

Qualitative and Probabilistic Models of Full Belief

Horacio Arló Costa

Carnegie Mellon University, Pittsburgh PA 15213, USA
hcosta@andrew.cmu.edu

Summary. Let L be a language containing the modal operator B – for *full belief*. An *information model* is a set \mathbf{E} of *stable* L -theories ([19], [21], [30], [27]). A sentence is valid if it is accepted in all theories of every model. We show that a formula A is valid if and only if it is a theorem of the modal system $S5$.

The second part of the paper considers the possibility of providing probabilistic foundations for full belief. We study a proposal presented by Bas van Fraassen in [31]. Some of the results proved in [31] are generalized (especially theorem 5.2) and the dynamic aspects of the model are discussed. We show that a characterizations of probability kinematics compatible with van Fraassen's proposal is *cumulative* - inputs introduced at any stage in a series of changes are rigidly maintained through subsequent changes. Finally we consider the range of applicability of cumulative models of belief change.

1 Qualitative models of full belief

Acceptance as true and *full belief* are closely related epistemic attitudes. Commitment to full belief is mirrored by commitment to accept as true; and commitment not to accept is mirrored by commitment not to fully believe. These ideas can be rendered formally by appealing to a regimented language L containing a modal operator B . In order to use a uniform notation we can focus on the set of *consistent theories* σ of L obeying the following two constraints:

- (A1) $A \in \sigma$ iff $B(A) \in \sigma$
- (A2) $A \notin \sigma$ iff $\neg B(A) \in \sigma$

σ can be seen in this context as a *commitment set* representing the doxastic commitments of an agent at certain instant t . One can say that an agent explicitly believes a finite set of sentences M , but that he is doxastically committed to the closure of M – under the classical notion of logical consequence. Then membership in σ represents commitment to accept and lack of membership in σ represents commitment no to accept.

The account of full belief summarized above was first articulated by Isaac Levi in [20] and [21] - following pioneering ideas of De Finetti.¹ Formally similar approaches have been proposed in computer science (see [27] and [12]) and philosophy [30]. Different models of the logic of full belief have been presented recently by Van McGee in [24] and [31] (among others). All these approaches share an interest in the study of the logic of certainty. Nevertheless, some of these approaches tend to prioritize, in the order of explanation, probability over (full)

¹ De Finetti's remarks appear in [9] chapter 2.

belief while others see this reduction as problematic. In the first part of this paper we will take the notion of full belief as a primitive and we will study its logic. The second part of the paper will be devoted to study recent attempts to explicate the notion of full belief via a probabilistic reduction.

The logical behavior of theories closed under A1-2 has been recently considered independently by several researchers in different disciplines. In computer science the aforementioned theories are called *stable* theories,² and the study of such theories is considered part of *autoepistemic logic*, a branch of intensional logic introduced by Robert Moore (see [27] for a foundational introduction to the subject). Autoepistemic logic has been, nevertheless, motivated by different ideas than the ones presented above.

The main intuition in autoepistemic logic is as follows: one starts with a *premise set*, which typically is not closed. Then the idea is to capture the final states that an agent might reach by reflecting on his beliefs and by making inferences from them and about them. As Stalnaker explains in [30] these final states must meet two intuitive conditions: first they must be stable, in the sense that no further conclusions can be drawn from them; and secondly they should be *grounded* in the initial (usually not closed) premise set. A theory T over the language L is *stable* as long as it is deductively closed and it is also closed under A1-2. *Groundedness* is usually captured via a syntactic definition of a fixed point.

Most of the work in autoepistemic logic has consisted in specifying useful formal definitions of groundedness. More recently a possible world semantics has been developed (see [27] for details and [12] for the first completeness result using possible worlds semantics). For the view presented here the notion of groundedness is unimportant and the space of all stable theories plays a different role than the one that they play in autoepistemic theories. We are interested in formulating *acceptance* rules for full belief, not in specifying truth conditions for (an eventually indexical) epistemic operator. Therefore we prefer to develop a purely syntactical modelling.

According to the view defended here the stable theories represent the *informational commitments* of the agent rather than the information explicitly available for the purposes of reasoning and acting. A thorough presentation of the importance of adopting the previous interpretation of stable theories rather than the one prevalent in autoepistemic logic is presented in [3]. I will only summarize some of the features of the view of saturated theories defended here. *First*, we characterize an operator of certainty or full belief (via acceptability conditions), not a doxastic notion used to represent the uncertainty of the agent. *Second*, we are tacitly assuming that both acceptance and belief carry commitments. *Thirdly* we focus on the linguistic description of *mental acts*. We are not giving an account of *speech acts* or an account of *doxastic reports*.³

² The terminology was probably introduced by Stalnaker in [30].

³ The last distinction has several formal and conceptual consequences. For example, one can define the following notion of consequence: Γ entails A iff A is believed in any stable theory grounded on Γ . Stalnaker proposed in [30] to use such notion of consequence to

An *information model* is a set of theories of L closed under A1-2. In order to distinguish the view proposed here from the more dynamical one used in autoepistemic logic we will call such theories *saturated*. They can be seen as the linguistic representation of the commitments carried by the set of feasible states of information of a rational agent at a given instant. A1 and A2 are now seen as acceptance rules for full belief. One can then ask which are the modal formulae that are accepted in all saturated theories of a given model. Following the notation used in [20], [7] and [6], as well as [21], we can call such formulae *positively valid in a model*, and *positively valid* when positively valid in all models. Let $Th(S5)$ denote the set of theorems of the system S5 of modal logic defined as follows:

We will use the meta-letters A, B, C,... to refer to formulas of L. S5 consists of the following axioms (1) All tautologies, (2) $(B(A) \wedge B(A \rightarrow B)) \rightarrow B(A)$, (3) the alethic axiom: $B(A) \rightarrow A$, (4) the positive introspection axiom: $B(A) \rightarrow B(B(A))$, (5) the negative introspection axiom: $\neg B(A) \rightarrow B(\neg B(A))$. The rules of inference are modus ponens and the rule that guarantees that $B(A)$ is a theorem of S5 whenever A is a theorem of S5.

Now we can establish the following fact:

Theorem 1. $A \in Th(S5)$ if and only if A is positively valid.

Proof. We need to show that a formula A of L is positively valid if and only if A is a theorem of S5. From left to right the proof is not demanding. From right to left the proof is slightly harder.

We will assume that α is not a S5 theorem and we will show that α does not belong to some saturated theory T. Take $K = S5 \cup \neg B(\alpha)$. K is consistent. For suppose that K were inconsistent. Then, by axiom (3), α is an S5 theorem.

We form now a list L of all formulae of L: $g_1, g_2, \dots, g_n, \dots$. With respect to this list, we construct an infinite sequence of sets

$$I_0, I_1, \dots, I_n, \dots$$

as follows. As I_0 we take K, i.e.,

$$I_0 = K$$

Then, for each positive integer n we set:

$$I_{i+1} = \begin{cases} Cn(I_i \cup \{g_{i+1}\}) & \text{if } I_i, g_{i+1} \text{ is consistent} \\ I_i & \text{otherwise} \end{cases}$$

Form, then, a Lindenbaum-type set: $\overline{K} = C(\cup I_i)$, where $\cup I_i$ denotes the union of all the infinitely many sets I_i . \overline{K} is a complete S5-theory. Nevertheless \overline{K} is not

formalize the (Gricean) notion of (generalized) conversational implicature. A criticism of this idea is presented in [3] where a different view of the former notion of entailment is offered.

saturated. It is interesting to see why not. Consider any formula A added at stage i . At stage $i + j$ one could consistently add $\neg B(A)$ - if L is such that $B(A)$ has not been added yet. This is so because the formula $A \wedge \neg B(A)$, although epistemically problematic, is logically consistent. To put it in a different way, the modal system S5 is unable to capture syntactically the paradoxical nature of Moore's paradox ('this is my hand, but I do not fully believe it').⁴

But the formula $A \wedge \neg B(A)$ does not belong to any consistent saturated theory. For say (by contradiction) that $A \wedge \neg B(A)$ is in a consistent saturated theory T . Then $\neg B(A) \in T$ and hence (by A2) $A \notin T$ which leads to an immediate contradiction. So, the standard Henkin method used in modal logic to prove completeness does not suffice. We propose the following Moore saturation⁵ of K :

$$\mathcal{M} = \{A : B(A) \in \overline{K}\}$$

\mathcal{M} is consistent. For suppose by contradiction that $\perp \in \mathcal{M}$. Then, $B(\perp) \in \overline{K}$. But since $S5 \subseteq \overline{K}$, $\neg B(\perp) \in \overline{K}$. Therefore we have a contradiction because it is easy to check that \overline{K} is consistent. \mathcal{M} is a S5 theory. It is easy to see that all S5 theorems are in \mathcal{M} . In fact, $B(A)$ is a S5-theorem whenever A is a S5-theorem and, by construction, all S5-theorems are in K . In general, assume that A entails B . Then $B(A \rightarrow B)$ is in \overline{K} . Assume that A is in \mathcal{M} . Then $B(A) \in \overline{K}$. Therefore $B(B) \in \overline{K}$ - \overline{K} is a complete S5-theory - and this guarantees that B is in \mathcal{M} . Assume that A, B are in \mathcal{M} . Then $B(A), B(B)$ are in \overline{K} . Now the formula $B(A) \wedge B(B) \rightarrow B(A \wedge B)$ is a S5-theorem, and therefore $A \wedge B$ is in \mathcal{M} .

We will check now that \mathcal{M} is closed under A1 and A2. First we will check A1. Assume that $B(A) \in \mathcal{M}$. Assume now by contradiction that $A \notin \mathcal{M}$. Then $B(A) \notin \overline{K}$, and since \overline{K} is a complete S5 theory, $\neg B(A) \in \overline{K}$. Therefore, by negative introspection, $B(\neg B(A)) \in \overline{K}$. This, in turn, entails that $\neg B(A) \in \mathcal{M}$, against the consistency of \mathcal{M} . Assume now that $A \in \mathcal{M}$. Then $B(A) \in \overline{K}$. By positive introspection, $B(B(A)) \in \overline{K}$. This yields the desired result, namely that $B(A) \in \mathcal{M}$.

Secondly we should check A2. Assume that $A \notin \mathcal{M}$. Then $B(A) \notin \overline{K}$, and since \overline{K} is a complete S5 theory, $\neg B(A) \in \overline{K}$. Therefore, by negative introspection, $B(\neg B(A)) \in \overline{K}$. This, in turn, entails that $\neg B(A) \in \mathcal{M}$, as desired. Finally assume that $\neg B(A) \in \mathcal{M}$. Assume by contradiction that $A \in \mathcal{M}$. Then $B(A) \in \overline{K}$. Positive introspection guarantees that $B(B(A)) \in \overline{K}$. But then $B(A) \in \mathcal{M}$, against the consistency of \mathcal{M} .

So, \mathcal{M} is a consistent and saturated S5 theory. But, by construction, $\neg B(\alpha) \in \mathcal{M}$. In fact, $\neg B(\alpha) \in K \subseteq \overline{K}$. Therefore, by negative introspection, $B(\neg B(A)) \in$

⁴ The philosopher G. E. Moore offered different variants of his paradox. A canonical version appears in [26].

⁵ Notice that the Moore saturation of a set K does not need to be an *extension* of \overline{K} - although it is constructed as a function of \overline{K} . For consider the case where the initial K is paradoxical (in the sense of G.E. Moore), i.e. it contains both A and $\neg B(A)$. The Lindenbaum extension of K will contain both A and $\neg B(A)$, but the saturation of K will only contain $\neg B(A)$ - and it will *not* contain A .

\bar{K} , which is enough to guarantee that $\neg B(\alpha) \in \mathcal{M}$. But now, by A2, $\alpha \notin \mathcal{M}$. This completes the proof. \square

1.1 First person belief and belief attributions

A better understanding of the above result can be achieved by embedding the previous model in a more complicated one, where time and the perspectives of different agents are represented.

Let $I = \{1, \dots, n\}$ be a set of rational agents. The underlying language is a Boolean standard language (L_0) augmented by the modal operators $B_{i,t}$ for each i in I and for each instant t . These various operators represent the (full) beliefs of different agents at different instants.

By the same token we can index our acceptance sets. So we will have sets $\sigma_{i,t}$ representing the sentences accepted by agent i at instant t . Notice that the interpretation of the epistemic operators, when embedded in the sets $\sigma_{i,t}$ is context dependent. In fact, $B_{i,t}(\dots) \in \sigma_{j,t'}$ means that agent j is certain at time t' that agent i is certain that ... is the case at time t . $B_{i,t}(\dots) \in \sigma_{i,t'}$ with $t < t'$ holds whenever agent i is certain that he *was* certain that $B_{i,t}(\dots) \in \sigma_{i,t'}$ with $t > t'$ holds whenever agent i is certain that he *will be* certain that Finally $B_{i,t}(\dots) \in \sigma_{i,t}$ holds whenever agent i holds 'I am certain that ...' at time t .

We can focus on the set of all possible certainty sets of agent i at time t . The only constraint that we impose on these sets is that they must be closed under logical consequence and under the principles (A1-2). No other requirement is imposed. i 's background knowledge at t can be reduced to the logical consequences of a Boolean sentence, say p , or i 's certainties at t can contain several other items, including some certainties about other agents – for example i could be certain that j is certain that $\neg p$.

The fact that the modelling used in the previous section did not appeal to indexes for agents and times, is only due to the fact that (A1-2) hold in the case in which the indexes in the acceptance sets and the indexes of the modal operators coincide:

- (A1) $A \in \sigma_{i,t}$ iff $B_{i,t}(A) \in \sigma_{i,t}$
 (A2) $A \notin \sigma_{i,t}$ iff $\neg B_{i,t}(A) \in \sigma_{i,t}$

Therefore (A1-2) can be formally studied without loss of generality by dropping indexation. Nevertheless conceptual clarity about the scope of the modelling might be lost.

Consider for example the *alethic* axiom:

- (3) $B_{i,t}(A) \rightarrow A$.

The completeness result shows that (3) is positively valid, and this means that all substitution instances of (3) belong to every possible saturated set of i at t . This, of course, *does not* mean that if i is certain that A at t , then A is the case.

What this result shows is that all rational agents at time t are certain of the truth of their certainties.

A coherent agent admits a representation by $\sigma_{i,t}$ containing both $B_{j,t}(A)$ with $j < i$ and $\neg A$. That only means that i *believed* A in the past, but now he thinks A is false. But i cannot say at t , on pain of incoherence, ‘I am certain that A , but A is false’. G.E. Moore’s paradox and the epistemic validity of (3) are two sides of the same coin.⁶

The important issue uncovered by G. E. Moore’s analysis is that there are sentences that are true in some situations (and false in others), but that cannot be coherently *judged as true* in any possible situation. The negation of sentence (3) is one of such sentences.

Much more can be said about how to treat (and dissolve) some well known epistemic paradoxes in the modelling proposed above. Such an account is offered in a companion paper [3]. The purpose of this section was to provide a minimal clarification of the behavior of the so-called *alethic* axiom in our framework.

A final comment. Consider the system that can be obtained from S5 by deleting the so-called alethic condition (axiom (3) in our presentation of S5) and by replacing it with the axiom $\neg B_{i,t}(\perp)$. This system is usually called KD45 and it is considered a good axiomatization of third-person belief (not certainty or full belief).

The logic of belief attributions might very well be axiomatized by KD45. This fact does not conflict with our result. We are interested in capturing the logic of first-person claims of certainty (or full belief). On the other hand both accounts can be bridged via the following fact:

(T) A is positively valid for agent i at t iff $\vdash_{KD45,t} B_{i,t}(A)$.

In other words epistemic validity can be captured in terms of (third-person KD45) belief attributions (via T). By the same token (T) explains away paradoxical sentences (like G.E. Moore’s) as *unbelievable* sentences (in accordance with the analysis offered by Hintikka in [15]).⁷

2 Probabilistic foundations

What is full belief? One can say that it is our strongest doxastic attitude. Probabilists have considered the following (more specific) alternatives:

(1) Full belief should be identified with maximal personal probability⁸

⁶ We alluded to G.E. Moore’s epistemic paradox previously during the proof of Theorem 1. We remind the reader that the standard reference can be found in [26].

⁷ To be more precise the paradoxical formulas are exactly the ones that are either KD45 unbelievable, or the ones whose negations should be believed by a KD45 agent – see [3] for details.

⁸ A terminological clarification: the rules of the form ‘ X fully believes the propositions that he deems maximally probable’ are usually called *rules of acceptance*. This notion

(2) Full belief should be identified with probability one.

The second reductive strategy is popular in Economics (where it is usually taken for granted). But unfortunately both strategies suffer from problems first pointed out by Henry Kyburg in [16].

EXAMPLE 1 (Kyburg): Think about a lottery with so many tickets that, for any i , an agent considers maximally probable that ticket number i will not win. Then it seems that it would be reasonable for this agent to have the following expectations:⁹

(3) Each ticket is expected to lose.

(4) Some ticket wins.

Now, in the presence of the following rule we have a contradiction

(Adjunction) If agent X fully believes A at t and X fully believes B at t , then X fully believes A and B at t .

For in the presence of Adjunction, (3) entitles us to expect that:

(5) Every ticket will lose.

So, finally a self-contradictory expectation is derived: 'some ticket wins and no ticket wins'.•

The only controversial step in the derivation of the paradox is (Adjunction). But it is easy to see that (Adjunction) is a constraint on the logic of full belief as presented in the former section. In fact, if σ is the view of X at t , it is easy to see that:

(Adjunction) If $B(A) \in \sigma$ and $B(B) \in \sigma$, then $B(A \wedge B) \in \sigma$

The second proposal suffers from what has been called a transfinite version of the lottery paradox (see [23]). The following example is a measure-theoretical version of the lottery paradox (that defeats (2)).

of acceptance bridging probability and qualitative attitudes should be distinguished from the notion used in the first part of this article. One can provide 'acceptability conditions' in the qualitative manner in which we provide them in section 1, without appealing at all to the notion of probability. In order to avoid confusions, we will not use the term 'acceptance rules' to refer to high probability rules – or any other rule of that sort, like (2). Our only use of the term 'acceptance' will be the qualitative one introduced in section 1.

⁹ I remind the reader that I am using the term 'expectation' as a shorthand for 'practical certainty' or 'full belief'.

EXAMPLE 2. Say that agent a assumes that the weight of a stone is representable by a real number in some interval, say between .5 and 1 pounds. a is practically certain of the hypothesis that says that the weight lies somewhere in the interval [.5, 1]. In other words, a does not reject any of the hypothesis H_n : ‘The weight is exactly x pounds, where x is a real value between .5 and 1’. I.e. a does not fully believe $\overline{H_n}$, for any x between .5 and 1’.¹⁰ Yet a might coherently assign 0 probability to the uncountably many rival hypothesis of the form H_n .•

EXAMPLE 2 has been used by Patrick Maher (in [23], pages 388-89) in order to argue that probability one is not a sufficient condition for full belief.¹¹ This is so at least if one insists that doxastic representations should be consistently completable. Otherwise, it would not be unreasonable for a to give probability one to the hypothesis that one of the H_n is true and probability one to all the $\overline{H_n}$.

Van Fraassen concludes in a more general way that: ‘In science we deal with continuous quantities. Therefore, in general, if we let science guide our opinion, the maximally likely propositions will not form a single picture – they will just give us a family of rival maximally likely pictures.’ ([31], page 350).¹²

Is it possible to improve on the previous probabilistic proposals? Bas van Fraassen presented an interesting alternative in [31]. The remaining sections of this paper will be devoted to study van Fraassen’s proposal and the constraints that it imposes on the dynamics of belief.

¹⁰ Here we are using informally some features of the logic of full belief developed in the former sections. Say that σ is the current ‘belief set’ of a . Then the *rejection* of H_n is represented as: $\overline{H_n} \in \sigma$. Failure to reject H_n as $\overline{H_n} \notin \sigma$. But then (A2) gives us that $\neg BH_n \in \sigma$. This argument is informally carried out in [23], page 387.

¹¹ Maher argued in his piece that probability 1 is not a sufficient condition for *acceptance*, not for full belief. Maher characterizes acceptance of a hypothesis H as the state of mind expressed by the sincere intentional assertion of H . Full belief, in turn, is characterized in terms of maximal personal probability. Then, he argues that probability one is neither a sufficient nor a necessary condition for acceptance. Here we have not appealed to the quasi-pragmatical characterization of acceptance used by Maher. But we are assuming as intuitive that an agent might fail to fully believe hypothesis to which he assigns maximal probability. One therefore needs a precise characterization of this intuitive notion. The following sections of the article will be devoted to study a proposal recently developed by Bas van Fraassen. A final historical clarification. van Fraassen has argued that full acceptance does not entail full belief. But, again, here full belief is the notion characterized in terms of maximal personal probability. In the new model developed by van Fraassen full belief no longer will be characterized in this manner, and in this new context the eventual remaining differences between full acceptance and full belief seem to be minimal.

¹² Henry Kyburg’s solution of the standard paradox consists in rejecting Adjunction. He has defended his view in a series of recent papers (see, [17], [18]). [17] is particularly useful in order to contrast the account of full belief presented here and Kyburg’s notion of ‘practical certainty’. In [3] the two accounts are contrasted. So, the gist of Kyburg’s solution to his own paradox consist in abandoning the ideal that doxastic representations should be consistently completable.

3 Full belief: a probabilistic account

The following sub-section will be devoted to presenting the probabilistic reduction of full belief proposed in [31].

3.1 Basic definitions

A *space* is a couple $S = \langle U, F \rangle$ with U a non-empty set and F a sigma-field on U . A 1-place probability measure on S is characterized by Kolmogoroff's axioms.

1. $0 \leq P(A) \leq 1 = P(U)$
2. $P(A \cup B) - P(A \cap B) = P(A) + P(B)$ (*finite additivity*)
3. If $\{E_n: n = 1, 2, \dots\}$ are disjoint, with union E , then $P(E) = \sum \{P(E_n): n = 1, 2, \dots\}$ (*countable additivity*)

Conditional probabilities can be defined in the usual manner. Alternatively one can take 2-place probability functions as primitives. This can be done as follows:

A 2-place probability measure $P(- | -)$ on space S is a map $F \times F$ into real numbers such that:

I. (Reduction Axiom) The function $P(- | A)$ is either a probability measure on S or else has a constant value 1.

II. (Multiplication Axiom)

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

for all A, B, C in F .

If $P(- | A)$ is a 1-place probability measure, A will be called *normal* and *abnormal* otherwise. Let '+' mark exclusive disjunction: $A + B = (A - B) \cup (B - A)$. Two facts are worth keeping in mind: If A is normal, so are its supersets (Fact 1); and If A is abnormal, so are its subsets (Fact 2). In van Fraassen's construction the notion of the *a priori* is the opposite of the idea of abnormality:

(A) A is *a priori* for P iff $P(A | X) = 1$ for all X , iff $U - A$ is abnormal for P .

Now we can introduce the notion of *superiority*:¹³

DEFINITION. $A >_P B$ iff $P(A | A+B) = 1$

¹³ van Fraassen traces back such notion to a similar one provided by De Finetti in [10].

Armed with this notion we can now introduce the following crucial concept:
 K is a *belief core* (for P) if and only if K satisfies the following axioms:

- (A1) K is normal
- (A2) If A is a non-empty subset of K while B and K are disjoint, then $A >_P B$
- (A3) The complement $U - K$ is normal.

Van Fraassen claims in [31] that the following property follows from the conditions A1-A2-A3.

- (A4) Finesse: all non-empty subsets of K are normal

Finesse is a rather strong condition. We will rely on it later on. Therefore, since van Fraassen does not offer a proof of it in his paper, we will prove it here. This proof is of independent interest. It shows the strength of the definition of cores – and a variant of it can be used, together with a lemma proved below, to show an important fact about Kyburg’s paradox (see [3]).

Lemma 1. *A4 follows from A2 and A3.*

Proof. Assume by contradiction that there is a core S and a subset X of S such that X is abnormal. Then, by (A), \overline{X} is a priori. Therefore:

$$P(\overline{X} | X + \overline{S}) = 1, \text{ by (A) again}$$

But since S is a core A2 gives us:

$$P(X | X + \overline{S}) = 1$$

A3 gives us that \overline{S} is normal. Therefore $X + \overline{S}$ is also normal, by (Fact 1). Therefore $P(\dots | X + \overline{S})$ is a 1-place probability function. But now by finite additivity we have:

$$P(U | X + \overline{S}) = P(X | X + \overline{S}) + P(\overline{X} | X + \overline{S})$$

which violates the fact that $0 \leq P(U | X + \overline{S}) \leq 1$, by axiom 1. □

The belief cores are nested. They form a chain linearly ordered by set inclusion. This fact can be easily proved by appealing to (A2) and (A4). This is the proof offered by van Fraassen:

Say that K and K' are belief cores, then we will show that either $K \subseteq K'$ or $K' \subseteq K$. Let A be $K - K'$ and $B = K' - K$. Each set is normal if non-empty, by (A4). Say that B is non-empty. Then, since A is disjoint from K' , it follows (by

(A2)) that $B >_P A$. A similar argument shows that if A is non-empty $A >_P B$. But we cannot have both unless $A + B$ is abnormal and therefore empty (this is guaranteed by (A4) and the independent fact that supersets of normal sets are normal). So, either A or B or both are empty, yielding the desired result (namely that at least one of K or K' is a subset of the other).

Now the innermost core (if it exists) can be seen as the strongest proposition believed by the agent. In other words, fully believed propositions will (intuitively) be supersets of the innermost core. This picture suggests that the innermost core represents the ‘belief set’ of the agent, while the remaining cores can be used to guide the dynamics of this ‘belief set’ – by ordering the several possible ‘contractions’ of this belief set.

Say that an agent’s epistemic state is represented by a probability function P . Say that P induces a system of cores with non-empty intersection B . One can intuitively see B as the strongest proposition (fully) believed by the agent. We will see below that this *is* the right qualitative model induced by van Fraassen’s system. But it will take some work to show that. For the moment we cannot define the full beliefs of an agent as the propositions entailed by the innermost core of his system of cores (if the agent has cores). In fact, even if the agent has cores, such an intersection is not obviously guaranteed to exist.

So, the following more cautious definition of full belief (for a given measure P) can be used.

DEFINITION. A is a *full belief* (for P) if and only if (i) there is a belief core, and $A >_P K$ for some belief core K ; or (ii) there is no belief core and A is *a priori* (in the sense that $P(A | X) = 1$ for all X).

It can be shown that the set of propositions that are superior to a given core are exactly the propositions that are entailed by the core. So, if an agent has cores and their intersection is non-empty the former definition is equivalent to saying that the full beliefs of the agent (in this case) are the propositions entailed by its innermost core.¹⁴

3.2 Fine-grained opinion and full belief

Summing up: a belief core is a proposition K such that:

(a) K and its complement are both normal; (b) K does not leave open any abnormal alternatives; (c) any alternatives left open by K are superior to any alternatives K excludes.

And a full belief is any proposition implied by a belief core (or a priori).

¹⁴ One of the surprising results proved below is that the intersection of a non-empty system of cores always exists. Therefore, we will be able to simplify the definition of full belief once the relevant result is proved.

Van Fraassen's construction is not completely paradox-free. In [3] it is shown that the lottery paradox reappears in the case of agents devoid of belief cores. But it can be shown (see also [3] for an explicit proof) that the measure-theoretical versions of the paradox can indeed be avoided when the underlying probability function induces at least one core.¹⁵ Here we will not focus on the problem of whether van Fraassen's construction manages or not to avoid epistemic paradoxes. Our goal will be to investigate the constraints on supposition and belief change induced by the proposal.

4 Supposing and qualitative supposing

$P(\dots | \dots)$ is, according to van Fraassen, a quantitative encoding of an elusive and somewhat neglected aspect of opinion: the notion of supposition. In fact, $P(B | A)$ represents the subjective grading attributed to B upon supposing A . You might judge that $P(A | U) = 0$. For example, in EXAMPLE 2 it is coherent to say that the hypothesis stating that the exact weight is x (for some value between .5 and 1) receives probability zero. But one can *suppose* that the weight is exactly .5 and in this hypothetical scenario one should be certain that the weight is less than 1. In this section we will begin to study the constraints on the qualitative notion of supposing induced by van Fraassen's model.

First, we will prove a result that will allow us to define full belief in a simpler manner.

Lemma 2. (*Descending Chains*). *The chain of belief cores induced by a non-coreless 2-place function P cannot contain an infinitely descending chain of cores.*

Proof. Assume by contradiction that there is a 2-place P , such that the chain of belief cores induced by it contains a core B_0 and a countable chain of cores $B_0, B_1, B_2, \dots, B_n, \dots$, such that B_0 is the outermost belief core of this subsystem of cores for P . Consider in addition a set of points x_j , with j in \mathbb{N} , such that for every index i and j in \mathbb{N} , $x_i \in B_i$, and $x_i \notin B_j$, if $j > i$.

For every index m , we have (by axiom A2 of belief cores) that

$$P(\{x_{m+1}\} | \{x_{m+1}\} + \{x_m\}) = 1$$

By axiom A4 the 1-place function $P(\dots | \{x_{m+1}\} + \{x_m\})$ is a probability function obeying finite additivity. Therefore:

$$(1) P(\{x_m\} | \{x_{m+1}\} + \{x_m\}) = 0$$

¹⁵ [3] offers several arguments showing that in order to avoid the paradox one should adhere to a fully Kolmogorovian version of probability. In particular countable additivity should be assumed as a constraint on 1-place probability functions. This assumption might be controversial. Pros and cons are discussed in [3].

Now notice that the Multiplication Axiom guarantees that for all propositions A, B, C in the sigma-field F

$$(E) P(A \mid A \cup B \cup C) = 0 \text{ if } P(A \mid A \cup B) = 0.$$

To see that E is true it is enough to consider the following instance of the Multiplication Axiom:

$$\begin{aligned} P((A \cup B) \cap A \mid A \cup B \cup C) \\ = P(A \cup B \mid A \cup B \cup C) \cdot P(A \mid (A \cup B) \cap (A \cup B \cup C)) \end{aligned}$$

Call C the countable set of points $\{x_0, x_1, \dots, x_m, \dots\}$. Therefore (1) and (E) guarantee that for every x_m in U :

$$(1') P(\{x_m\} \mid C) = 0$$

Since $C \subseteq B_0$, A4 implies that $P(\dots \mid C)$ is a 1-place probability function. By the Multiplication Axiom and the fact that $P(\dots \mid C)$ is *countably additive* we have the desired contradiction: $1 = P(C \mid C) = \sum P(x_i \mid C) = 0$. \square

Remark about Finesse: Notice that Finesse *follows* from conditions A1-2-3 - see the proof above. Therefore nothing substantial has been used in the proof aside from countable additivity. In addition, nothing in the above proof requires the full force of Finesse. We only need to assume that each pair of points in U is normal.

Remarks about the Lemma of Descending Chains: van Fraassen stated in his paper some related results. For example, [T5.2] shows that the intersection of a non-empty countable family of cores is a core. But he did not prove the general result stated in our lemma, from which it follows (in the general case) that if a function P has cores, then the system of cores has a non-empty intersection. After reading a previous version of this paper, Wolfgang Spohn pointed out that he showed a similar result in Lemma 5 of [29]. Although the result resembles the one presented here, Lemma 5 in Spohn seems to be used to state a different fact - his chains of 'cores' have the structure of Lewis' system of spheres, while van Fraassen's system of cores have a different structure (for example, they are not closed under unions).¹⁶

Now we can provide an alternative definition of full belief. Consider a 2-place function P . If there are belief cores, let I be the innermost core of a system of cores

¹⁶ Some special acknowledgments concerning Lemma 2. First, thanks are due to Teddy Seidenfeld who first suggested to me in conversation the need to prove an important corollary of the lemma in order to guarantee the existence of the intersection of non-coreless probability functions. He also emphasized the role of countable additivity in the proof and gave me hints about the shape of the proof of a (differently stated) lemma. Thanks are due to Rohit Parikh who detected an error in a previous formulation of the lemma. His comments were crucial to put the result in perspective and to understand its role in the context of the larger argument offered in the paper.

induced by P . The former lemma guarantees the existence of I whenever there are cores (for P).

DEFINITION: A is *full belief* (for P) iff A is entailed by I (or A is *a priori*).

The former lemma does not rule out infinitely *ascending* chains of cores. For example, let $X = \{x_0, x_1, \dots\}$. Let $\text{rank}(x_i) = i$. Then given A, B , such that $A \subseteq B \subseteq X$, let $P(A | B) = 1$ if the lowest ranked element of B is also in A ; and let $P(A | B) = 0$ otherwise. Any set $\{x_0, \dots, x_n\}$ is a core.

In the following section we will use the previous lemma to obtain more data about the constraints on supposition and probability kinematics induced by van Fraassen's model.

5 Cumulativity

In the previous section we use the letter I to denote the innermost core of a system S of cores induced by a function P (for which there are cores). Here it will be convenient to use a more precise notation. Let IP be the innermost core corresponding to a function P (for which there are cores).

Say that the 2-place function P represents the epistemic state of the agent X at time t . Say also that P induces a system of cores with innermost core IP . Van Fraassen's account provides a model of the subjective grading that X will attribute (at t) to any item B conditional on the supposition that A . This account leaves open an important question. Which is the structure of the (qualitative) suppositional state generated by hypothetically perturbing P with A ? One response could be that an item B is *fully supposed* in the scenario obtained by hypothetically perturbing P with A if and only if $P(B | A) = 1$. But the idea that what is fully supposed after supposing A is given by what is maximally likely given A , is as problematic as the corresponding idea that a sufficient condition for full belief is maximal unconditional probability.

Here is a natural alternative suggestion for recovering full suppositions out of 2-place functions. Say that P is the current epistemic state and that this state has cores. Then IP is the state of full belief that correspond to P . The qualitative counterpart of the result of perturbing P with the supposition that A is the innermost core of $P(- | - \cap A)$ - if such function has cores. We can abbreviate $P(- | - \cap A)$ by P^A . Sequences of hypotheses are indicated as follows: $P^{A,B,\dots}$.

The previous proposal is the one tacitly used in Popper's axiomatization of 2-place functions (see, [5]). Also in [5] it is shown that the previous definition is needed in order to preserve an iterated version of the connection between infinitesimal probabilities and 2-place functions proposed by McGee in [24]. Moreover in [5] it is also shown that the notion of supposition embedded in the probabilistic semantics of conditionals requires the previous proposal for

identifying bodies of full supposition. Finally in [5] and [4] it is shown that the only available axiomatization of qualitative supposing (encoded in 2-place probability functions) [14] is also compatible with the previous proposal.

The aforementioned connection between conditional reasoning and supposition is particularly interesting. When a conditional ‘If A, then B’ is fully believed with respect to the epistemic state characterized by a non-coreless 2-place function P? An interesting idea is to check whether B is entailed by IP^A . If the answer is yes, the conditional is fully believed (or accepted) with respect to P. If the answer is no, the conditional is not accepted. The intuition underlying this test is that the conditional is accepted whenever B is fully supposed in the hypothetical state IP^A – [1].

The previous idea is not the one usually assumed in the literature on probabilistic semantics [1]. Consider, for example, Van McGee’s *Improved Adams hypothesis*, which appeals to 2-place functions. According to this variant of Adams hypothesis the probability of ‘If A, then B’ is determined by a two-place function $P(B | A)$ - see [24], page 190. But the improved hypothesis is still problematic. For say that the agent modeled in EXAMPLE 2 has his innermost core in the real interval $].5, .75[= H$. Then the improved thesis requires accepting: ‘If H, then the weight is given by an irrational number’.¹⁷ Such an unintuitive result might, nevertheless, not be required by the alternative acceptance test suggested above - as long as IP^H contains at least a rational number. This new ‘test’ for acceptance of probabilistic conditionals seems to improve on the ‘Improved Adams hypothesis’ – see, [2] for a discussion of this topic.¹⁸

The following sections will be devoted to study how bodies of full belief induced by 2-place functions evolve when perturbed by iterated suppositions. A definition is needed at this juncture:

DEFINITION: A is *sizable* (for P) iff there is core C (for P) such that $C \cap A$ is non-empty and for every core of the system of cores S (for P) $U - (C \cap A)$ is normal.

Van Fraassen’s construction entails that if both A, B and $A \cap B$, are sizable for P:

(Cumulativity) $IP^{A,B} \subseteq A$ ¹⁹

¹⁷ This requires to assume rather mild constraints on P.

¹⁸ It would be nice to know whether there are interesting probability spaces forcing the collapse of these tests. [2] offers a preliminary result showing the coincidence of the tests for countable spaces.

¹⁹ I follow here the terminology used in [5] and other recent literature on probability kinematics. Nevertheless I would like to avoid a terminological confusion by distinguishing Cumulativity from a homonymous property thoroughly investigated in non-monotonic

The following observations will be useful to prove the above claim.

Proposition 1. *If S is the system of cores of $P(-|-)$ and A is sizable for P , then all the elements of*

$$S' = \{C \cap A : C \in S \text{ and } C \cap A \neq \emptyset\}$$

are elements of the system of belief cores induced by P^A .

Proof. Let C be a core for P . Since A is sizable, there is C such that $C \cap A$ is non-empty. Consider $C \cap A = C'$. We need to show that C' is a core for P^A .

(1) C' is normal for P^A . This is guaranteed by the fact that A is sizable and by A4 – which says that all non-empty subsets of a core are normal. So

$P(-|C')$ is a 1-place probability function and $P^A(-|C')$ is also a 1-place probability function.

(2) Let S be a non-empty subset of C' . Let B be an arbitrary proposition with empty intersection with C' . We should show that S is superior (modulo P^A) to B . One should show that

$$P^A(S | (S + B)) = P(S | (S + B) \cap A) = 1.$$

first subcase: B intersects A . Then:

$$P^A(S | (S + B)) = P(S | (S + B) \cap A) = P(S | (S + (B \cap A))) = 1$$

This is so because C is a belief core for P and therefore S is superior to any proposition outside C . Therefore S is superior (for P) to $(B \cap A)$, since by construction B does not intersect C' and $(B \cap A)$ cannot intersect $C - C'$.

second subcase: B does not intersect A . Then,

$$P^A(S | (S + B)) = P(S | (S + B) \cap A) = P(S | S) = 1$$

(3) $U - C'$ is normal for P^A . This is guaranteed by the fact that A is sizable (for P). \square

The following observation shows that the qualitative effect of updating $P(-|-)$ with A is to make all $\neg A$ points unfeasible. More technically, let S be the system of cores of $P(-|-)$ and S' the system of cores of P^A . Then:

logic and the theory of theory change. The latter property can be presented as follows: If $IP^A \subseteq C$ and $IP^B \subseteq C$, then $IP^{A \wedge B} \subseteq C$. There is certain consensus about the need for a constraint of that sort in belief revision (see, for example [20] for a defense of this constraint). But, of course this property *is not* the one under investigation here. While the property that we are alluding here under the name of Cumulativity is not adopted by any of the standard axiomatizations of belief change, its homonymous counterpart is adopted by almost all of them. It is therefore of certain interest to show that the constraints on probability kinematics required by van Fraassen's construction enforce Cumulativity in the first sense.

Proposition 2. *For any core C in S' , $C \cap \neg A = \emptyset$*

Proof. Assume by contradiction that S' contains a core C such that there is a $\neg A$ -point x and $x \in C$. Since $P^A(- | x) = P(- | x \cap A) = P(- | \emptyset)$ and $P(\alpha | \emptyset) = 1$, for all α ; x is abnormal for P^A , and so is any other $\neg A$ subset of C (all such subsets are *a priori* equivalent for P^A). But, by Finesse, any subset of a core is normal. Contradiction! \square

The previous observation shows that any proposition entailing $\neg A$ is immediately non-sizable for P^A . This also shows that *all* propositions are a priori for $P^{A, \neg A}$.

Proposition 3. *If both A , B and $A \cap B$, are sizable for P (and P has cores): (Cumulativity) $IP^{A,B} \subseteq A$*

Proof. Let S be the system of cores of P . Since $A \cap B$ is sizable (for P), there is a minimal core C of S interesting $A \cap B$ (this is guaranteed by the Lemma of Descending Chains). $C \cap A$ is clearly in S' (the system of cores of P^A). This is guaranteed by Observation 1. This also guarantees that $CI = (C \cap A) \cap B$ is in S'' (the system of cores of $P^{A,B}$). Moreover CI is identical or a superset of the minimal core in S'' (because cores are nested). Of course, the existence of the minimal core in S'' is guaranteed by the Lemma of Descending Chains. Therefore,

$$IP^{A,B} \subseteq C \cap (A \cap B) \subseteq A. \quad \square$$

The role of Cumulativity (C) as a normative constraint on belief change is somewhat controversial. The standard models of (qualitative) belief change do not satisfy C – see [5] for a review of several of the possible alternatives. One of the reasons for this fact is that C introduces certain rigidity on the belief revision process. Items introduced at some point in a sequence of updates are rigidly maintained through the process. Of course this rigidity is uncontroversial when the series of inputs forms a consistent set that is, in turn, consistent with the initial state of full belief. The standard account of belief change, for example, presented by Peter Gärdenfors in [11] is compatible with this restricted form of (C). The same holds for standard conditioning. But (C) is a principle of greater generality explicitly incompatible with the account of belief change offered in [11]. (C) does allow for sequences of mutually incompatible inputs. And even in the case that they form a consistent set, they can be incompatible with the initial state of full belief.

Some authors, ([25], [4]) have argued that C is adequate to represent the process of suppositional or hypothetical change involved in the acceptance of some types of conditionals. It should be stressed, nevertheless, that many of the notions of supposition used in possible worlds semantics *are not* cumulative – see, for example, [22]. The same holds for some of the recent attempts to characterize acceptance of conditionals by appealing to belief revision techniques (see, for example [20]).

Researchers committed to probabilistic semantics (Adams or McGee are salient examples) have always defended that ‘If A, then if B, then C’ is equivalent to ‘If A and B, then C’ (Export-Import principle). Therefore they tacitly accept (C) – although the acceptance of this principle tends to be justified conceptually rather than formally (the development of models of iterate probabilistic kinematics are rather recent). (Export-Import) is, nevertheless, rejected by almost *all* other semantic systems. So, (C) is not universally accepted as an encoding of the notion of hypothetical change involved in the acceptance of conditionals.

2-place probability functions can be defined in such a way that they do not enforce cumulativity as a matter of definition.²⁰ But then certain connections between 2-place functions and non-standard probability measures are lost. In fact, Vann McGee proved in [24] a theorem establishing a natural correspondence between 2-place functions and nonstandard probability measures. Nevertheless in [5] it is shown that such correspondence holds in general only if C is enforced. So, van Fraassen’s construction possesses the features needed to bridge 2-place functions and non-standard measures, and to validate axioms universally defended in the literature of probabilistic semantics. Nevertheless, these very same features do not seem to be universal constraints on a notion of belief change with a wide application in the formalization of inquiry.

Acknowledgments: Thanks are due to Teddy Seidenfeld for many valuable comments on previous drafts. Thanks are also due to Isaac Levi, Petr Hájek and Hans Rott who commented on a previous version of this paper, read at the 1998 Logic Colloquium in Prague. Finally I would like to thank both Rohit Parikh, Bas van Fraassen and an anonymous referee. They read the final manuscript and made several useful comments.

References

1. Ernest Adams. *The logic of conditionals* (Reidel, Dordrecht, 1975).
2. Horacio Arló-Costa. ‘On the ‘Improved Adams hypothesis’,’ manuscript, Carnegie Mellon University, January, 1999.
3. Horacio Arló-Costa. ‘Practical certainty and epistemic paradox,’ manuscript, Carnegie Mellon University, first draft September 1997.
4. Horacio Arló-Costa. ‘Belief revision conditionals: Basic iterated systems,’ *Annals of Pure and Applied Logic*. forthcoming.
5. Horacio Arló-Costa and Richmond Thomason ‘Iterative probability kinematics,’ submitted for publication.
6. Horacio Arló-Costa and Isaac Levi. ‘Two notions of epistemic validity,’ *Synthese*, 109:217–262, 1996.
7. Horacio Arló-Costa. ‘Epistemic conditionals, snakes and stars’. In *Conditionals, from Philosophy to Computer Science*, vol. 5 of Studies in Logic and Computation [Series

²⁰ [5] offers such definition.

- editor: D.M. Gabbay], L. Fariñas del Cerro, G. Crocco, A. Herzig (eds.), Oxford University Press, 193-239, 1995.
8. L. Jonathan Cohen. *An essay on belief and acceptance*, OUP, 1992.
 9. Bruno De Finetti. *Theory of Probability*, Vol. 1, (eds.) A. Machi, A. Smith, John Wiley and Sons, 1990.
 10. Bruno De Finetti. 'Les probabilités nulles,' *Bulletin des Sciences Mathématiques*, 1936, 275-288.
 11. Peter Gärdenfors. *Knowledge in Flux*, Bradford Book, MIT Press, Cambridge, Mass, 1988.
 12. Melving Fitting 'Moore's nonmonotonic logic and S5,' manuscript CUNY, November 1983.
 13. Adam Grove. 'Two modellings for theory change,' *Journal of Philosophical Logic* 17:157-170.
 14. Alan Hajek and Willam Harper. Full belief and probability: Comments on van Fraassen. *Dialogue*, 36, 1996.
 15. J. Hintikka. *Knowledge and Belief: An Introduction to the Logic of the two Notions* (Ithaca, New York, Cornell University Press, 1962).
 16. Henry E., Jr. Kyburg. *Probability and the logic of rational belief*, Wesleyan University Press, Middletown, 1961.
 17. Henry E., Jr. Kyburg. 'Probabilistic inference and non-monotonic inference,' *Uncertainty in Artificial Intelligence*, R.D Schachter, T.S. Evitt, L.N. Kanal J.F. Lemmer (eds.), Elsevier Science (North Holland), 1990.
 18. Henry E., Jr. Kyburg. 'The rule of Adjunction and reasonable inference,' *Journal of Philosophy*, March 1997.
 19. Isaac Levi. 'Serious Possibility,' in *Essays in Honor of Jaakko Hintikka*, Dordrecht: Reidel, 219-36, 1979.
 20. Isaac Levi. *For the Sake of Argument: Ramsey Test Conditionals, Inductive Inference, and Nonmonotonic Reasoning*. Cambridge University Press, Cambridge, England, 1996.
 21. Isaac Levi. 'The logic of full belief,' in *The Covenant of Reason: Rationality and the Commitments of Thought*, CUP, 40-69, 1997.
 22. David Lewis. *Counterfactuals*
 23. Patrick, Maher. 'Acceptance without belief,' PSA 1990, volume 1, 381-392.
 24. Vann McGee. 'Learning the impossible,' in E. Eells and B. Skyrms (eds.) *Probability and Conditionals: Belief Revision and Rational Decision*, CUP, Cambridge, UK, 179-199, 1994.
 25. Vann McGee. 'A counterexample to modus ponens,' *Journal of Philosophy* 82, 462-471, 1985.
 26. G.E. Moore. 'Russell's 'Theory of Descriptions,' in *Philosophical Papers* by G.E. Moore, ed. H.D. Lewis (London: Muirhead Library of Philosophy, 1959) [2nd ed. 1963].
 27. Robert Moore. *Logic and Representation*, CSLI Lecture Notes No. 39, 1995.
 28. Frank.P. Ramsey. *Philosophical Papers*, H. A. Mellor, ed., (Cambridge University Press, Cambridge, 1990).
 29. Wolfgang Spohn 'The representation of Popper measures,' *Topoi* 5, 69-74, 1986.
 30. Robert Stalnaker. 'A note on non-monotonic modal logic,' *Artificial Intelligence* 64, 183-196, 1993.
 31. Bas van Fraassen. 'Fine-grained opinion, probability, and the logic of full belief,' *Journal of Philosophical Logic* 24, 349-377, 1995.