we now develop an abstract account of iterative conditional probability.

---

[10]McGee's examples are discussed at length in [2].

[11]To establish this claim we will show that extending a representation result presented by McGee in [29] that links infinitesimals and Popper functions requires several forms of cumulativity. Simple Cumulativity was independently assumed (as an intuitive constraint) by McGee in [28]. See Theorems 3.2 and 3.3, below. Cumulativity seems to be an unavoidable feature of this probabilistic semantics rather an optional part of the program.

[12]For those who examine this argument carefully, U is operating on extended probability functions; see below, immediately after Definition 2.1.

[13]This idea is somewhat similar to Vann McGee's idea of for modifying Stalnaker's conditional semantics in [28, p. 469].

Definition 2.1. Probability measure.

Let $\mathcal{B}$ be a Boolean algebra with zero $\bot$ and unit $\top$. A function P from $\mathcal{B}$ to $[0,1]$ is a *probability measure* on $\mathcal{B}$ if and only if:

(1) $P(\bot) = 0$,

(2) $P(\top) = 1$, and

(3) if $p \wedge q = \bot$ then $P(p \vee q) = P(p) + P(q)$.

We adopt the policy in [29] of replacing cases where update is undefined with cases where it yields "the absurd probability function." In particular, let $1_{\mathcal{B}}$ be the function that returns 1 for each $p \in \mathcal{B}$. An extended probability function, or "eprobability function" on $\mathcal{B}$, is either a probability measure on $\mathcal{B}$ or is $1_{\mathcal{B}}$.

Definition 2.2. Iterative Index.

Let $\mathcal{B}$ be a Boolean algebra. An *iterative index i* on $\mathcal{B}$ is a pair $\langle P_i, \square_i \rangle$, where (1) P_i is an eprobability function on $\mathcal{B}$ and (2) $\square_i \in \mathcal{B}$.

Definition 2.3. Partial Iterative System.

A *partial iterative system* $\mathcal{P}$ is a system $\langle \mathcal{I}, \mathcal{B}, /, \omega \rangle$ such that:

(1) $\mathcal{B}$ is a Boolean algebra.

(2) $\mathcal{I}$ is a family of iterative indices over $\mathcal{B}$ and $\omega \in \mathcal{I}$, where $P_\omega = 1_{\mathcal{B}}$ and $\square_\omega = \bot$.

(3) $/$ is a partial function taking indices and propositions to indices, i.e., $i/p \in \mathcal{I}$ if $i/p \in \mathcal{I}$ is defined.
Where $p \in \mathcal{B}$, $i \in \mathcal{I}$, we write i/p for the value that $/$ assigns to i and p. And we write $i/p_0, p_1, \ldots, p_n$ for $(\ldots((i/p_n)/\ldots/p_1)/p_0)$.

A normal index is any index other than ω.

Definition 2.4. Iterative System.

An iterative system is a partial iterative system $\langle \mathcal{I}, \mathcal{B}, /, \omega \rangle$ such that $/$ is total.

The conditions characterizing an iterative system are very weak: they merely place typing constraints on the system's components. We now add three general conditions that are motivated by the interpretation of the revision operator as a form of probability conditionalization.

Definition 2.5. Iterative Probability System.

A [partial] iterative probability system $\mathcal{P}$ is a [partial] iterative system $\langle \mathcal{I}, \mathcal{B}, /, \omega \rangle$ such that, for all p and $i \in \mathcal{I}$:

(4) $P_{i/p}(p) = 1$.

(5) $P_i(\square_i) = 1$ and $P_{i/p}(\square_i) = 1$.

(6) If $\square_i \leq p$ then $i/p = i$.

(7) If $\square_i \leq \neg p$ then $i/p = \omega$.

7

Conditions (4) and (5) ensure that conditionalization on the proposition p at the index i will give probability 1 to p and to any proposition that is necessary at i. Condition (6) is motivated by the idea that the effects of conditionalizing on p should in some sense be minimal; it says that conditioning i on a proposition that is already necessary at i will leave i unperturbed. Condition (7) ensures that the abnormal index is selected on updating with a necessarily false proposition.

We wish to mention three additional conditions which we believe are well motivated by intuitions concerning iterative probability. They characterize a natural class of cases in which the abnormal index is unique, and can only be reached by conditioning on epistemically impossible propositions.[14]

Definition 2.6. Conservative Iterative Probability System.
A conservative [partial] iterative probability system $\mathcal{P}$ is a [partial] iterative probability system $\langle \mathcal{I}, \mathcal{B}, /, \omega \rangle$ such that

(8) If $i/p = \omega$ then $\square_i \leq \neg p$.

(9) If $P_i = 1_{\mathcal{B}}$ then $i = \omega$.

(10) $\square_i = \bot$ iff $i = \omega$.

Several less general conditions are of special importance. We state these separately.

Definition 2.7. A-Cumulative Probability System.
A [partial] iterative probability system $\mathcal{P}$ is *a-cumulative* (acceptance-cumulative) if it satisfies the following condition.

(11) For all $i \in \mathcal{I}$, $\square_{i/p} \leq \square_i$.

Any iterative probability system will assign hypotheses probability 1 on conditionalization; this is the import of Condition (4) of Definition 2.5. An a-cumulative system does this one better, treating hypotheses as necessary.

We note in passing that there are iterated probability systems that are not a-cumulative. In the simplest case, the algebra of propositions is generated by just one proposition p, which alternates endlessly between probabilities 1 and 0 as you revise with p and $\neg p$.

Definition 2.8. Conjunctive Cumulativity.
A [partial] iterative probability system $\mathcal{P}$ satisfies *Conjunctive Cumulativity* if and only if for all $i \in \mathcal{I}$, $i/p, q = i/p \wedge q$.

We will use the following lemma below.

Lemma 2.1. Let $\mathcal{P}$ be a conservative iterative probability system. If $i/p = i/q$, then for all propositions s, $P_{i/r}(p \wedge s) = P_{i/r}(q \wedge s)$.

> *Proof.* If $i/p = i/q$ then $i/p, \neg p = i/q, \neg p$. But $i/p, \neg p = i/p \wedge \neg p$ by Conjunctive Cumulativity, so $i/q \wedge \neg p = \omega$ by Definition 2.5(7) and Conjunctive Cumulativity. So by Definition 2.6(8), $\square_i \leq \neg(q \wedge \neg p)$. By similar reasoning, $\square_i \leq (p \equiv q)$, so $\square_i \leq ((p \wedge s) \equiv (q \wedge s))$. So by Definition 2.5(5), $P_{i/r}(p \wedge s) = P_{i/r}(q \wedge s)$.

---

[14]Conservative IPSs will not be required for various important results stated below. Nevertheless, the proof of Theorem 3.4 does require the use of conservative properties of IPSs.

Definition 2.9. Conjunctive A-Cumulativity.

This is the following condition on [partial] iterative probability systems.

(12) For all $i \in \mathcal{I}$, $\Box_{i/p} = p \wedge \Box_i$.

Definition 2.10. Bayesian Iterative Probability System.

A [partial] iterative probability system $\mathcal{P}$ is *Bayesian* if it satisfies the following condition.

(13) (i) $P_{i/p} = \lambda q P_i(q \mid p)$ if $P_i(p) \neq 0$ and (ii) $\Box_{i/p} = p \wedge \Box_i$.

A Bayesian system uses Bayesian conditionalization for hypotheses that have nonzero probability. In the context of iterative probability systems, we construe Bayesian conditionalization to include not only the familiar transformation of probabilities, but a conservative approach to supposition that transforms the background necessities by adding only the hypothesis proposition to them; i.e., a Bayesian system satisfies Conjunctive A-Cumulativity.

Note that Condition (6) follows from Condition (13) and the remaining conditions of Definition 2.5.

Definition 2.11. Representing a Popper function as an iterative probability system.

We can assume without loss of generality that a Popper function C is defined over a Boolean algebra $\mathcal{B}$ of propositions. (If the Popper function is defined over a language, there is an equivalent Popper function defined over the language's Lindenbaum algebra.) For a Popper function C defined over a Boolean algebra $\mathcal{B}$, the corresponding iterative probability system $\mathcal{P}^C = \langle \mathcal{I}^C, \mathcal{B}^C, /^C, \omega^C \rangle$ is defined as follows:

(i) $\mathcal{I}^C = \{\langle P, p \rangle \ / \ p \in \mathcal{B} \text{ and } P = \lambda q C(q, p)\}$.

(ii) $\mathcal{B}^C = \mathcal{B}$.

(iii) $i/^C p = \langle \lambda q C(q, p \wedge \Box_i), p \wedge \Box_i \rangle$.

(iv) $\omega^C = \langle 1_\mathcal{B}, \bot \rangle$.

Theorem 2.1. Every Popper function C on a complete[15] Boolean algebra $\mathcal{B}$ can be regarded as a Bayesian iterative probability system $\mathcal{P}$ on $\mathcal{B}$.

Proof. Let C be a Popper function. Then the corresponding iterative system $\mathcal{P}^C$ satisfies conditions (4), (5), (6), (7) and (13).

Definition 2.12. Iterative Probability σ-System.

An iterative probability σ-system is like an iterative probability system, but with a σ-algebra and a σ-additivity constraint on the probability functions.

Without going into details, we remark that the case in which the range of P is restricted to $\{0, 1\}$ provides an interpretation for a conditional logic.

---

[15]The completeness condition can, of course, be weakened. We assume completeness here for the sake of simplicity.

Definition 2.13. Generators for $\mathcal{P}$.

Let $\mathcal{P}$ be an iterative probability system. $\mathcal{P}$ is *generated by* $\mathcal{I}_0 \subseteq \mathcal{I}$ if for all i: $i \in \mathcal{I}$ iff for some $i_0 \in \mathcal{I}_0$, either $i = i_0$ or $i = i_0/q_1, \ldots, q_m$ for some $q_1, \ldots, q_m$. $\mathcal{P}$ is *rooted in* i_0 iff $\mathcal{P}$ is generated by $\{i_0\}$.

Lemma 2.2.   Let $\mathcal{P} =$ be an iterative probability system rooted in i_0, that satisfies Conjunctive Cumulativity. Then $i \in \mathcal{I}$ iff $i = i_0/q$ for some $i_0 \in \mathcal{I}$ and q.

Proof. By Conjunctive Cumulativity, $i = i_0/q_1, \ldots, q_m$ iff $i = i_0/q_1 \wedge \ldots \wedge q_m$.

3.   Nonarchimedean representations

3.1.   McGee's representation theorem

In [29], a theorem is proved that is intended to show the equivalence of two quite different approaches to probability kinematics, in the case where nontrivial updates on impossible information are allowed. Using an operation NS taking an arbitrary Popper function C into a nonstandard probability measure $NS(C)$, McGee's representation theorem goes as follows.

Theorem 3.1. [29, pp. 182–183].

(1) If C is a Popper function, then: $NS(C)$ satisfies

   (1.1) $[NS(C)](p) = 0$ iff $C(q, p) = 1$ for all q, and
   (1.2) for all p, q, $C(q, p) = std(NS(C)(q \mid p))$ if $[NS(C)](p) \neq 0$.

(2) If P is a nonstandard probability measure, then there is a Popper function C such that
$$C(q, p) = std(P(q \mid p)).$$

Here, $std(x)$ is the unique standard real that differs from x by at most an infinitesimal amount.

3.2.   Generalizing nonarchimedean representations to an iterative setting

This theorem does in fact establish a natural correspondence between Popper functions and nonstandard probability measures. But as stated and proved, it presupposes fairly strong constraints on update; and worse, these constraints are built into the update notation.

To see this, suppose that we have an iterative update operator $U_p(P, q)$ that produces the result of updating an absolute probability measure P with p, relative to background suppositions compiled into the proposition q. Can we account for this sort of iterative update using a Popper function C? The identity

$$U(\lambda r\, C(r, q), p) = \lambda r\, C(r, q \wedge p)$$

provides an account of single revisions in terms of C. (Note that the revisions are only defined over one-place probability functions that are recoverable by conditionalizing C.) But this update produces only a probability function; it does not provide the background assumptions that are required for further update by U. To iterate the revision process, we must make further assumptions.

The simplest such assumption is the following cumulativity condition.

> *Conjunctive Cumulativity for Iterative Update of a Probability Measure:*
>
> (CCIP) $U_{p_1,\ldots,p_n}(\lambda r\ C(r,\top),q) = \lambda r\ C(r, q \wedge p_1 \wedge \ldots \wedge p_n)$.

If we take McGee's representation theorem for Popper functions to be canonical, then in view of Theorem 3.3, below, we are pretty much committed to CCIP. Thus, the combination of Popper functions, nonstandard probability interpretations, and a view of iterative update that satisfies CCIP presents a coherent picture of probability kinematics, though it does make rather strong assumptions about iterative update, which (at least on some interpretations) can seem implausible. However, to establish the coherence of the position, it is necessary to generalize McGee's representation theorem, since his proof does not work for the iterative case. To see the difficulty, we restate the nonarchimedian representation in a general setting, using iterative probability systems.

Definition 3.1. M-Representation.

Let NS be a map from the normal indices of an iterative probability system $\mathcal{P}$ to a family of nonstandard probability measures. Where P is such a measure and $P(p) \neq 0$, let P/p be the result of conditionalizing P on p:

$$P/p = \lambda q \frac{P(p \wedge q)}{P(p)}.$$

NS *M-represents* $\mathcal{P}$ in case NS satisfies the following two conditions. (These are straightforward analogues of McGee's conditions in the generalized setting.)

> If $i \in \mathcal{I} - \{\omega\}$, then:
>
> (1) for all p, q, $P_{i/p}(q) = std([NS(i)]\,(q|p))$ if $[NS(i)](p) \neq 0$.
> (2) for all p, $[NS(i)](p) = 0$ iff $\square_i \leq \neg p$.

The following definition provides a more useful version of Condition (2), which is available in case we have a complete Boolean algebra, and which implies Condition (2) for finite Boolean algebras.

Definition 3.2. MC-representation.

Let NS be a map from the normal indices of an iterative probability system $\mathcal{P} = \langle \mathcal{I}, B, \omega \rangle$, where B is a complete Boolean algebra, to a family of nonstandard probability measures. NS *MC-represents* $\mathcal{P}$ in case NS satisfies Condition (1) of Definition 3.1, as well as the following condition.

If $i \in \mathcal{I} - \{\omega\}$, then:

$(2')$ $\square_i = \bigwedge \square(NS(i))$, where $\square(P) = \{p \ / \ P(p) = 1\}$.

Definition 3.3. Revision Homomorphic M-representation.

Let NS be a map from the indices of an iterative probability system $\mathcal{P}$ to nonstandard probability functions. NS is *Revision Homomorphic* in case

$$NS(i/p) \doteq \lambda q [NS(i)](q \mid p).$$

Here, $\doteq$ stands for strong equality on partial functions: $F(x) \doteq G(x)$ means that $F(x)$ is defined iff $G(x)$ is defined, and whenever $F(x)$ is defined, $F(x) = G(x)$. A revision homomorphic M-representation, then, renders the representation of an update in the iterative probability system equal to the update of its representation.

Theorem 2.1 showed that every Popper function corresponds to a Bayesian iterative probability system. And these systems satisfy a strong form of cumulativity, Conjunctive A-Cumulativity. The following two theorems establish connections between nonarchimenidean representability and forms of cumulativity.

Theorem 3.2. Let $\mathcal{P}$ be an iterated probability system that is M-represented by a revision homomorphic mapping NS. Then $\mathcal{P}$ is a-cumulative.

Proof. It follows from condition (2) of Definition (3.1) that $[NS(i)](q) = 1$ iff $\square_i \leq q$, for all i and q. In particular, $[NS(i/p)](\square_i) = 1$ iff $\square_{i/p} \leq \square_i$.

(i) Suppose that $\square_i \not\leq \neg p$. Then by Revision Homomorphism, $NS(i/p)(\square_i) = \dfrac{NS(i)(\square_i \wedge p)}{NS(i)(p)}$. Again from Condition 2 of Definition (3.1), $[NS(i)](\square_i) = 1$ iff $\square_i \leq \square_i$, so we have $[NS(i)](\square_i) = 1$. Therefore, $NS(i/p)(\square_i) = \dfrac{NS(i)(p)}{NS(i)(p)} = 1$. So $\square_{i/p} \leq \square_i$.

(ii) Suppose that $\square_i \leq \neg p$. Then by Condition (7) of Definition 2.5, $i/p = \omega$, so $\square_{i/p} = \bot$. So, trivially, we have $\square_{i/p} \leq \square_i$ in this case.

Theorem 3.3. Suppose that NS provides a Revision Homomorphic MC-representation of an iterative probability system $\mathcal{P}$ over a complete Boolean algebra. Then $\mathcal{P}$ satisfies Conjunctive Cumulativity.

Proof.

Case 1. $NS(i)(p \wedge q) > 0$.

12

Then $NS(i/p \wedge q)$ is defined, since $NS(i/p \wedge q)(r) \doteq [NS(i)](r \mid p \wedge q)$, which is defined by hypothesis. Also, $NS(i/p, q)$ is defined, since $NS(i/p, q)(r) \doteq [NS(i/p)](r \mid q) \doteq NS(i)(r \mid p \wedge q)$. Furthermore, this calculation shows that $NS(i/p \wedge q) = NS(i/p, q)$.

By Condition $(2')$ of Definition 3.2, we have $\square_{i/p,q} = \bigwedge \square(NS(i/p, q))$.
But $\bigwedge \square(NS(i/p, q)) = \bigwedge \square(NS(i/p \wedge q))$. Again by Condition $(2')$ of Definition 3.2, we have $\square_{i/p \wedge q} = \bigwedge \square(NS(i/p \wedge q))$.
So $\square_{i/p,q} = \square_{i/p \wedge q}$.

Case 2. $NS(i)(p \wedge q) = 0$.

By Condition $(2')$ of Definition 3.2, we have $\square_i \leq \neg(p \wedge q)$.
Then, by Condition (7) of Definition 2.5, $i/p \wedge q = \omega$.
By the argument of Case 1, we have $NS(i/p \wedge q) \doteq NS(i/p, q)$. So, since $NS(i)/p \wedge q$ is undefined, $NS(i/p, q)$ is also undefined. But NS is defined on all indices other than ω; therefore $i/p, q = \omega$. So again, $i/p \wedge q = i/p, q$.

Case 3. $NS(i)$ is undefined. Then $i = \omega$, so $i/p \wedge q = i/p, q = \omega$.

Under very general conditions, then, the only iterative systems that are representable using nonstandard probability measures will satisfy Conjunctive Cumulativity. This means, among other things, that our happiness with the nonstandard representations should depend crucially on our willingness to accept Conjunctive Cumulativity as a constraint on probability kinematics.

3.3. Problems with McGee's proof

McGee's proof does not apply to the iterated case; his construction does not guarantee the Revision Homomorphism condition. This is shown by the following example.

Example 3.1.

Take the propositional language generated over the atoms $\{p, q\}$. Define the behavior of C as follows.

1. $C(p, \top) = .5$
2. $C(q, \top) = 0$
3. $C(p, q) = 0$

Let $\mathcal{P}^C$ be as in Definition 2.11. We show that $\mathcal{P}^C$ is not revision homomorphic under McGee's map NS. According to McGee's construction,

$[NS(i/^C p)](q) = \epsilon(1 - \epsilon)/(1 - \epsilon^2) = \epsilon/(1 + \epsilon)$, where ϵ is an infinitesimal.

And $[NS(i)](q \mid p) = x/y$ where $x = [NS(i)](p \wedge q)$ and $y = [NS(i)](p)$, for $y \neq 0$.

13

Now, $x = \epsilon^2(1 - \epsilon)/(1 - \epsilon^3)$.
And $y = [(1 - \epsilon)/(1 - \epsilon^3)][(1 + 2\epsilon^2)/2]$.
So $x/y = 2\epsilon^2/(1 + 2\epsilon^2)$. Now, assume that revision homomorphism holds. In particular, assume:

$$[NS(i/^C p)](q) = [NS(i)](q \mid p) = x/y.$$

Then we have
$$\epsilon/(1 + \epsilon) = 2\epsilon^2/(1 + 2\epsilon^2).$$

So $1 + 2\epsilon^2 = 2\epsilon + 2\epsilon^2$.
So $\epsilon = 1/2$.

But this contradicts the assumption that ϵ is infinitesimal.

3.4. A generalized result

In the following proof, we assume that B is finite. The proof relies heavily on McGee's construction of a series of formulas that enables us to formulate patterns of conditionalization.

Definition 3.4. M_n.
   With respect to a fixed iterative probability system rooted in i_0, let
$D(p) = \bigvee\{q \ / \ q \leq p$ and $P_{i_0/p}(q) = 0\}$. Define N and $M_0, \ldots, M_{N+1}$ as follows.

$M_0 := \top$.
$M_{n+1} := D(M_n)$ if $D(M_n) \neq \bot$.
$N := n$ if $D(M_n) = \bot$.
$M_{N+1} := \bot$.

Lemma 3.1. If $n \leq N$ and $P_{i_0/M_n}(p) = 0$ for each $m < n$, then $p \leq M_n$.
   Proof. Straightforward, by induction on n. (This lemma is proved in McGee [29].)

Lemma 3.2. If $P_{i_0/M_n}(p) = 0$ for all $n, 0 \leq n \leq N$, then $P_{i_0/p} = 1_B$ and $\square_{i_0} \leq \neg p$.

   Proof. Suppose that $P_{i_0/M_n}(p) = 0$ for all $n, 0 \leq n \leq N$. Then, in particular, $P_{i_0/M_N}(p) = 0$. By Lemma 3.1, $p \leq M_N$. By Definition 3.4, we have $D(M_N) = \bot$. Now, $D(M_N) = \bigvee\{q \ / \ q \leq M_N$ and $P_{i_0/M_N}(q) = 0\} = \bot$. Therefore, $p \leq D(M_N)$. So $p \leq \bot$, i.e. $p = \bot$. Then $\square_{i_0} \leq \neg p$ and $P_{i_0/p} = 1_B$.

Lemma 3.3. If Conjunctive A-Cumulativity is satisfied then $\square_i \leq \neg(p \wedge q)$ iff $\square_{i/p} \leq \neg q$.

   Proof. First, suppose that $\square_{i/p} \leq \neg(p \wedge q)$. By Conjunctive A-Cumulativity, $\square_{i/p} = \square_i \wedge p$, and, in particular, $\square_{i/p} \leq p$. So $\square_{i/p} \leq \neg q$.

   Second, suppose that $\square_{i/p} \leq \neg q$. By Conjunctive A-Cumulativity, $\square_i \wedge p \leq \neg q$, so $\square_i \wedge p \leq \neg(p \wedge q)$. By Conjunctive A-Cumulativity, $\square_{i/p} \leq \neg(p \wedge q)$.

14

Definition 3.5. $K(p)$.

If $P_{i_0/p} \neq 1_B$ then $K(p) :=$ the least n such that $P_{i_0/M_n}(p) > 0$.

If $P_{i_0/p} = 1_B$ then $K(p) := N + 1$.

Lemma 3.4. If $\square_i \leq \neg q$ and $\square_i \not\leq \neg p$ then $P_{i/p}(q) = 0$.

Proof. Suppose that $\square_i \leq \neg q$ and $\square_i \not\leq \neg p$. Now, by Definition 2.5(7), $P_{i/p} \neq 1_B$. So $P_{i/p}(q) = 1 - P_{i/p}(\neg q)$. And $P_{i/p}(\neg q) = 1$, by Definition 2.5(4).

Lemma 3.5. $P_{i_0/p}(q) = \dfrac{P_{i_0/M_{K(p)}}(p \wedge q)}{P_{i_0/M_{K(p)}}(p)}$.

Proof. By Lemma 3.1, $p \leq M_{K(p)}$. Then $p = M_{K(p)} \wedge p$, so

$$P_{i/p}(q) = P_{i/M_{K(p)} \wedge p}(q).$$

But

$$P_{i/M_{K(p)} \wedge p}(q) = P_{i/M_{K(p)}, p}(q)$$

by Conjunctive Cumulativity. And since $P_{i/M_{K(p)}}(p) > 0$, we have

$$P_{i/M_{K(p)}, p}(q) = \dfrac{P_{i/M_{K(p)}}(p \wedge q)}{P_{i/M_{K(p)}}(p)}.$$

The following theorem shows that McGee's result can be generalized to the iterative case where conservative Conjunctive Cumulativity is satisfied.

Theorem 3.4. Let $\mathcal{P}$ be a conservative iterative probability system over a finite Boolean algebra, rooted in i_0, that satisfies Conjunctive Cumulativity. Then if $i_0 \neq \omega$ there is a nonstandard probability MC-representation NS of $\mathcal{P}$ satisfying the Revision Homomorphism condition.

Proof.

This proof diverges from McGee's. Define $\epsilon_0, \ldots, \epsilon_{N+1}$ as follows.

$\epsilon_0 := 1$.
If $n < N$ then $\epsilon_{n+1} :=$ an arbitrary constant that is infinitesimal with respect to ϵ_n.
$\epsilon_{N+1} := 0$.

Then let

$$[NS(i_0)](p) := \Sigma_{n=0}^{N}(\epsilon_n - \epsilon_{n+1})P_{i_0/M_n}(p)$$

$$[NS(i_0/p_1,\ldots,p_m)](q)) := \frac{\Sigma_{n=0}^{N}(\epsilon_n - \epsilon_{n+1})P_{i_0/M_n}(p_1 \wedge \ldots \wedge p_m \wedge q)}{\Sigma_{n=0}^{N}(\epsilon_n - \epsilon_{n+1})P_{i_0/M_n}(p_1 \wedge \ldots \wedge p_m)},$$

where it is assumed that $i_0/p_1,\ldots,p_m \neq \omega$; in case $i = \omega$, $NS(i)$ is undefined.

Note that NS is presented as a two-place function from the index $i_0/p_1,\ldots,p_n$ and q to the nonstandard reals. But it is defined in terms of the probabilities of $p_1 \wedge \ldots \wedge p_n \wedge q$ and $p_1 \wedge \ldots \wedge p_n$. Lemma 2.1 resolves this difficulty, showing that NS is well defined. (It is here that we appeal to Conservativity.)

We now show that the conditions for MC-representation are met, as well as the Revision Homomorphism condition.

1. First, Condition (1) on MC-representation: we wish to show that for all p, q, that the difference

$$P_{i/p}(q) - NS(i)(q \mid p) \qquad (1\text{-}1)$$

is infinitesimal, if $[NS(i)](p) \neq 0$.

We can suppose without loss of generality that $i = i_0$. Now, in view of Lemma 3.5 and Conjunctive Cumulativity, (1-1) becomes

$$\frac{P_{i_0/M_{K(p)}}(p \wedge q)}{P_{i_0/M_{K(p)}}(p)} - \frac{\Sigma_{n=K(p)}^{N}(\epsilon_n - \epsilon_{n+1})P_{i_0/M_n}(p \wedge q)}{\Sigma_{n=K(p)}^{N}(\epsilon_n - \epsilon_{n+1})P_{i_0/M_n}(p)}. \qquad (1\text{-}2)$$

Let
$$P_n = P_{i_0/M_{K(p)+n-1}}(p)$$
$$J = N + 1 - K(p),$$
$$Q_n = P_{i_0/M_{K(p)+n-1}}(p \wedge q),$$
$$\delta_n = \epsilon_{K(p)+n-1} - \epsilon_{K(p)+n}.$$

Then (1-2) becomes
$$\frac{Q_1}{P_1} - \frac{\delta_1 Q_1 + \ldots + \delta_J Q_J}{\delta_1 P_1 + \ldots + \delta_J P_J}. \qquad (1\text{-}3)$$

Simplifying (1-3) and combining terms, we obtain
$$\frac{\delta_2(P_2 Q_1 - P_1 Q_2) + \ldots + \delta_J(P_J Q_1 - P_1 Q_J)}{P_1(\delta_1 P_1 + \ldots + \delta_J P_J)}. \qquad (1\text{-}4)$$

Now, choose
$$\gamma = max\left(\left\{\frac{\delta_n(P_n Q_1 - P_1 Q_n)}{\delta_{n-1}P_1 P_{n-1}} \middle/ 2 \leq n \leq J\right\}\right).$$

16

The quantity γ is infinitesimal, because $\delta_n \ll \delta_{n-1}$ in view of the original choice of the ϵ_n, and P_n and Q_n are standard reals.

And (1-4) is bounded by γ since γ was defined so that for all n such that $2 \le n \le J$ we have:

$$\delta_n(P_n Q_1 - P_1 Q_n) \le \gamma \delta_{n-1} P_1 P_{n-1}.$$

2. Second, Condition $(2')$ on MC-representation.

First, we establish Condition (2) of Definition 3.1. That is, we show $[NS(i)](p) = 0$ iff $\square_i \le \neg p$.

Because of Lemma 2.2, we can assume without loss of generality that $i = i_0/q$.

First, suppose that $[NS(i)](p) = 0$.
Then $\Sigma_{n=0}^{N}(\epsilon_n - \epsilon_{n+1})P_{i_0/M_n}(q \wedge p) = 0$.
Then, by Lemma 3.2, $\square_{i_0} \le \neg(q \wedge p)$. So by Lemma 3.3, $\square_{i_0/q} \le \neg p$,
i.e., $\square_i \le \neg p$.

Conversely, suppose that $\square_i \le \neg p$, i.e., $\square_{i_0/q} \le \neg p$. Now, by Conjunctive Cumulativity, we have $(q \wedge \square_{i_0}) \le \neg p$, so that $\square_{i_0} \le \neg(q \wedge p)$.
Therefore, by Definition 2.5(5), $P_{i_0/M_n}(q \wedge \neg p) = 0$ for all n. So, by the definition of NS, $[NS(i_0/q)](p) = 0$. I.e., $[NS(i)](p) = 0$.

This establishes Condition (2).

Since B is finite, $\bigwedge \square(NS(i))$ exists, for any $i \in \mathcal{I}$.
By Condition (2), $\square_i \le \bigwedge \square(NS(i))$; and clearly, $\bigwedge \square(NS(i)) \le \square_i$. So Condition $(2')$ is proved.

3. Third, Revision Homomorphism.

Show that $NS(i/p) \doteq \lambda q [NS(i)](q \mid p)$.

Let $i = i_0/r$. Where $NS(i/p)$ is defined, i.e. where $i/p \ne \omega$, this follows immediately from the definitions of $NS(i/p)$ and $NS(i)$.

Suppose, then, that $i/p = \omega$. Then by Definition 2.5(8) $\square_i \le \neg p$. So by Condition (2) on MC-representation, established above, we have $[NS(i)](p) = 0$.
So $[NS(i)](p) = 0$, so $\lambda q [NS(i)](q \mid p)$ is undefined.
Also, since $i = i_0/r$, we have $\square_{i_0} \le \neg(p \wedge r)$. So by Condition (2), $[NS(i_0)](p \wedge r) = 0$, so by Definition 2.5, $NS(i/p)$ is undefined.
On the other hand, if $\lambda q [NS(i)](q \mid p)$ is undefined, i.e., if $[NS(i)](p) = 0$, then by Condition (2) on MC-representation, established above, $\square_i \le \neg p$.
So $NS(i/p)$ is also undefined.

This proof assumes crucially that the Boolean algebra is finitely generated. Thus, the method works only for finite languages. McGee [29, p. 182] claims that the result can be extended to the infinite case by an analogue of the ultraproduct construction. We do not see how to generalize that construction to prove compactness for this logic; the problem of whether these representation theorems can be generalized to the infinite case remains open.

4. Quantitative and qualitative methods for epistemic update

According to a popular qualitative view, the locus of inquiry for an ideal agent is a set S of propositions—the ones that are fully believed by the agent. And it is usual to model these propositions using sets of possible worlds. The intersection of the members of S, which we can call K, compiles the agent's beliefs into a single proposition.

This traditional model remains fairly open concerning other components of an agent's epistemic state. Initially, it was also silent about the origin of beliefs, and about how K is (or should be) updated when a new proposition is learned or supposed, at least in the nonmonotonic case where this proposition is inconsistent with K.

This last problem has been intensely scrutinized during the last two decades. This scrutiny has been performed, nevertheless, in the context of *pluralists* forms of Bayesianism where qualitative belief and probability are taken as separate epistemological primitives. Therefore, belief change has been studied independently of probability change. The pioneering AGM theory remains an important contender. We review this theory briefly in Section 7.

The main goal of the following sections will be to make explicit the qualitative commitments imposed by a cumulative account of iterative probability kinematics. The intended outcome of this study is to reveal the qualitative commitments imposed by a unified account of generalized probability kinematics considered in previous sections. The idea, therefore, is to study changes of belief as supervening on (a primitively given) notion of probability change. In other words, the idea is to study the (qualitative) notion of belief change induced by a form of *radical probabilism* where the only primitive is a generalized notion of conditional probability. Surprisingly, this form of belief change supervening on probability change has hardly been studied in the literature. This is so even when such a form of radical probabilism has been recently advocated both on theoretical and applied grounds.[16] We will proceed as follows. First we will review some of the standard qualitative axiomatizations of belief change and we will evaluate their adequacy as representations of the notions of learning and/or supposing. Then we will focus on making explicit the qualitative structure of *probability-based* change induced by a notion of probability kinematics obeying cumulativity. This will be accomplished in two steps.

First we will focus on determining the notions of *full belief* and *qualitative expectations* for a primitively given Popper function C. This first step will rely on previous work done by Bas van Fraassen in [41] (and recently extended and slightly modified in [3], [6], [4] and [5]). The basic idea is to show that for each Popper function C (modulo some constraints)

---

[16]Bas van Fraassen has defended this form of probabilism in various papers, of which [41] is one of the most salient for our purposes. Various applications in Economics are summarized in [19]. Related applications in computer science appear in [16] – especially concerning the method called 'Bayesian qualitative updating with hard evidence', which will be discussed later on in more detail.

there is a nested set of propositions called *probability cores*. We will propose to interpret the largest core of C as the encoding of full beliefs for C, and the smallest as the expectations for C.

Secondly we will study how the system of cores for C changes when C changes as a result of supposing a new item. This, of course, also will give us a picture of how full beliefs and qualitative expectations for a given C evolve when C changes for the sake of the argument. The resulting notion of qualitative supposition will then be axiomatized and compared with other accounts of epistemic change.

We remind the reader that in section two we proposed an alternative definition of Popper functions that does not impose cumulativity by fiat. This representation kept track of a qualitative notion of *epistemic necessity* $\Box i$ associated with each index i. But, of course, a subclass of such indices will represent standard Popper functions - where cumulativity *is* imposed by definition. In the following sections we will show that those indices can be naturally represented as pairs $\langle P_i, \Box_i \rangle$, where the first component $P_i(q)$ is $C(q \mid p)$ for some two-place function C and proposition p; and the second component is the largest core of this C. This analysis will shed some light on the intended interpretation of the notion $\Box_i$ used in previous sections. When a set of indices represents a function C, $\Box i$ can be seen as encoding the full beliefs for $\lambda q \, C(q \mid p)$. It will also give us the tools to clarify some open issues in the abstract axiomatization on Popper functions. Finally, this analysis will give us the tools for commenting on the intended interpretation of the notion of core change induced by probability kinematics.

So, our goal in the following sections is not to question any form of cumulativity built into probability kinematics, but to understand the qualitative constraints that follow from assuming it. In order to make the last point more transparent it would be convenient to sum up here what we have accomplished so far. We first offered an account of iterative probability kinematics that does not have cumulativity built into it. Then we showed that cumulativity needs to be assumed in order to obtain a natural representation of Popper functions in terms of non-standard probability. The conclusion is that a unified account of Popper functions and non-standard probability is coherent with a view of iterative update satisfying Conjunctive Cumulativity (CCIP). Moreover such a unified account *requires* CCIP. This result is to some extent surprising. Usually non-standard measures, or Popper functions, or both are presented as an unproblematic and general extension of standard probability theory.[17] Our result seems to indicate that an equivalence result linking Popper functions and non-standard probability will seriously constrain the set of possible interpretations compatible with these formalisms (namely to those interpretations for which cumulativity might make sense). The goal in the following sections is to make explicit the nature of the qualitative constraints induced by cumulative accounts of probability kinematics. We will do so by axiomatizing the form of probability-based qualitative change induced by an account of iterative probability kinematics that *has* cumulativity built into it. The notion that thus arises will be then compared with other qualitative models of supposition and learning. The conclusion will be that the resulting notion of probability-based change is a notion seldom studied before (and inconsistent with most accounts of belief change offered in recent years).

Our immediate goal will be to consider how probabilistic kinematics and qualitative

---

[17]At least this is so for probabilistic spaces of limited size.

kinematics connect. Although there is a fair amount of recent work devoted to formalizing notions of belief dynamics, almost nothing focuses on capturing the qualitative constraints presupposed by two-place probability functions. We are aware of only one such theory, namely the one offered by Hájek and Harper in [18]. Unfortunately this account only covers single-shot hypothetical changes, and we will argue below that an unmodified version of it is inadequate. Nevertheless Hájek and Harper's approach can be reconciled with ours. The last section of this essay shows that a *modified* version of Hájek and Harper's axioms follows from our account.

Section 5 introduces two different theories of qualitative belief change. This brief but self-contained review is needed to understand both the gist of our own proposal and the sort of technical changes that are needed to reconcile Hájek and Harper's theory with ours. Our ultimate goal in the coming sections will be to offer an account of the qualitative properties needed to extend McGee's mapping.

5. Knowledge in flux: two theories of qualitative change

5.1. AGM

A *basic model* of change is a triple $\langle \Omega, \mathcal{P}, * \rangle$, where Ω is a non-empty set of states, $\mathcal{P}$ is a field of propositions with respect to Ω and $*$ is a function $* : 2^{\Omega} \times \mathcal{P} \to 2^{\Omega}$, obeying the following constraints. Subsets K of Ω represent bodies of full belief.

$(5.1.1)$ $K * p \subseteq p$ (Success)

$(5.1.2)$ If $K \cap p \neq \emptyset$, then $K \cap p = K * p$ (Preservation)

$(5.1.3)$ If $p \neq \emptyset$, then $K * p \neq \emptyset$ (Consistency Preservation)

$(5.1.4)$ If $(K * p) \cap q \neq \emptyset$,

         then $(K * p) \cap q = K * (p \cap q)$ (Conjunction)

More constraints can be added to the models, but the ones presented above are enough for our purposes.

The so-called "AGM approach" [15] satisfies these five conditions. It is prima facie unclear whether the intended interpretation of the axioms is suppositional or not. Some constraints, like Success, seem constitutive of supposing (and inappropriate for inquiry). But we will argue that Preservation precludes a suppositional interpretation. The argument is based on an example initially proposed by McGee in [28].

5.2. Cumulativity and the inadequacy of AGM as a theory of supposing

The AGM approach leaves iterated change relatively unconstrained. It has been criticized (perhaps unfairly) because of this limitation. But the five constraints listed above do impose *some* constraints on iteration. For example, in [2] it was noted that the following postulate is a consequence of Preservation:

If $K \subseteq p$ and $K \neq \emptyset$, then $(K * p) * q = K * q$.   (Invariance)

On the suppositional interpretation, Invariance says that if r follows upon supposing first p and then q, then r should follow upon just supposing q, provided that p is believed in the current (consistent) view. But consider the following example presented by McGee in [28]:

Example 5.1. (Vann McGee)

> Opinion polls taken just before the 1980 election showed the Republican Ronald Reagan decisively ahead of the Democrat Jimmy Carter, with the other Republican in the race, John Anderson, a distant third. Those apprised of the poll results believed, with good reason: If a Republican wins the election, then if it's not Reagan who wins it will be Anderson. A Republican will win the election. Yet they did not have reason to believe If it's not Reagan who wins, it will be Anderson.

It is reasonable to conclude that Anderson will win upon supposing first that the winner will be a Republican, and second that he won't be Reagan. And one should conclude that *Carter* will win after supposing that Reagan won't win. But Invariance tells us that in this situation, where it is believed initially that a Republican will win, one should reject this second conclusion and conclude instead that Anderson will win, supposing that Reagan won't. This result is, to say the least, implausible.

So, although AGM appeals to some postulates (like Success) that seem adequate as universal constraints on supposition, it admits other postulates (like Preservation and its corollary, Invariance) which fail on a suppositional interpretation of update. Moreover, there are convincing arguments that the AGM postulates capture only *some* features of supposition, and *some* features of rational inquiry. (See Levi [26] and Collins [10] for arguments showing that AGM is not a completely adequate theory of supposition. See Levi [25] for arguments showing that AGM doesn't fit some features of rational inquiry.) These considerations suggest that it may be worthwhile to reconsider the axiomatization of qualitative belief kinematics in light of a thorough reexamination of the sources of the intuitions that motivate the theory.

In their foundational critique of AGM, Friedman and Halpern suggest that postulates need to pay closer attention to the "underlying ontology," where

> An ontology must make clear what the agent's epistemic state is, what types of observations the agent can make, the status of observations, and how the agent goes about revising the epistemic state. [14, p. 430]

Friedman and Halpern confine themselves to intuitions relating to rational inquiry; their point is that there may be important differences in these intuitions depending on the underlying agent model, and that for this reason the axiomatization of qualitative belief revision should be accompanied by more detailed explicit models of the rational agent. To this point—which we endorse—we wish add that the recommendation applies equally well to the suppositional construal of qualitative conditionalization.

What about Cumulativity? A purely qualitative version of the postulate follows:

$$(K * p) * q = K * (p \cap q) \tag{Cumulativity}$$

Obviously, Conjunction is weaker than Cumulativity. Moreover AGM is inconsistent with Cumulativity, because Cumulativity clashes with Success and Consistency Preservation.[18] Thus, Cumulativity runs against the basic account of theory change incorporated in the AGM theory.

5.3. Non-invariant theories: Darwiche and Pearl

In [11], Darwiche and Pearl presented a theory of belief change inspired by previous ideas of Wolfgang Spohn [37]. Their theory (called "DP" from now on) is less inadequate than AGM as an account of supposing.

An *extended model* of change is a quintuple $\langle \mathbf{E}, \Omega, \mathcal{P}, *, \rho \rangle$, where $\mathbf{E}$ is a non-empty set, Ω is a snon-empty et, $\mathcal{P}$ is a field of propositions with respect to Ω, $*$ is a function $* : \mathbf{E} \times \mathcal{P} \to \mathbf{E}$, and ρ is a function $\rho : \mathbf{E} \to 2^{\Omega}$.

Intuitively, $\mathbf{E}$ is a set of epistemic states, Ω is a set of states or worlds, * maps epistemic states and propositions to epistemic states, and ρ maps epistemic states to their associated belief sets. Darwiche and Pearl constrain their extended models by imposing the following *basic postulates*:

$$\rho(E * p) \subseteq p \qquad \qquad \text{(Success)}$$
$$\text{If } \rho(E) \cap p \neq \emptyset, \text{ then } \rho(E) \cap p = \rho(E * p) \qquad \text{(Preservation)}$$
$$\rho(E * p) \text{ is consistent if } p \text{ is.} \qquad \text{(Consistency Preservation)}$$
$$\text{If } \rho(E * p) \cap q \neq \emptyset, \text{ then}$$
$$\rho(E * p) \cap q = \rho(E * p \cap q) \qquad \text{(Conjunctive Revision)}$$

In addition to these basic postulates, Darwiche and Pearl impose the following special postulates for iterated revision or *C-postulates*:

(C1) If $p \subseteq q$, then $\rho((E * q) * p) = \rho(E * p)$.
(C2) If $p \subseteq \neg q$, then $\rho((E * q) * p) = \rho(E * p)$.
(C3) If $\rho(E * p) \subseteq q$, then $\rho((E * q) * p) \subseteq q$.
(C4) If $\rho(E * p) \not\subseteq \neg q$, then $\rho((E * q) * p) \not\subseteq \neg q$.

In this setting, Cumulativity is $\rho((E * p) * q) = \rho(E * (p \cap q))$. Clearly, this equation does not follow from DP. In fact, Cumulativity is inconsistent with DP, because it clashes with Consistency Preservation and Success. DP, nevertheless, is less inadequate than AGM as an account of non-cumulative forms of supposing (see section 8 below). Examples relating to this point, and an attempt to modify DP to treat supposition, are considered in [2].

Several useful examples are discussed in [11]. For example, epistemic states can be encoded as *rankings*. A ranking is a function from the set Ω of worlds into the class of ordinals. A ranking F is extended to (consistent) propositions by requiring that the rank $F(p)$ of a proposition be the smallest rank assigned to a p-world:

$$F(p) = min(\{F(w) \ / \ w \in p\}).$$

---

[18]To see this, consider the case $q = \neg p$, with p, q consistent.

So, we can consider the following instance $\langle \mathbf{F}, \Omega, \mathcal{P}, *, \rho \rangle$, of an extended model, where $\mathbf{F}$ is the set of ranking functions, ρ is defined as the set $\{w \: / \: F(w) = 0\}$, and the function * is defined as:

$$(F * p)(w) = \begin{cases} F(w) - F(p) & \text{if } w \in p \\ F(w) + 1 & \text{otherwise} \end{cases}$$

Even if we focus on a suppositonal interpretation of theories like DP, nested English conditionals with contradictory antecedents are hardly intelligible. So one might consider the following, milder version of cumulativity.

> If $p \cap q \neq \emptyset$, then $\rho((E * p) * q) =$
> $\qquad \rho(E * (p \cap q))$ 　　　　　　　　(Weak Cumulativity)

The *-method proposed above fails to satisfy even Weak Cumulativity. Let the set Ω of worlds of an extended structure be $\{w_1, w_2, \ldots\}$, and consider the following epistemic state κ:

$F(w)$	Possible world w
0	w_1, w_4
1	w_2, w_3

Let $\text{Odd} = \{w_i \: / \: i \text{ is odd}\}$, and $\overline{\text{One}} = \{w_i \: / \: i \neq 1\}$. It is easy to check that $\rho((F * \text{Odd}) * \overline{\text{One}}) = \{w_3, w_4\}$, whereas $\rho(F * (\text{Odd} \cap \overline{\text{One}})) = \{w_3\}$, although $\text{Odd} \cap \overline{\text{One}} \neq \emptyset$.

6.　From two-place probability functions to belief

In the previous sections we briefly reviewed some of the salient theories of belief change and we considered their adequacy as theories of supposing. In this section we will consider which qualitative notion of supposition is adequate for the non-standard probability functions studied in the first three sections of this article. This will require tackling two hard problems in formal epistemology. The first problem can be formulated as follows: which is the set of certainties (beliefs) naturally correlated to each infinitesimal probability function (or, equivalently, to each Popper function). The second problem is related to the form of *probability-based* belief revision induced by probability kinematics—in other words, to what laws dictate how a body of certainties for a function P changes when P is updated.

This section will mainly focus on the first problem. It will also prepare the ground for an axiomatic treatment of the second problem in sections 6.1 and 7. It is well known that some easy solutions to this problem are defective. (For example, the solution based on using probability-one events to define certainty is known to be problematic [23].) To bridge two-place functions and belief, we will appeal to a proposal first offered by Bas van Fraassen. Van Fraassen's approach is sufficiently well behaved to deliver a paradox-free account of the certainties associated with Popper functions, at least if we restrict ourselves to countable probability spaces.[19]

---

[19]See [7], [34], [4] for further discussion about this topic.

Let's briefly review the main axioms obeyed by two-place measures: (I) for any fixed p, the function $C(x, p)$ as a function of x is either a probability measure,[20] or has constant value 1 and (II) $C(q \wedge r, p) = C(q, p)C(r, q \wedge p)$ for all propositions p, q, r. The unconditional probability of p, $pr(p)$, is $C(p, \top)$. Axiom II (usually called the Multiplication Axiom) differs from the usual rule for computing conditional probabilities by allowing conditioning on events of measure zero.

As we explained above, in order to determine the qualitative constraints on belief and conditional belief imposed by this account, we need to determine the body of belief associated with each two-place measure and how these bodies of belief change when the measures change.

Paradox-free accounts of the first problem were considered by van Fraassen in [41] and then refined, extended and slightly modified in [3] and [6]. The problem these accounts need to circumvent is the so-called *lottery paradox* [23], which arises when full belief is identified with measure one. The account presented here, following the proposal in [6], dissolves the paradox by making probability one a necessary but not a sufficient condition for full belief. This approach has the additional advantage of producing the best fit between some well known models of qualitative belief change and probability kinematics.

If $C(x, p)$ is a probability measure as a function of x, then p is *normal*; otherwise it is *abnormal*. The notion of normality is closely connected to an epistemic analysis of *a prioricity*: (A) p is *a priori for C* iff $C(p, x) = 1$ for all x. Thus p is *a priori for C* iff the complement of p is abnormal for C.

A probability *core* is a set K which is normal and satisfies the *strong superiority condition* (SSC): for all nonempty subsets A of K and all B disjoint from K, $C(B, A \cup B) = 0$.

The family of cores induced by a non-coreless two place probability function C is nested. That this holds in general has been shown in [41]. Moreover it can also be shown for spaces of arbitrary size that the chain of belief cores induced by a non-coreless 2-place function C cannot contain an infinitely descending chain of cores (the proof is presented in [3]).[21]

Thus the cores are well ordered under inclusion and closely resemble the so-called Grove system of spheres [17], widely used as a semantical tool in belief revision - or Spohn's 'ordinal conditional functions' [37]. When the space is countable it can be shown that there is a smallest as well as a largest core. Moreover, the smallest core has measure 1. And, more importantly, this innermost core is constituted by all 'heavy' points carrying unconditional positive measure (this is shown in [6]). In the following we will restrict our analysis to finite or at most countable spaces—[4] provides arguments in favor of restricting applications related to models of supposition to such spaces.

We will rely heavily below on some recent studies revealing how the web of cores of a function C changes when updated by a proposition p. Let S^C be the system of cores for C. Let, in addition, $C(x, y \wedge p)$ be the update of $C(x, y)$ with the proposition p.[22] Then in [5] it is shown that, provided that p intersects the largest core for C, the system of cores of $P_{i/p}$

---

[20]Results requiring countable additivity will be flagged below. The axiom will not be assumed in general.

[21]This result requires postulating countable additivity. Although the property is useful, the main results used in the following sections will not depend on it. On the other hand, the existence of a smallest core is, of course, automatically guaranteed for finite spaces

[22]As we explained above, we will build cumulativity into this account of probability change. Our interest is to study the form of qualitative change induced by this form of probability kinematics.

is:

$$S' = \{c \cap p \; : \; c \in S^{P_i} \text{ and } c \cap p \neq \emptyset\}$$

From now on we will refer to the fact just alluded as "Update of Cores." When p does not intersect the largest core for C, the unique core of $C(-, p)$ is set identical to the empty set.

Two salient epistemic notions are definible using a system of cores for a function C: the largest core for C, and the smallest core for C. Both propositions carry measure one, but only the former is *robust with respect to suppositions*. I.e. if $c(C)$ is the largest core for C, it can be shown that $C(c(C), x) = 1$ for all propositions x. The smallest core also carries measure one but it does not exhibit the 'epistemic robustness' with respect to suppositions of the largest core. We will argue below that the largest core should be seen as encoding the *full beliefs* for C, while the smallest core should be seen as encoding the *expectations* for C.[23] We will also argue that, in general, $\square_i$ should be seen as encoding the full beliefs for each index i.

Now we will turn to the idea of representing Iterative Probability Systems by Popper functions. We remind the reader that Iterative Indices (i.e. the indices defined in section 2) are pairs of one-place standard measures P_i, and a proposition $\square_i$. For each index i, i/p 'selects' the 'closest' standard measure for P_i and a new necessary proposition.

Our goal here is to focus on a particular sub-class of such indices, namely the ones such that the first component $P_i(q)$ of an index i is $C(q, p)$ for some (fixed) two-place function C and a proposition p. The second component of the index i is the largest core of this function $\lambda q C(q, p)$. On the other hand, for each index i (whose first component is $C(q, p)$) and proposition r, the first component of i/r is $C(q, p \wedge r)$ and the second component is the largest core of $C(q, p \wedge r)$.

The first component of the index i_0 is $\lambda p C(p, \top)$, and the second component is the largest core of this function. In addition we assume that all indices are rooted in i_0 – i.e. are reachable from it via a finite sequence of updates.

The idea is to show that the notion of epistemic necessity $\square_i$ used in section 2 can be naturally represented in this setting by the largest core of $\lambda p C(p, q)$, where $\lambda p C(p, q) = \lambda p P_i(p)$. Since, for every proposition q, $\lambda p C(p, q)$ satisfies the axioms for one-place probability measures (by Axiom I adopted in section 6), we are focusing on a legitimate and 'natural' sub-class of probability indices. On the other hand only cumulative transitions among indices will be allowed - as in standardly defined Popper functions.

Definition 6.1. Representation of an Iterative System by a Popper function.

---

[23] The model respects the intuitive fact that every fully believed proposition is expected – while the converse does not necessarily holds. We remind the reader that a proposition A is fully believed if and only if it is a superset of the largest core, and a proposition is expected as long as it is a superset of the innermost core. The expected inclusion among the *theories* induced by the largest and innermost core also holds. If $\text{Th}(c(C))$ is the (closed) set of sentences true in the largest core, and $\text{Th}(\rho(C))$ the theory corresponding to the innermost core, $\text{Th}(c(C)) \subseteq \text{Th}(\rho(C))$.

Let $\mathcal{P}$ be an iterative system over a Boolean algebra $\mathcal{B}$, rooted in i_0. $\mathcal{P}$ represents a (normal and non-coreless) Popper probability function C on $\mathcal{B}$ if

(1) $P_{i_0} = \lambda p C(p, \top)$,

(2) for all $p_1, \ldots, p_n$, $P_{i_0/p_1,\ldots,p_n} = \lambda p C(p, p_1 \wedge \ldots \wedge p_n)$,

(3) for all $p_1, \ldots, p_n$, $\square_{i_0/p_1,\ldots,p_n}$ is the largest core of $\lambda p C(p, p_1 \wedge \ldots \wedge p_n)$.

(4) If $p = \bot$, then, $P_{i/p} = P_\omega$.

(5) $P_{\omega/p} = P_\omega$.

Now we can establish the following fact:

Theorem 6.1. Representation of $c(P_i)$ as epistemic necesity

Let $\mathcal{P}$ be an iterative system over a Boolean algebra $\mathcal{B}$, rooted in i_0 representing a Popper probability function C. $\mathcal{P}$ is an Iterative Probability System obeying Conjunctive Cumulativity.

Proof. We need to check the properties 4 to 7 required in Definition 2.5.

(4): We need to show that: $P_{i/p}(p) = 1$, for all p. When $i = i_0$ this is easily established by the fact that $C(p, p) = 1$. In fact, $P_{i_0/p} = \lambda p C(p, p) = 1$. Since $\mathcal{P}$ is an iterative system over a Boolean algebra $\mathcal{B}$, rooted in i_0, for any index j $\neq i_0$ we have a sequence $p_1, \ldots, p_n, \ldots$, such that $j = i_0/p_1, \ldots, p_n, \ldots$. But then the Multiplication Axiom guarantees that $P_{i_0/p_1,\ldots,p_n,\ldots,p} = \lambda p C(p, p_1 \wedge \ldots \wedge p_n \wedge \ldots \wedge p)$ $= 1$. In general once the i_0-case is established, proving the correspondent fact for an index different from i_0 mimics the proof used in the i_0-case. So in the following we will only establish (5), (6) and (7) for i_0.

We will establish the following fact before proving (5) and (6):

(Identity) For all propositions r, q, if $C(r, q)$ is normal, $C(r, q) = C(r, q \wedge c(C))$, where $c(C)$ is the largest core of C.

In order to see that Identity holds consider the following instance of the Multiplication Axiom:

$C(r \wedge c(C), q) = C(c(C), q) \cdot C(r, q \wedge c(C))$

Since $c(C)$ is a priori for C,[24] we have that $C(c(C), q) = 1$. On the other hand, since $C(r, q)$ is assumed to be normal, finite additivity guarantees the following fact (here we will use the notation $\overline{p}$ to denote the complement of the proposition p).[25]

---

[24]We are assuming here that the space is countable.

[25]Previously we used the notation $\neg p$ to denote complements of propositions. We change notation in this section in order to improve readability—we are working here with complex propositions like the one encoding the cores of a system of cores.

$C(r, q) = C(r \wedge c(C), q) + C(r \wedge \overline{c(C)}, q).$

The second term of the previous sum is zero. This is so because $c(C)$ is a priori and therefore $\overline{c(C)}$ has zero measure, conditional on q. Therefore, by finite additivity and the assumption that the proposition q is normal, any cell of any finite partition of $\overline{c(C)}$ should carry measure zero as well (conditional on q).[26] So, we have:

$C(r, q) = C(r, q \wedge c(C))$

as desired. Now, let p be *epistemically possible* for P_{i_0}, if and only if p intersects $c(C)$, where $P_{i_0} = \lambda p C(p, \top)$. Otherwise p is *epistemically impossible* for P_{i_0}. An immediate consequence of Identity is that if p is epistemically impossible for P_{i_0}, the function $C(-, p)$ is abnormal, and $pr_{i_0/p} = pr_\omega$. Now we are ready to prove (5) and (6).

(5) We need to establish that $P_{i_0}(c(C)) = 1$, where $P_{i_0} = \lambda p C(p, \top)$. This follows from the fact that $c(C)$ is a core for C and every core for every function carries probability one.

Now, consider $pr_{i_0/p}(c(C))$. If p is epistemically impossible for P_{i_0}, we have $pr_{i_0/p} = pr_\omega$, and therefore $pr_{i_0/p}(c(C)) = 1$. Otherwise, Update of Cores guarantees that $p \cap c(C)$ is a core for $C(-, p)$. Therefore, since all cores receive probability one we have $pr_{i_0/p}(c(C)) = 1$.

(6) Here we need to show that if $\square_{i_0}$ entails p, then $i_0/p = i$. So, assume that $c(C)$ entails p. We have to show that $P_{i_0/p} = P_{i_0}$ and $\square_{i_0/p} = \square_{i_0}$.

$P_{i_0/p} = \lambda q C(q, p)$, where $C(q, \top)$ is assumed to be normal and non-coreless.[27] Therefore, according to Identity,

$C(q, c(C)) = C(q, \top).$

We need to show that $P_{i_0/p} = \lambda q C(q, p) = P_{i_0} = \lambda q C(q, \top)$. Given that $C(q, \top)$ is assumed to be normal, Identity guarantees that $C(q, \top) = C(q, c(C))$. On the other hand, Update of Cores guarantees that $C(q, p)$ is normal and non-coreless with $c(C)$ as its largest core. Therefore $C(q, c(C)) = C(q, p)$, given that $\square_{i_0}$ entails p. So, we can establish that $P_{i_0/p} = \lambda q C(q, p) = \lambda q C(q, \top) = P_{i_0}$. In addition, Update of Cores also guarantees that $\square_{i_0/p} = \square_{i_0}$.

(7) Assume that $\square_{i_0} = c(C)$ entails $\overline{p}$. We need to establish that $i_0/p = \omega$. By Identity, $P_{i_0/p} = \lambda q C(q, p) = \lambda q C(q, p \wedge c(C))$. But then $C(q, p \wedge c(C)) = C(q, \perp) = P_\omega$. On the other hand, since $P_{i_0/p} = \lambda q C(q, p) = \lambda q C(q, p \wedge c(C))$, the unique core of $C(q, p)$ is empty, making $\square_{i_0/p} = \perp$.

---

[26] We should dispel a possible confusion here. Since $c(C)$ is a priori for C, we also have that $\overline{c(C)}$ is abnormal for C. But this only means that $P(-, \overline{c(C)})$ is not a probability function. $C(-, q)$ *is* nevertheless a probability function obeying finite additivity. Therefore $C(r \wedge \overline{c(C)}, q) + C(\overline{r} \wedge \overline{c(C)}, q) = C(\overline{c(C)}, q) = 0$.

[27] A slightly more general result can also be proved for functions C that need not be normal. The needed assumption is that the C function should be assumed to be non-coreless, if normal.

27

Finally, it is easy to see that Conjunctive Cumulativity is obeyed. Moreover a Popper Iterative System is also a Bayesian Iterative Probability System.

Notice that when an Iterative System is used to represent a two-place function C, other epistemic notions defined via the system of cores for C fail to provide a representation of epistemic necessity. In particular, the *qualitative expectations* encoded in the smallest core of $C(p, \top)$, are not a viable representation of $\Box_{i_0}$—properties like Identity will not be satisfied.

It is also important to see that the crucial role of $\Box_i$ is to determine the space of *epistemic possibilities* for P_i. In order to see this, notice that according to clause (4) in the representation of Iterative Systems by a Popper function, the inconsistency of p is only a sufficient, but not a necessary condition of the inconsistency of $P_{i/p}(P_{i/p} = P_\omega)$. This leaves open the possibility that revising P_i with logically *consistent* propositions might also yield P_ω as an outcome. The previous results shows that this is indeed the case. For example, $P_{i_0/p} = P_\omega$ for any (non-empty) proposition p failing to intersect $c(C)$, i.e. for any proposition that is *epistemically impossible* for C. This and other reasons summarized in [6] and [5] strongly suggest that, for every C, the largest core for C (alternatively our $\Box_i$) encodes the certainties (full beliefs) for P_i, determining the space of epistemic possibilities for P_i. McGee adopts in [29] a semi-formalized principle incorporating the former ideas:

$P_{i/p}$ is to take the constant value 1 only if p is epistemically utterly impossible. There is no need for even an ideally rational agent to formulate a plan for how she would revise her beliefs upon learning that p, if she is absolutely certain that p is false, in this extreme case, we set $P_{i/p}$ equal to 1 arbitralily.

Notice that principle is not formulated exactly in [29]. This is so because this paper lacks the resources to exactly define what is 'epistemically utterly impossible' for a probability function P_i. A theory like van Fraassen's is needed to define this notion probabilistically in terms of P_i. Here we have offered such definition, by appealing to the theory presented in [6] and [4]. McGee's principle can therefore be exactly formulated as follows for Popper functions:

(Consistency) $U_p(C)$ is abnormal only if $p \cap c(C) = \emptyset$[28]

Iterative systems provide an alternative formulation of the same principle in a more general setting:

(Consistency) $P_{i/p} = P_\omega$ only if $p \cap \Box_i = \emptyset$.

Notice that as long as conditions 9 and 10 of conservative iterated probability systems are satisfied, condition 8 and Consistency are also satisfied. As a matter of fact, both condition 8 and Consistency hold in the presence of the RTL half of condition 10. This condition guarantees that the necessary proposition ($\Box_i$) for the index i is the falsity as long as the P-component of i is $1_\mathcal{B}$. So, we have as a corollary that if $\mathcal{P}$ be an iterative system over a Boolean algebra $\mathcal{B}$, rooted in i_0 representing a Popper probability function C and satisfying

---

[28]We adopt here the notation proposed in section 2 for encoding changes with sequences of propositions $\{q_0, \ldots, q_n\}$.

conditions 9 and 10 (Definition 2.6), then $\mathcal{P}$ is a Conservative Iterative Probability System obeying Conjunctive Cumulativity.

Although consistency is a widely accepted principle in the orthodox Bayesian tradition, the forms of Bayesianism that nurtured the development of theories of belief revision and belief-contravening supposition do not accept the principle. To a large extent the main current theories of belief change and supposition allow for the formulation of policies for changing view upon learning (supposing) that p, when the agent is absolutely certain that p is false. The probabilistic tradition based on the use of Popper functions is not equipped to deal with such (full)belief-contravening changes.

The bodies of *expectations* for a measure C (encoded via the innermost core for C) can, nevertheless, be revised with propositions incompatible with those expectations, as long as these propositions are epistemically possible, i.e. compatible with the largest core for P. We will verify in the coming section that this form of expectation change *also* obeys a form of cumulativity. And we will also verify, of course, that the notion of (full) belief change induced by probability kinematics is also constrained by conjunctive cumulativity.

6.1. The notion of belief change induced by cumulative probability kinematics

We will adopt here the notation proposed in section 2 for encoding changes with sequences of propositions $\{q_0, \ldots, q_n\}$. $U_{q_0, \ldots, q_n}(C)(p \mid \top)$ denotes such sequential change. Moreover, according to the adopted definition of update, we have $U_{q_0, \ldots, q_n}(C)(p \mid \top) = C(p \mid q_0 \cap q_1 \cap \ldots \cap q_n)$. Then the following properties are satisfied, for non-coreless Functions P over a countable space.[29]

Inclusion $\rho(C) \subseteq c(C)$

Expansion $c(C) \cap p = c(U_p(C))$

Success $\rho(U_p(C))$ entails p

Preservation If $\rho(U_p(C)) \cap p \neq \emptyset$, then $\rho(C) \cap p = \rho(U_p(C))$

Restricted Consistency Preservation If $\rho(C) \neq \emptyset$, and $c(C) \cap p \neq \emptyset$, then $\rho(U_p(C)) \neq \emptyset$.

Fixity If $C = 1_{\mathcal{B}}$, then $\rho(C) = c(C) = \emptyset$ and $U_p(C) = 1_{\mathcal{B}}$.

Cumulativity for Expectations $\rho(U_{p,q}(C)) = \rho(U_{p \wedge q}(C))$

Entertainability If $c(C) \cap p = \emptyset$, then $U_p(C) = 1_{\mathcal{B}}$.

And, of course, Expansion yields Cumulativity for full belief or certainity:

Cumulativity for Full Belief $c(U_{p,q}(C)) = c(U_{p \wedge q}(C))$

In addition, Expansion, together with Fixity, entails the principle of Consistency presented in the previous section:

---

[29]The satisfaction of Fixity requires non-triviality constraints on the size of the underlying space - the universe of the space needs to contain more than one point.

7.   Hypothetical revision and cumulativity

A qualitative account of change of view can be abstracted from the previous account of probability change. In order to do that one can focus on an epistemic state E, rather than a probability function C. Such state is supposed to have two epistemic dimensions. On the one hand there is a body of expectations $\rho(E)$ associated with it as well as a body of certainties $c(E)$. These two dimensions of E obey, in turn, a series of constraints reflecting the properties mentioned above.

A *cumulative* model of change is a tuple $\langle\langle \mathbf{E}, \Omega, \mathcal{P}, *, \rho\rangle, c\rangle$, where $\langle \mathbf{E}, \Omega, \mathcal{P}, *, \rho\rangle$, is an extended model and c is a function $c : \mathbf{E} \to 2^{\Omega}$.

While $\rho(E)$ is supposed to yield the strongest proposition expected in state E, $c(E)$ captures the strongest proposition fully believed in E. The following axioms are then motivated by the previous account of iterated probability kinematics:

(7.1c) $\rho(E) \subseteq c(E)$.	(Inclusion)
(7.2c) $c(E) \cap p = c(E * p)$.	(Expansion)
(7.3c) $\rho(E * p) \subseteq p$.	(Success)
(7.4c) If $\rho(E) \cap p \neq \emptyset$, then $\rho(E) \cap p = \rho\,(E * p)$.	(Preservation)
(7.5c) If $\rho(E) \neq \emptyset$ and $p \cap c(E) \neq \emptyset$, then $\rho(E * p) \neq \emptyset$.	(Restricted Consistency Preservation)
(7.6c) If $c(E) = \emptyset$, then $c(E * p) = \emptyset$.	(Fixity)
(7.7c) $\rho((E * p) * q) = \rho(E * (p \cap q))$.	(Cumulativity)

Fixity indicates that the revision of any state whose (full) beliefs are inconsistent yields in turn a state preserving this feature. This strong condition is compatible with Cumulativity, but not required by it. Nevertheless, such a condition is required in Conjunctively Cumulative probability systems. Of course, the abstractly specified state E can not only be instantiated by a Popper function, but also by a Grove system of spheres or, even more directly, a system of cores defined on Ω, or a ranking function, etc.

The previous axiomatization has been independently proposed in [4] and [5] under the name of *hypothetical revison*. A computational prodedure obeying the axioms is offered in [16]. The following example provides some intuitive idea of how the method can be used to model the act of supposing.

Consider a spinner on a dial divided into four equal parts 1, 2, 3 and 4. Being in E means that you expect that the winners are either 1 or 4.

E	Possible worlds
CI	w_1, w_4
	w_2, w_3
	Rest of worlds

E is represented via a system of cores whose innermost core is CI and its largest core is $\{w_1, w_2, w_3, w_4\}$. According to intuition, hypothetical revision sanctions the acceptability of:

(1) If the winner lands in an odd number, then if the number is different from 1, the number is 3.

The initial act of supposing Odd is modeled as the opening of a 'suppositional window' that only includes Odd-options. This suppositional window preserves as much as possible of the structure encoded in the initial core system. For example, the ordering of worlds induced by E is not altered. So, our first act of supposing will lead us to:

$E * Odd$	Possible worlds
CI	w_1
	w_3
	Rest of worlds: $\{w_2, w_4\}$

where CI is the innermost core and $\{w_1, w_2\}$ is the largest core.[30] It is easy to check that the previously reviewed methods of change fail to capture the intuitive acceptability of (1). In particular, an analogue in this setting of the DP rule sanctions the acceptability of the following conditional with respect to E:

> (1) If the winner lands in an odd number, then if the number is different from 1, the number is either 3 or 4.

The examples considered above were first studied in [28]. In this article McGee used a qualitative model of supposition similar to the one presented here. McGee was also aware that this model of supposition requires the adoption of Cumulativity, a property which he did not find questionable. Here we have shown that the adoption of a cumulative modeling of supposition is not optional, but mandatory for probabilism. We have also shown that Cumulativity is required in order to prove that Popper functions and infinitesimal probability are two faces of the same coin - in the sense precisely specified in section3. While this clarifies the shape of a cumulative theory of supposition, it leaves the status of Cumulativity open. In fact, many accounts of supposition reject Cumulativity.

---

[30] The previous diagrams used to represent cores can be easily adapted in order to represent raking functions (or Spohn's ordinal conditional functions) of the type considered in section 5.3. We only need to assign ordinals to each rank (row in the diagrams) - or to each core, depending on the intended interpretation of the diagrams. The updating algorithm tacitly used in the examples assigns to each p-point w the updated rank $F * p(w) = F(w) - F(p)$ – see section 5.3. ¬p-points get assigned the infinite rank after a p-update. This method has been dubbed 'Bayesian updating with hard evidence' in [16], where it is recommended for computational reasons. Probabilistically the idea is that the points in the infinite ranks after a p-update are the abnormal points of $U_p(C)$ – if the initial core system corresponded to a function C. The points in the lowest rank are the ones corresponding to the innermost core and the remaining p-points configure the largest core of $U_p(C)$, which, on the other hand, are also the normal points for $U_p(C)$. All the points with positive (unconditional) measure appear in the lowest rank. The remaining points all receive unconditional measure zero. But there are various gradations distinguishing those zeros. The ones in the infinite rank are abnormal points. Epistemologically those points are impossible. The remaining normal zeros are, in turn distinguished in terms of their plausibility by intermediate cores. So, in the example, after updating with Odd, the points 2 and 4 are rendered epistemically impossible, while 1 is expected (with measure one). 3 is considered epistemically possible, but maximally unexpected. In plainer terms, after the update the agent is *certain* that the number is either 1 or 3. Given his background knowledge he is *almost sure* (or he *expects*) that the winner is 1; and he is certain that the outputs 2 and 4 are ruled out.

8. Rival views of supposing

In the previous sections, we presented a purely qualitative model of the theory of change presupposed by iterative probability kinematics. We argued that the theory can be viewed as a theory of supposition, and we showed that some of its most fundamental axioms are incompatible with AGM and DP.

Cumulativity also collides with other theories of supposition. Here we will briefly review three different theories of supposing, none of which is cumulative.

When conditionals are considered as propositions, their truth conditions with respect to a view K can be represented as follows:

> *Suppositional construal of conditional belief (SCCB)*
> $K \subseteq p > q$ iff $K\#p \subseteq q$.

The notion of change $\#$ satisfying SCCB has been proposed as an account of supposition by Gärdenfors in [15]. Previously, David Lewis proposed a probabilistic version of $\#$ in [27]. Lewis' name for this operation (*imaging*) is now more or less standard in this field. Gärdenfors' notion is incompatible with AGM [15]. Lewis' imaging is incompatible with standard conditioning and with Popper functions. Both fail to satisfy the qualitative or quantitative versions of Cumulativity and Preservation.

Instead of Preservation, $\#$ requires:

$$\text{If } K \subseteq p, \text{ then } K\#p = K. \qquad\qquad \text{(Weak Preservation)}$$

This in turn requires:

$$\text{If } K = \emptyset, \text{ then } K\#p = \emptyset. \qquad\qquad \text{(Inconsistency Preservation)}$$

A recent theory of supposing proposed by Hájeck and Harper also implements Inconsistency Preservation—see their postulate (2), [18, p. 3]. We will consider this theory below.

The $\#$ operator not only fails to obey Preservation, it also fails to satisfy:

$$\text{If } K \nsubseteq p \text{ and } K \nsubseteq \neg p, \text{ then } K\#p = K \cap p. \qquad \text{(Open Revision)}$$

Isaac Levi argued in [25] that Open Revision is an adequacy condition which needs to be satisfied by theories of supposing that extend the views first suggested by F.P. Ramsey in [31]. He offered in [25] a theory of supposition that does obey Open Revision, and fails to obey Preservation (Levi calls this method Ramsey revision). He also defended the appropriateness of using such a notion to provide acceptability conditions for conditionals. As in the case of $\#$, Levi's notion of change fails to obey Cumulativity.

Lewis' notion of supposing has played a prominent role in recent years due to the central role played by the notion of supposing in contemporary *causal decision theory*. In fact, most causal decision theorists claim that the notion of conditional probability used in the calculation of expected utility is imaging rather than standard conditioning. James Joyce has recently articulated the mathematical details of this claim in [22].

Brian Skyrms proposed yet another notion of supposition in [35]. Skyrms clearly emphasizes the differences between *learning* and *supposing* in a probabilistic context. He also makes clear how supposing (rather than learning) is the act required to evaluate certain conditionals. The following lines present the issue clearly:

> Updating subjective belief to assimilate a given piece of information and supposing what the world would be like were that bit of information true, are distinct mental acts for which different rules are appropriate. The difference is often marked by the distinction between the indicative and the subjunctive mood. The Warrenite will assert: 'If Oswald didn't kill Kennedy then someone else did' but deny 'If Oswald hadn't killed Kennedy someone else would have'.

Without going into the details of Skyrms' probabilistic models of learning and supposing, we will focus here only on a few aspects of his proposals. Our main goal is to indicate the role played by cumulative methods of change (probabilistic or not) in the theory.

The less controversial point made by Skyrms is the following: the processes of learning and supposing need to be conceptually distinguished. They are governed by different rules. There is indeed a robust consensus in the probabilistic literature concerning the need to distinguish between stochastic supposition and learning, as methods 'for which different rules are appropriate'. The distinction is less consensual among researchers who use qualitative methods to model epistemic change. Nevertheless, the distinction is easy to export from a quantitative to a qualitative setting, and we will take it for granted here.

Controversy arises at the point where the rules governing supposing need to be made explicit. Skyrms suggests (in the paper mentioned above) that learning should be modeled by conditioning *and* that this method is *also* appropriate to model the form of supposition revealed by the acceptance of indicative conditionals like 'If Oswald didn't kill Kennedy then someone else did'. On the other hand, Skyrms seems to equate the act of supposing with the form of subjunctive supposing involved in the acceptance of subjunctive conditionals like: 'If Oswald hadn't killed Kennedy someone else would have'. The probabilistic rule required to model such a mental act is, according to Skyrms, the one involved in minimizing the so-called Kullback-Leibler information measure (MAXENT).

Recently, James Joyce [22] offered an analysis of Skyrms' ideas and of their (conflictive) relationship with Lewis' analysis of (subjunctive) supposition in terms of imaging.

> I agree completely with the claim that learning and subjunctive supposition are different, and with the claim that learning and the process captured by minimizing Kullback-Leibler information are fundamentally different. What I deny is that the process captured by maximizing Kullback-Leibler information has anything to do with *subjunctive* supposition.

In the dispute concerning how to model subjunctive supposition, Joyce sides here with Lewis, proposing imaging, rather than MAXENT as the correct method. Of course, neither method is cumulative. So we find in the literature at least three models of subjunctive supposition, of which two are probabilistic (MAXENT and imaging) and one is qualitative. The methods are MAXENT (Skyrms), imaging (Lewis) and Ramsey Revision (Levi). Again, none of these methods is cumulative.

Joyce also differs from Skyrms when it comes to indicative (or 'matter-of-fact') supposition. The idea is that conditioning is not powerful enough to model all forms of indicative hypothesizing. This is based on the simple observation that there are plenty of cases in which one evaluates indicative conditionals with respect to bodies of information that assign

probability zero to their antecedents. Skyrms' own example can be used to illustrate Joyce's point. According to the Warrenite, Oswald, acting alone, killed Kennedy. Therefore, according to the Warrenite, the antecedent of 'If Oswald didn't kill Kennedy then someone else did' might carry zero measure. This is so, for example, in case the Warrenite fully believes that Kennedy's only killer was Oswald. The conclusion is that in order to model indicative supposition one needs either Popper functions or infinitesimal probability.

According to Joyce, when the unconditional probability of A is zero, one can entertain A for the sake of argument in two different ways. One way appeals to (Lewis') *image* of X given A, while another uses the two-place Popper function $P(X, A)$. The first measures the probability that X *would* have if C *were* true; while the second determines the probability that X has if C *is* true.[31]

McGee also proposed to treat indicatives via two-place functions in [29], the same article in which he offered his mapping between Popper functions and infinitesimal probability:

(8.4) *The probability of a simple conditional $A > B$ is $C(B, A)$ where C is a Popper function on the factual sentences.*

This probabilistic test improves on a previous version offered by Adams, according to which all conditionals whose antecedents carry zero measure should carry probability one. But this improved version of Adams' test is still limited; it only applies to non-nested conditionals. (In McGee's formulation of the test, 'simple' means non-nested.) In a different essay [28], McGee did consider the problem of iterated conditionals. Here, he recommends the so-called Exportation and Importation laws, both for indicatives and subjunctives. These laws require identical acceptance conditions for the conditionals of the shape 'If A, then if B, then C' and 'If A and B, then C'. In our setting, this is tantamount to requiring conjunctive cumulativity. Nevertheless Lewis' conditionals *do not* obey the Exportation and Importation laws. This is just the counterpart of the fact that Lewis' imaging does not obey Conjunctive Cumulativity. Unfortunately McGee's analysis in [28] is not probabilistic—he appeals to a possible worlds semantics in order to defend the Exportation and Importation laws. In addition a test of acceptance for iterated probabilistic conditionals has been recently considered in [4]. The test makes essential use of a generalized version of McGee's Improved Adams' Hypothesis (see 8.4 above). The resulting view of iterated probability conditionals requires the use of Export-Import laws.

Cumulativity and Conjunctive Cumulativity have emerged as fundamental parts of the probabilistic program in semantics pioneered by Adams and later continued by McGee, Edgington and other researchers. Cumulation principles were tacitly invoked by McGee in order to explain away the fact that a rational agent can accept both a conditional of the form 'If A, then if B, then C' and its antecedent and reject its consequent. Edgington has recently offered similar arguments in [13], together with an attempt to explain away the acceptance conditions of subjunctive conditionals without appealing to imaging.

The idea of developing and defending a unified notion of supposition encoded in terms of generalized conditional probability seems to be congenial to contemporary defenders of

---

[31]Although this view is more or less widespread in the literature, we should stress here that it is far from being consensual. For example, both Adams and Levi disagree with it, although for different reasons.

Adams' program in the semantics of conditionals.[32] Here we do not intend to offer arguments pro or con this sort of unification.[33] We have focused instead on revealing the probabilistic and qualitative commitments induced by the extension of McGee's mapping to the iterated case, concluding that a strong form of Cumulativity is needed to establish this mapping. Moreover we have provided an axiomatization of the cumulative notion required (at the qualitative level), along with an interpretation of the axioms based on the work of Darwiche and Pearl. Of course, Cumulativity has not been popular in the context of belief change. Far from that, all standard accounts of belief change are incompatible with Cumulativity.

Although Cumulativity seems too strong a constraint for an inquiry-based interpretation of generalized conditional probability, a suppositional interpretation might suit it. It is less clear whether *all* forms of supposing are Cumulative. Many authors disagree with this. A more modest claim, perhaps, is that the constraint applies to 'matter of fact' forms of supposing, in Jeffrey's terminology.[34] In any case, the shared conclusion is that a general mapping between Popper functions and infinitesimal probability imposes constraints that limit the range of possible interpretations of these two different generalizations of standard conditional probability.

The next and final section is devoted to a notion of supposition explicitly proposed to study the qualitative constraints on change imposed by non-iterated Popper functions. We will verify an interesting convergence between this view and the one encoded in Hypothetical Revision.

8.1. Hájek and Harper on supposition

The theory of supposition proposed by Hájek and Harper in [18] is closely related to the purposes of the present study.[35] Hájek and Harper defend the appropriateness of extending conditional probability as a representation of conditional belief by taking full belief as primitive and defining extended conditional probability as coherent degree of belief relative to suppositions. This idea is similar to the one considered by van Fraassen, but, in the order of explanation, it places full belief and supposition before extended conditional probability. Therefore this view is more akin to the non-reductive proposal considered here. In fact, Hájek and Harper start with full belief and then introduce supposition axiomatically, as we have done in the previous sections.[36]

---

[32]This view has been independently defended by some philosophers of science like van Fraassen and Harper. The gist of the idea is to make the notion of *matter of fact* supposing involved in the acceptance of indicatives the central suppositional operation, and to explain other types of supposition in terms of added time shifts.

[33]In [4] an argument is offered in defense of the view that commitment to Export-Import laws depends on the kind of Bayesian Epistemology assumed in the analysis. Commitment to Export-Import laws is only required when the underlying epistemology is a form of *radical probabilism* whose *only primitive* is a notion of generalized conditional probability. Natural extensions of 8.4 in this setting require the Export-Import laws. But the mere use of generalized conditional probability in the context of a different Bayesian epistemology might not require these laws (see the observations in the first paragraph of section 8.1 for more details).

[34]See [21]. Nevertheless, this view is far from being uncontroversial. The issue is related to the status of the Exportation and Importation laws. Many researchers are reluctant to accept those laws, even for forms of matter-of-fact supposing.

[35]The theory is based on a previous account offered by Harper in [20].

[36]Therefore Hájek and Harper's analysis tacitly rejects the use of an underlying form of radical probabilism whose only primitive is conditional probability. They offer instead a probabilistic analysis compatible with

Their notion of change is explicitly conceived as the qualitative counterpart of the notion of supposing encoded in non-iterated Popper functions. Our notion of *hypothetical revision* is supposed to do the same job for iterated Popper functions. Therefore, it is natural to ask whether our notion of *hypothetical revision* is an extension of Hájek and Harper's notion. The lemmas below show that if Hájek and Harper's axioms are modified in one well-motivated way, they indeed are a special case of hypothetical revision.

A *basic* model of supposition is a triple $\langle \Omega, F, * \rangle$, where Ω is a non-empty set of states, F is a field of propositions with respect to Ω and $*$ is a function $* : 2^{\Omega} \times F \to 2^{\Omega}$, obeying the following constraints. Subsets K of Ω represent bodies of full belief.

(k1) $K * p \subseteq p$.

(k2) If $K \subseteq p$, then $K = K * p$.

(k3) If $p \subseteq q$ and $K * q \cap p \neq \emptyset$, then $K * p = K * q \cap p$.

(k4) If $p \subseteq q$ and $K * q = \emptyset$, then $K * p = \emptyset$.

The view of supposition encoded by the former postulates is incompatible with AGM and Ramsey revision. In fact, it satisfies Inconsistency Preservation (which is entailed by (k2)); a postulate inconsistent with both AGM and Ramsey revision. The notion does not fit the idea of supposition encoded by # either. In fact, k3 entails Preservation, while the crucial idea of # is to construct a non-Preservative notion of change.

Moreover, Hájek and Harper's postulates do not seem apt to model a well-motivated notion of supposition,[37] unless we confine our attention to non-nested conditionals. In fact, Postulates (k2) entails Invariance. But the notion proposed by Hájek and Harper is almost right, as we will see. Invariance can be circumvented by adopting an axiomatization of the type advocated by Darwiche and Pearl (and adopted in our presentation of Hypothetical revision).

(E1) $\rho(E * p) \subseteq p$.

(E2) If $\rho(E) \subseteq p$, then $\rho(E) = \rho(E * p)$.

(E3) If $p \subseteq q$ and $\rho(E * q) \cap p \neq \emptyset$, then $\rho(E * p) = \rho(E * q) \cap p$.

(E4) If $p \subseteq q$ and $\rho(E * q) = \emptyset$, then $\rho(E * p) = \emptyset$.

This presentation of the axioms advocated by Hájek and Harper is indeed entailed by the axioms of hypothetical revision.

Lemma 8.1. E2 follows from 6.2c and 6.4c.

Proof. If $\rho(E) \neq \emptyset$, k2 follows immediately from (6.2c). If $\rho(E) = \emptyset$, then $\rho(E * p) = \emptyset$, by (6.4c).

Lemma 8.2. E3 follows from 6.2c, 6.5c and 6.6c

---

the classical views of De Finetti and Savage, where a qualitative notion of full belief (certainty) is assumed as an epistemological primitive not derivable from (conditional or unconditional) probability - see [4], section 1, for details.

[37] Or to model the notion of supposition required by McGee's extension of Adams' program.

Proof. Assume $p \subseteq q$ and $\rho(E * q) \cap p \neq \emptyset$. Now, by (6.2c) $\rho(E * q) \cap p = \rho((E * q) * p)$. (6.5c) yields $\rho((E * q) * p) = \rho(E * (q \cap p))$. Now, by the hypothesis and (6.6c) we get that $\rho(E * (q \cap p)) = \rho(E * p)$.

Lemma 8.3. E4 follows from 6.5c and 6.6c.

Proof. Assume the antecedent of E4. Since $p \subseteq q$, $p = p \cap q$, (6.6c) guarantees that $\rho(E * p) = \rho(E * (p \cap q))$. By (6.5c), $\rho(E * p) = \rho(E * q) * p$. Since $\rho(E * q) = \emptyset$, then, by (6.4c), $\rho(E * q) * p = \emptyset$.

The previous lemmas show an interesting convergence of ideas between the view proposed in this article and the one offered in [18]. Hájek and Harper focused on capturing the notion of supposition presupposed by non-iterated Popper functions. We just saw that the modified version of their axioms is indeed part of our more comprehensive study of the qualitative commitments induced by Popper iterative probability kinematics. It is also interesting to see that neither theory is compatible with some of the standard theories of theory change, or with other, differently motivated, notions of supposition.[38]

---

[38]Notice also that Hájek and Harper's postulates are embeddable in different iterated extensions. As we explained above, unlike van Fraassen, they do not subscribe to a form of radical probabilism whose only primitive is conditional probability. As it is argued in [4], this leaves open the possibility of non-cumulative iterated extensions of their theory. For example, their postulates might be obeyed by other forms of matter-of-fact supposing which are not consistency preserving (like the form of consensus revision, proposed by Levi in [26], section 2.5.).

Bibliography

References

[1] Ernest W. Adams. *The Logic of Conditionals*. D. Reidel Publishing Co., Dordrecht, 1975.

[2] Horacio Arló-Costa. Belief revision conditionals: Basic iterated systems. *Annals of Pure and Applied Logic*, 96:3–28, 1999.

[3] Horacio Arló-Costa. Qualitative and probabilistic models of full belief. In S. Buss, P.Hájek, and P. Pudlák, editors, *Proceedings of Logic Colloquim'98*. Association of Symbolic Logic and A. K. Peters, 1999.

[4] Horacio Arló-Costa. Bayesian epistemology and conditionals: The role of the export-import laws. Unpublished manuscript, Philosophy Department, Carnegie Mellon University, May 2000.

[5] Horacio Arló-Costa. Hypothetical revision and matter-of-fact supposition. In *Eighth International Workshop on Non-Monotonic Reasoning (NMR'2000). Special Session*, Belief Change, Theory and Practice, 2000. Computer Research Repository, Los Alamos e-Print Archive, ACM and NCSTRL, forthcoming in *Journal of Applied Non-Classical Logic*.

[6] Horacio Arlo-Costa and Rohit Parikh. Two place probabilities, full belief and belief revision. In Paul Dekker, editor, *Proceedings of the Twelfth Amsterdam Colloquium*, pages 1–6. ILLC/Department of Philosophy, University of Amsterdam, Amsterdam, 1999.

[7] D. Balckwell and L. Dubins. On the existence and non-existence of proper, regular, conditional distributions, *The Annals of Probability*, 3, 441-752, 1975.

[8] Lawrence Blume, Adam Brandenburger, and Eddie Dekel. Lexicographic probabilities and choice under uncertainty *Econometrica* 59(1), 61-79, 1991.

[9] Michael Bratman. *Intentions, Plans and Practical Reason*. Harvard University Press, 1987.

[10] John Collins. *Belief Revision*. Ph.d. dissertation, Princeton University, Princeton University, 1991.

[11] Adnan Darwiche and Judea Pearl. On the logic of iterated belief revision. *Artificial Intelligence*, 89(1–2):1–31, 1997.

[12] Bruno De Finetti. Les probabilités nulles *Bulletins des Sceinces Mathématiques (premiere partie)* 60, 275-288, 1936.

[13] Dorothy Edgington. On conditionals. *Mind*, 104(413):235–329, 1995.

[14] Nir Friedman and Joseph Y. Halpern. Belief revision: A critique. In Luigia Carlucci Aiello, Jon Doyle, and Stuart Shapiro, editors, *KR'96: Principles of Knowledge Representation and Reasoning*, pages 421–431. Morgan Kaufmann, San Francisco, California, 1996.

[15] Peter Gärdenfors. *Knowledge in Flux*. The MIT Press, Cambridge, Massachusetts, 1988.

[16] Moisés Goldszmidt and Judea Pearl. Qualitative probabilities for default reasoning, belief revision, and causal modeling. *Artificial Intelligence*, 84(1–2):57–112, 1996.

[17] Adam Grove. Two modelings for theory change. *Journal of Philosophical Logic*, 17, (1988): pp. 157-170.

[18] Alan Hájek and Willam Harper. Full belief and probability: Comments on van Fraassen. *Dialogue*, 36, 1996.

[19] Hammond, P. J. 'Elementary Non-Archimedean representations of probability for decision theory and games,' in Humphreys (ed.), *Patrick Suppes: Scientific Philosopher*; vol 1, Kluwer, 1994.

[20] William L. Harper. Rational belief change, Popper functions and the counterfactuals. *Synthése*, 30:221–262, 1975.

[21] Richard C. Jeffrey. Matter of fact conditionals. *Aristotelian Society, Supplementary Series*, 65:161–183, 1991.

[22] James M. Joyce. *The Foundations of Causal Decision Theory*. Cambridge University Press, Cambridge, England, 1999.

[23] Henry E. Kyburg, Jr. *Probability and the Logic of Rational Belief*. Wesleyan University Press, Middletown, Connecticut, 1961.

[24] Daniel Lehmann and Menahem Magidor. What does a conditional base entails? *Technical Report TR 90-10 of the Dept. of Computer Science, Hebrew University of Jerusalem* June 1990, later published in Artificial Intelligence **55** 1-60.

[25] Isaac Levi. *The Fixation of Belief and Its Undoing: Changing Beliefs through Inquiry*. Cambridge University Press, Cambridge, England, 1991.

[26] Isaac Levi. *For the Sake of Argument: Ramsey Test Conditionals, Inductive Inference, and Nonmonotonic Reasoning*. Cambridge University Press, Cambridge, England, 1996.

[27] David K. Lewis. Probabilities of conditionals and conditional probabilities. *The Philosophical Review*, 85:297–315, 1975.

[28] Vann McGee. A counterexample to modus ponens. *Journal of Philosophy*, 82:462–471, 1985.

[29] Vann McGee. Learning the impossible. In Ellery Eells and Brian Skyrms, editors, *Probability and Conditionals: Belief Revision and Rational Decision*, pages 179–199. Cambridge University Press, Cambridge, England, 1994.

[30] Karl R. Popper. *The Logic of Scientific Discovery*. Hutchinson, London, revised edition edition, 1968.

[31] Frank P. Ramsey. General propositions and causality. In R.B. Braithwaite, editor, *The Foundations of Mathematics: Collected Papers of Frank P. Ramsey*, pages 237–255. Routledge and Kegan Paul, London, 1931.

[32] Alfred Renyi. *Foundations of Probability*. Holden Day, 1970.

[33] Abraham Robinson. *Non-standard Analysis* Revised Edition, Princeton University Press, Princeton, 1996.

[34] Seidenfeld, T. Schervish, M., and Kadane J. Improper regular conditional distributions. Manuscript CMU, October 4, 2000.

[35] Brian Skyrms. Updating, supposing and MAXENT. *Theory and Decision*, 22:225–226, 1987.

[36] Wolfgang Spohn. The representation of Popper measures *Topoi* 5 69-74, 1986.

[37] Wolfgang Spohn. A general non-probabilistic theory of inductive inference. *Causation in Decision, Belief Change and Statistics*, eds. W. Harper and B. Skyrms, Dordrecht:Reidel, 105-134, 1988.

[38] Robert C. Stalnaker. A theory of conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory*, pages 98–112. Basil Blackwell Publishers, Oxford, 1968.

[39] Bas C. van Fraassen. Representation of conditional probabilities. *Journal of Philosophical Logic*, 5(3):417–430, 1976.

[40] Bas C. van Fraassen. Erratum. *Journal of Philosophical Logic*, 6:365, 1977. Erratum to [39].

[41] Bas C. van Fraassen. Fine-grained opinion, probability, and the logic of full belief. *Journal of Philosophical Logic*, 24(4):349–377, 1995.