Old Evidence, Belief Revision, and Ranking Theory

Nathan M Bice

December 13, 2008

Abstract

The advantages of Bayesian Epistemology are well-known, as are its difficulties. After briefly introducing Bayesian Epistemology and making a few remarks concerning dichotomous belief, I will focus on the infamous Problem of Old Evidence. I will first examine several attempts to resolve the Problem of Old Evidence and in the process introduce both AGM Belief Revision and Ranking Theory. I will conclude by presenting a novel Ranking-Theoretic resolution of the problem.

Contents

1 Bayesian Epistemology: A Brief Introduction

Bayesian Epistemology gets off the ground with the assumption that an epistemic agent assigns degrees of belief to propositions, and in particular these degrees of belief can be captured formally as subjective probabilities. According to the Bayesian, a minimal requirement for rationality is that the agent's degrees of belief satisfy the axioms of probability theory. More specifically, the function from any proposition P to the agent's degree of belief in P should be a *probability measure*.

Definition. Let W be a nonempty set of possibilities. Let $\mathcal{A}$ be a Boolean algebra of subsets of W.[1] The elements of $\mathcal{A}$ are called *propositions*. Pr is a *probability measure* on $\mathcal{A}$ iff Pr is a function from $\mathcal{A}$ into $\mathbb{R}$ such that for all $A, B \in \mathcal{A}$:

(a) $0 \leq Pr(A) \leq 1$,

(b) $Pr(W) = 1$,

(c) if $A \cap B = \emptyset$, then $Pr(A \cup B) = Pr(A) + Pr(B)$.

This definition has several important consequences. I will mention two: $Pr(\bar{A}) = 1 - Pr(A)$ and if A entails B (in our propositional framework: if $A \subseteq B$), then $Pr(A) \leq Pr(B)$. Notice that since we are proceeding at the level of *propositions* rather than sentences in a particular language, we need not worry about logically-equivalent sentences being assigned the same probability, for (in our framework, at least) two propositions are logically equivalent exactly if they are in fact the same proposition. For instance, there is a unique tautological proposition; namely, the set of all possibilities W. So we are still assuming that the agent is *logically omniscient* in the sense that the agent treats all logically equivalent sentences as expressing the same proposition by assigning that proposition a unique subjective probability. In particular, the agent must recognize all sentences which express the tautology as expressing a proposition to be assigned a probability of 1.

[1] W can be, for example, the set of outcomes of a single toss of a six-sided die or even the set of all possible worlds. In many cases one can let $\mathcal{A}$ be the entire set of all of subsets of W, but this cannot be done in every case. However, such technical details are irrelevant for our purposes.

Furthermore, the agent must never assign a logical consequence of a proposition a lower probability than the original proposition. These assumptions are indeed controversial, but they are not the focus of this paper and we will not consider them further here.

Bayesian Epistemology also includes constraints on how agents ought to *update* their degrees of belief in light of new evidence. Before stating these constraints, we must define *conditional* probability. The conditional probability of A *given* B is defined as $Pr(A|B) = \frac{Pr(A \wedge B)}{Pr(B)}$, provided that $Pr(B) \neq 0$, and is undefined otherwise. Intuitively, $Pr(A|B)$ is the probability of A under the assumption that B. Notice in particular that $Pr(A|B) = \frac{Pr(A \wedge B)}{Pr(B)} = \frac{Pr(B|A)Pr(A)}{Pr(B)}$. This is a special case of Bayes' Theorem.

We can now state the dynamic law of updating via Strict Conditionalization: If Pr is the agent's probability measure at time t and E is the (propositional content of the) evidence that the agent receives between t and t', then the agent's new probability measure Pr' at time t' should be defined for all $A \in \mathcal{A}$ as $Pr'(A) = Pr(A|E)$.

Notice that after updating by strict conditionalization, the agent will assign a probability of 1 to the evidential proposition E (and its logical consequences). Additionally, for all propositions A such that $Pr(A) > 0$, if $Pr(E) = 1$ then $Pr(E|A) = 1$. So the agent cannot reduce the probability of E via conditionalization after initially conditionalizing on E. This approach has two obvious shortcomings. For one, updating via strict conditionalization cannot account for cases in which the evidence learned is not learned *with certainty*; i.e. it is not assigned a probability of 1 after being learned. Additionally, even if one insists that there is always *some* evidence that is learned with certainty, it isn't clear that there will always be an evidential proposition E that is both the propositional content of that evidence and is a subset of the set of possibilities W.

Jeffrey Conditionalization (Jeffrey, 1983) is a more general account of updating degrees of belief. Here is the dynamic law of updating via Jeffrey Condtionalization: If Pr is the agent's probability measure at time t, $\{E_1, ..., E_n\}$ is a partition of W, and the only direct, non-inferential effect of the evidence that the agent receives between t and t' is to shift the probabilities of $E_1, ..., E_n$ to $Pr'(E_1), ..., Pr'(E_n)$, then the agent's new probability measure Pr' at time t' should be defined for all $A \in \mathcal{A}$ as $Pr'(A) = Pr(A|E_1)Pr'(E_1) +$

$... + Pr(A|E_n)Pr'(E_n)$.

There is clearly much more that can be said about the strength and robustness of the Bayesian framework, but let's move on to some of its weaknesses.

2 Dichotomous Belief

As we saw in the last section, Bayesian Epistemology deals primarily with *degrees* of belief, which are treated formally as subjective probabilities. But in traditional epistemology, one mainly deals with *dichotomous* belief; that is, either the agent *believes* that P or the agent does not believe that P. In the latter case the agent may believe that $\bar{P}$ or suspend judgment regarding P. Can the Bayesian simply define "s believes that P" as "s assigns a probability of at least x to P", for some suitably large x? Unfortunately, this move leads to an infamous paradox.

Consider the Lottery Paradox (Kyburg, 1961). Assume, for *reductio*, s *believes* that P iff s assigns a subjective probability of at least $1 - \varepsilon$ to P, where ε is sufficiently small. Then consider a fair lottery with at least $\frac{1}{\varepsilon}$ tickets. The probability that some ticket will win is 1, and s knows this, so s believes that some ticket will win. However, for any particular ticket a, the probability that a will lose is greater than $1 - \varepsilon$, and s knows this as well, so for any particular ticket a, s believes that a will lose. Furthermore, if belief is closed under conjunction, s will have the contradictory beliefs that some ticket will win and every ticket will lose.[2] Since ε was arbitrary, this result holds for any definition of belief in terms of a subjective probability of less than 1.

The Bayesian could define "s believes that P" as "s assigns P a subjective probability of 1", but this seems unsatisfactory. For one, recall that once a proposition is assigned a probability of 1 its probability cannot be reduced via conditionalization. But surely an agent should be able to give up beliefs in light of new evidence. Furthermore, insisting

[2]Must belief be closed under conjunction? It is difficult to see how s could believe that P and believe that Q while failing to believe that $P \wedge Q$, especially since we are already assuming logical omniscience. In particular, this "resolution" of the lottery paradox would result in s having a set of beliefs that cannot all be true. In any case, we won't consider this attempted resolution here.

that beliefs are assigned a probability of 1 conflates "believing that P" with "being certain that P", and surely we do not want to insist that for any P, if s *believes* that P, then s is *certain* that P.

A staunch Bayesian might respond, "So much the worse for dichotomous belief, a vague concept which has no place in serious epistemology or science." However, just because a concept is *vague*, it doesn't follow that that concept is *unfit* for serious epistemic inquiry. Rejecting dichotomous belief outright seems like a drastic response to the lottery paradox. Indeed, such a move is radically at odds with much of traditional epistemology. Dichotomous beliefs are capable of being true or false. One can speak of the *accuracy* of degrees of belief, but not the *truth* of degrees of belief. Also, it seems that one is justified in asking when it is rational to accept a hypothesis *as true* rather than merely more probable than its negation (or an alternative hypothesis). So we have identified one weakness of the Bayesian framework: the seeming inability to account for dichotomous beliefs. However, there are more pressing problems in Bayesian Confirmation Theory, and it is there that we will turn next.

3 The Problem of Old Evidence

Confirmation has been defined in many different ways throughout the Bayesian literature. Here we will consider several definitions of an evidential proposition E *confirming* a hypothesis H:

(1) E *confirms* H exactly if learning E raises the probability of H.

(2) E *confirms* H exactly if the probability of H increases after conditionalizing on E.

(3) E *confirms* H exactly if $Pr(H|E) > Pr(H)$.

Notice that definition (1) is ambiguous between (at least) two readings:

(1a) One *does* learn E, and this *does* raise the probability of H.

(1b) If one *were* to learn E, this *would* raise the probability of H.

Notice that if E confirms H according to definition (1a), E also confirms H according to definitions (1b), (2), and (3). However, several other implications do not hold. For example, if one does not learn E, then E may confirm H according to (1b) but not (1a). Also, if $Pr(H|E) > Pr(H)$ and yet E is not (or cannot be) learned, E confirms H according to (3) but not (1a). Furthermore, if one updates by Jeffrey conditionalization, $Pr(H|E)$ may be greater than $Pr(H)$ even if the probability of H does not increase after conditionalization[3], and even if one updates by strict conditionalization, conditionalization may never actually occur. In particular, (2) is *diachronic* while (3) is *synchronic*. Hence E can confirm H according to (3) but not (2).

Let's now consider the Problem of Old Evidence (presented in Glymour, 1980). Consider the following case: the evidence captured by evidential proposition E is known prior to the formulation of theory H and H entails E. Recall that by Bayes' Theorem, $Pr(H|E) = \frac{Pr(E|H)Pr(H)}{Pr(E)}$. Since E is already known (that is, E is already a member of background information $\mathcal{K}$), $Pr(E) = 1$. Since H entails E, $Pr(E|H) = 1$. Thus $Pr(H|E) = Pr(H)$. So according to definition (3), E does not confirm H. Since E has already been learned and cannot be re-learned, the confirmation relation is undefined according to definition (1a). Similarly, since E is already an element of $\mathcal{K}$ and cannot be re-conditionalized on, the confirmation relation is undefined according to definition (2). We will consider definition (1b) later.

However, there are many such cases in which it seems that E *ought* to confirm H. Here is a famous historical example: Astronomers were aware of the difference between the actual advance of the perihelion of Mercury (E) and the advance predicted by Newtonian Physics for quite some time. Einstein eventually proved that his General Theory of Relativity (H) correctly predicted the advance of the perihelion, and most scientists considered this to be strong confirmation of H. But since E was already known and H

[3]This depends on what is meant by "conditionalizing on E", for if this simply means that E is learned with certainty and the new probability function is determined by strict conditionalization, then conditionalization on E may never occur even though $Pr(E)$ increases after observation. If this simply means that it is observed that E (although E is not necessarily learned with certainty), then $Pr(E)$ may not increase if $Pr(E)$ was initially very high. In any case conditionalization may never actually occur.

entails E, $Pr(H|E) = Pr(H)$. Hence, according to the above definitions of confirmation (except, perhaps, (1b)), either the advance of the perihelion of Mercury does not confirm the General Theory of Relativity or the confirmation relation is undefined, but this seems incorrect.

Here is one way that a Bayesian could try and escape this difficulty.[4] In our description above, we assumed without argument that if E is a member of background information $\mathcal{K}$, then $Pr(E) = 1$. But if the Bayesian were justified in asserting that $Pr(E) < 1$ *even though* E is a member of $\mathcal{K}$, then it seems that E would indeed confirm H according to definition (3) (since $Pr(H|E) = \frac{Pr(H)}{Pr(E)} > Pr(H)$ when $Pr(E) < 1$). Furthermore, a Bayesian can account for uncertain background information and uncertain learning by abandoning updating by strict conditionalization in favor of updating by Jeffrey conditionalization.

Additionally, it seems we have good reason to assign a subjective probability of less than 1 to contingent propositions such as E, even after conditionalization. For one, we learn very little (if anything) *with certainty.* I seem to see a black raven, but I may still assign a nonzero probability to the hypothesis that the object in front of me is a raven-facsimile, I am hallucinating, I am a brain in a vat, etc. Furthermore, propositions assigned a subjective probability of 1 are *unrevisable*; that is, the probability of such propositions cannot be reduced via conditionalization in light of new evidence. But there are many cases in which we may want to reduce our confidence in old evidence. For instance, if I see a black raven and then later learn that some fiendish character has been planting raven-facsimiles around my neighborhood, it seems that I ought to reduce my subjective probability that the object I saw was a black raven in light of this new evidence.[5] So perhaps we have an easy solution to the problem of old evidence.

[4] I will consider several attempts to resolve the problem of old evidence, but due to considerations of brevity I will not consider *all* such attempts. It should go without saying that this paper makes no attempt to give a comprehensive account of its subject-matter. For the purposes of this paper, I merely intend to establish that Ranking Theory can provide an intuitive resolution of this problem, and hence a lack of successful alternative resolutions would strengthen my argument but is not essential to it.

[5] A staunch proponent of strict conditionalization may retort that the evidence being learned with certainty is not that "object a is a black raven" but rather something like "I seem to see a black raven"

However, Earman (1989) points out that while this may solve what he calls the *qualitative* problem of old evidence, a *quantitative* problem remains; that is to say, E may confirm H, but the *degree* of confirmation will oftentimes be very small. In the GTR case above, perhaps scientists weren't *certain* that their measurement of the advance of the perihelion of Mercury was correct, but they at least assigned its correctness a high probability ($Pr(E) = 0.99$, say). Thus $Pr(H|E) = \frac{Pr(H)}{Pr(E)} > Pr(H)$, but the difference between $\frac{Pr(H)}{Pr(E)}$ and $Pr(H)$ will then be very small. So E confirms H, but only to a minute degree, which is again counterintuitive. So resolving the problem in this way simply creates a new problem.

However, Fitelson (unpublished) demonstrates that while the degree of confirmation will indeed be minute according to the *difference* measure, if we adopt the *likelihood-ratio* measure then the degree of confirmation can be arbitrarily large.[6] Furthermore, Fitelson (2001) has independently argued that the likelihood-ratio measure should be the preferred measure of degree of confirmation. So perhaps we can use Fitelson's result to resolve the quantitative problem of old evidence.

Unfortunately, this will not even resolve the *qualitative* problem. Consider the many controversial assumptions being made here. We are assuming that the agent updates via Jeffrey conditionalization, and in particular the agent *never* assigns a probability of 1 to a contingent proposition such as E (if some such E' were assigned probability 1, then the problem of old evidence would resurface relative to any hypothesis H such that H entails E'). We are also assuming that the correct definition of "E confirms H" just *is*

(although even establishing the certainty of *that* evidence requires further argument). However, our set of possibilities W may not be expressive enough to capture such subtle differences between propositions, and as Jeffrey (1983) points out, there may be no proposition E such that what the agent learned with certainty from her observation is that E is true. In any case, we will soon see that insisting on updating via Jeffrey conditionalization does not avoid the problem of old evidence, and hence it is unnecessary to resolve this issue here.

[6] Degree of confirmation according to the *difference* measure is defined as the difference between $Pr(H|E)$ and $Pr(H)$, while degree of confirmation according to the *likelihood-ratio* measure is defined as $\frac{Pr(E|H)}{Pr(E|\bar{H})}$. For details see Fitelson (2001).

$Pr(H|E) > Pr(H)$, since this resolution will not work under definitions (1a) or (2).[7] But what motivation could one have for defining "E confirms H" as $Pr(H|E) > Pr(H)$ apart from the fact that conditionalizing on E raises the probability of H to $Pr(H|E)$? Since we're assuming updating by Jeffrey conditionalization, $Pr(H)$ won't rise to $Pr(H|E)$ after conditionalization unless $Pr(E)$ rises to 1, and we are also assuming that $Pr(E)$ will *never* rise to 1.[8] So it seems one might need to appeal to *counterfactuals* in order to motivate definition (3) in this case, and it isn't clear how to do this without collapsing into something like definition (1b).

So the qualitative problem would not be resolved merely because $Pr(H|E) > Pr(H)$. One natural solution is to modify definition (1a) or (2) in order to account for cases of old evidence. One natural way to do this is to insist that although E has already been learned and hence it doesn't make sense to say that learning E raises the probability of H, it remains true that in the cases where intuitively E should confirm H, if E *wasn't* in our background information $\mathcal{K}$ and E *could* be learned now, then E *would* raise the probability of H. This is a form of definition (1b). So if we could *contract* E from $\mathcal{K}$, we could then determine whether E confirms H relative to $\mathcal{K}$ contracted by E. This assumes, of course, that we have a well-defined set of background information $\mathcal{K}$, a reasonable way of contracting E, and a reasonable way of assigning probabilities to E and H after contraction. These assumptions will be addressed in the next section.

4 Contraction

Howson has defended modifying the Bayesian definition of confirmation so as to include confirmation relative to $\mathcal{K}$ contracted by E as the best way to solve the problem of old evidence. According to Howson, "When you ask yourself how much support E gives H, you are plausibly asking how much a knowledge of E *would* increase the credibility of

[7] If we follow definitions (1a) or (2), since E has *already* been learned (and conditionalized on), it is *irrelevant* that $Pr(E) < 1$, for the confirmation relation will be *undefined*.

[8] In fact, in many cases it doesn't seem that $Pr(E)$ should rise at all, since one has not learned anything new about E.

H, which is the same thing as asking how much E boosts the credibility of H relative to what *else* you currently know. The 'what else' is just $\mathcal{K} - \{E\}$." (1991, pg. 548)

A natural first question is: "What exactly is $\mathcal{K}$?" $\mathcal{K}$ is a set of propositions; namely, the agent's background information. A natural suggestion is that $\mathcal{K}$ is simply the set of all propositions A such that $Pr(A) = 1$. Notice that if we define $\mathcal{K}$ in this way, $\mathcal{K}$ will automatically be closed under logical consequence (if $Pr(A) = 1$ and A entails B, then $Pr(B) = 1$). But this definition seems to have several problems. For one, background information can be given up, revised, etc. in light of new evidence, while propositions assigned probability 1 cannot (so long as we update by conditionalization). Additionally, why should we discount background information to which we have assigned a subjective probability less than 1? Most importantly, it is unclear how we can reasonably define the *unique* contraction of E from $\mathcal{K}$, since there can be many different ways of contracting E from $\mathcal{K}$. For example, if $\mathcal{K}$ contains (among other things) the propositions E, 'if A then E', 'if C then A', and C, then in order to successfully contract E from $\mathcal{K}$, we must also remove one of the other three propositions. But only one of them needs to be removed, and it isnt clear which one ought to be removed. We will return to this difficulty below.

Howson denies that $\mathcal{K}$ should be closed under logical consequence. Howson claims that this is unnecessary since any proposition A entailed by $\mathcal{K}$ will be assigned probability 1 with respect to $\mathcal{K}$ (assuming that every element of $\mathcal{K}$ is *itself* assigned probability 1). So $\mathcal{K}$ is a collection of propositions assigned probability 1, but not every proposition assigned probability 1 is an element of $\mathcal{K}$. So one should think of $\mathcal{K}$ as an independent axiomatization of the agent's background information. Then one can regard $\mathcal{K} - \{E\}$ as the simple set-theoretic removal of E from $\mathcal{K}$.

However, if A is a member of the agent's background information and A entails B, it seems odd to deny that B is a member of the agent's background information, especially since $Pr(B) \geq Pr(A)$ and hence $Pr(B) = 1$ if $Pr(A) = 1$. Furthermore, how does one construct this independent axiomatization of $\mathcal{K}$? It cannot simply be, for instance, the set of evidential propositions conditionalized upon in the past, for this would not necessarily result in a logically-independent set of propositions. For example, if one learns that object a is colored (E_1), and then later learns that object a is red (E_2), then one can still run into some of the problems mentioned above, for how can one contract E_1 without also contracting E_2 (since E_2 entails E_1 and hence otherwise E_1 will *still* be assigned probability 1 with respect to $\mathcal{K} - \{E_1\}$ since $Pr(E_2) = 1$)? Howson and Urbach (1993) admit that this solution does not apply to cases where the elements of $\mathcal{K}$ are not logically independent, but they give us no procedure for constructing the needed logically-independent set $\mathcal{K}$. So we do not yet have a satisfactory resolution.

Furthermore, it is unclear what the probabilities of E and H ought to be after contracting E from $\mathcal{K}$. Howson insists that $Pr(E)$ and $Pr(H)$ are external elements and it doesn't matter that there is no general procedure for computing them, for this is not the job of Bayesian Epistemology. Howson claims that there is no *in principle* reason for thinking that they can't be computed. Maybe so, but we are now assuming that our agent has not only assigned a subjective probability to each proposition, but has also assigned *counterfactual* probabilities to propositions based upon what probabilities she *would* assign were she to not have certain pieces of information that she does in fact have. This assumption seems problematic, to say the least.

So perhaps Howson has "solved" the problem of old evidence, but only by introducing a variety of new problems that the Bayesian will need to overcome. But there is hope, for Howson and Urbach mention that the field of Belief Revision deals explicitly with such issues as contracting propositions from sets closed under logical consequence. Perhaps the Bayesian will be able to solve the problems which plague Howson's solution by appealing to the Belief Revision literature.

Let's begin by considering a more general problem: if $\mathcal{K}$ is a set of propositions closed under logical consequence, how ought one *contract* a proposition E from $\mathcal{K}$? There will often be many different ways of contracting a proposition E, as in the example above. Alchourron, Gärdenfors, and Makinson (1985) provide a now-standard account of such contraction.[9]

[9] Since we will soon see that AGM Contraction cannot save Howson's resolution, I've ommitted most of the more technical details here.

Let's begin by defining a *belief set*: $\mathcal{K}$ is a *belief set* iff $\mathcal{K}$ is a subset of $\mathcal{A}$ such that $W \in \mathcal{K}$; $\emptyset \notin \mathcal{K}$; if $A, B \in \mathcal{K}$ then $A \cap B \in \mathcal{K}$; and if $A \in \mathcal{K}$ and $A \subseteq B \in \mathcal{K}$, then $B \in \mathcal{K}$. So intuitively $\mathcal{K}$ contains the tautology, does not contain the contradiction, and is closed under both conjunction and logical consequence.

An *AGM contraction function* $\div$ characterizes an agent's doxastic state at some time t and includes information on how to contract any non-tautological proposition from the agent's belief set $\mathcal{K}$. That is, given a non-tautological proposition E, $\div(E)$ determines a new belief set $\mathcal{K}'$. This contraction function can be modeled by an *ordering of disbelief* over the set of possibilities W; that is, one can construct a unique contraction function from such an ordering (and vice versa).

We can now state a law of simple conditionalization: If the contraction function $\div$ characterizes the doxastic state of the subject s at time t and E is the propositional content of the evidence s receives between t and t', then $\div(\bar{E}) \cap E$ is s's belief core at t'; that is, s *believes* A at t' exactly if $\div(\bar{E}) \cap E$ is a subset of A.[10]

There is a big problem with this law: the prior doxastic state is represented by a contraction function, but the posterior doxastic state is merely represented by a belief set. This violates what is known as the Principle of Categorical Matching: Prior doxastic states and posterior doxastic states should be represented in the same format. Otherwise, we seem to have no way of handling *iterated* belief changes. This is the Problem of Iterated Belief Revision. As a result, we still do not have a satisfactory account of contraction.

Can one simply use the *same* contraction function in order to continue updating one's belief set after the first contraction? Darwiche and Pearl (1997) have given several examples of how such reasoning can go horribly wrong. I reproduce one such example here:

"*Example 1* We see a strange new animal X at a distance, and it appears to be barking

[10] Although this law could be stated more elegantly in terms of a *revision* function, I've chosen to avoid introducing additional formal definitions, especially since we are mostly interested in contraction. Additionally, contraction functions and revision functions are interdefinable by the Levi and Harper Identities (ommitted here). This also applies to the example below, which refers to a revision function rather than a contraction function.

like a dog, so we conclude that X is not a bird, and that X does not fly. Still, in the event that X turns out to be a bird, we are prepared to change our mind and conclude that X flies. Observing the animal closely, we realize that it actually can fly. The question now is whether we should retain our willingness to believe that X flies in case X turns out to be a bird after all. We submit that it would be strange to give up this conditional belief merely because we happened to observe that X can fly. Yet, we provide later an AGM-compatible revision operator $\circ$ that permits such behavior."

Let's return to the problem of old evidence. Recall that we wanted to contract an evidential proposition E from a set of background information $\mathcal{K}$ closed under logical consequence. So even though we haven't found a *complete* characterization of contraction, we only need to contract a single proposition from $\mathcal{K}$ and then compute probabilities with respect to $\mathcal{K} - \{E\}$ in order to determine whether or not E confirms H. So perhaps we have a satisfactory account of contraction after all.

However, there is a serious problem with this approach. As currently formulated, $\mathcal{K}$ is a set of propositions with probability 1. So we can't construct the required ordering of disbelief over W in order to construct a unique contraction of E, since all of the elements of $\mathcal{K}$ are maximally believed![11] However, a contraction function doesn't have to be based upon an ordering of disbelief over W. Gärdenfors (1988) originally proposed an *epistemic entrenchment* ordering, where instead of propositions being ordered by plausibility, they are ordered by "usefulness in inquiry and deliberation". When forced to choose between giving up different propositions, one is supposed to give up the least-entrenched propositions. But such an entrenchment ordering will not come from an agent's probability function, and hence would have to be postulated as an additional primitive element of the agent's doxastic state. Furthermore, it isn't obvious that our contraction of E from $\mathcal{K}$ should be based upon the epistemic "usefulness" of particular propositions as opposed

[11] If one includes propositions with probability less than 1 in $\mathcal{K}$, then one will need a more precise account of what counts as background information and what does not. Also, one is then committed to updating via Jeffrey conditionalization, since contraction of E is undefined if E is maximally believed. Most importantly, this won't avoid the difficulty addressed in the following paragraph.

to their plausibility.

Even if we find a way around this difficulty, a more serious problem remains. What probabilities should be assigned to E and H after contracting E from $\mathcal{K}$? E must of course be assigned a probability of less than 1, for otherwise we have made no progress. AGM Belief Revision won't help us here, and as we discussed above, it seems counterintuitive to suppose that our agent has assigned *counterfactual* probabilities to propositions based upon what probabilities she *would* assign were she to not have certain pieces of information that she does in fact have. For the moment, this problem appears intractable.

So it seems that we need some way of *combining* an account of dichotomous belief changes with *degrees* of belief. And thus we finally turn to Ranking Theory.

5 Ranking Theory

A *negative ranking function* (originally presented as an *ordinal conditional function* in Spohn 1988) assigns a nonnegative real number or ∞ to every proposition over W, where the real numbers are meant to represent degrees of *disbelief.* More formally, κ is a *negative ranking function* for $\mathcal{A}$ iff κ is a function from $\mathcal{A}$ into $\mathbb{R}^+ = \mathbb{R} \cup \{\infty\}$ such that for all $A, B \in \mathcal{A}$:

(a) $\kappa(A) \geq 0$, $\kappa(W) = 0$, and $\kappa(\emptyset) = \infty$,

(b) $\kappa(A \cup B) = \min\{\kappa(A), \kappa(B)\}$.

Notice that either $\kappa(A) = 0$ or $\kappa(\bar{A}) = 0$ or both, since $\min\{\kappa(A), \kappa(\bar{A})\} = \kappa(A \cup \bar{A}) = \kappa(W) = 0$.

Intuitively, negative ranking functions capture both an agent's dichotomous belief set and the *firmness* of the agent's beliefs. That is, an agent *believes* that P exactly if the agent assigns some positive rank of disbelief to $\bar{P}$. The agent assigns a rank of 0 to both P and $\bar{P}$ exactly if the agent suspends judgment regarding P. One can then define the *belief set associated with a negative ranking function* κ, $\mathcal{K}(\kappa)$, as the set of all propositions

P such that $\kappa(\bar{P}) > 0$. One can easily show that this belief set is indeed a belief set as defined above; in particular it will be consistent and closed under logical consequence.

As we will soon see, negative ranking functions behave similarly to probability functions. First of all, for any $A, B \in \mathcal{A}$ such that $\kappa(A) < \infty$, the *conditional rank of B given A* is defined as $\kappa(B \mid A) = \kappa(A \cap B) - \kappa(A)$. As a consequence, $\kappa(A \cap B) = \kappa(B \mid A) + \kappa(A)$. So intuitively disbelief in A and disbelief in B given A add up to disbelief in both A and B.

We can now define the conditionalization of a ranking function: Let κ be a negative ranking function for $\mathcal{A}$ and $A \in \mathcal{A}$ such that $\kappa(A), \kappa(\bar{A}) < \infty$, and $x \in \mathbb{R}^+$. Then the *$A \to x$ conditionalization* $\kappa_{A \to x}$ of κ is defined as $\kappa_{A\to x}(B) = \min\{\kappa(B \mid A), \kappa(B \mid \bar{A}) + x\}$.

Intuitively, after conditionalization the possibilities in A are shifted downward so that $\kappa_{A\to x}(A) = 0$ and the possibilities in $\bar{A}$ are shifted upward so that $\kappa_{A\to x}(\bar{A}) = x$. So after conditionalization A is believed with firmness x. This is of course inspired by Jeffrey conditionalization. Furthermore, this process satisfies the Principle of Categorical Matching and is clearly iterable.

So we can state a dynamic law of conditionalization for ranking functions: If the prior doxastic state of the subject s at time t is characterized by the ranking function κ and if s receives evidence with propositional content E and firmness n between t and t', then the posterior state of s at t' is characterized by the $E \to n$ conditionalization of κ.

In particular, we can now define the contraction of a ranking function κ by evidential proposition E as follows: Let κ be a negative ranking function for $\mathcal{A}$ and $E \in \mathcal{A}$ such that $\kappa(\bar{E}) < \infty$. Then the *contraction $\kappa_{\div E}$ of κ by E* is defined as

$$\kappa_{\div E} = \left\{ \begin{array}{ll} \kappa, & \text{if } \kappa(\bar{E}) = 0 \\ \kappa_{E\to 0}, & \text{if } \kappa(\bar{E}) > 0 \end{array} \right\}.$$

And the *single contraction $\div_\kappa$ induced by κ* is defined as the function assigning to each $A \in \mathcal{A}$ such that $\kappa(\bar{A}) < \infty$ the belief set $\div_\kappa(A) = \mathcal{K}(\kappa_{\div A})$.

Notice that contraction is fully iterable and is defined uniquely given an agent's ranking function. Also, the single contraction induced by κ is an AGM contraction function (Spohn, unpublished).

Before returning to the problem of old evidence, let's consider an objection to ranking theory. It is quite simple. What are the agent's ranks supposed to be? If they merely order an agent's disbelief in propositions, have we really gained anything over the AGM model? That is, we have assigned real numbers to propositions rather than merely indicated their position in an ordering, but do we need any *arithmetical* properties of these real numbers? In particular, do *differences* between ranks matter, or is, for instance, ranking propositions P, Q, and R as 1, 2, and 3, no different (doxastically speaking) from ranking them as 2, 10, and 1,000?

The differences do indeed matter. In fact, Hild and Spohn (2008) have demonstrated that one can measure ranks on a ratio scale via iterated contraction functions. In particular, if an iterated contraction function behaves in an appropriate way, one can construct a difference comparison from the contraction function and use that as a difference measurement in order to construct a full ranking function.

What does "behaves in an appropriate way" mean? Hild and Spohn show that one can construct an acceptable difference comparison precisely when the iterated contraction function satisfies 6 axioms $(IC1) - (IC6)$ (stated in the appendix). Hild and Spohn also provide strong arguments that these axioms are intuitive axioms for iterated contraction. So they seem to solve two problems at once: axiomatizing iterated contraction and measuring ranks. Intuitively, differences between ranks correspond to differences in contracting propositions from a belief set. So given an agent's full belief set and information on how the agent *would* contract her belief set if she needed to remove particular propositions, provided that such contraction satisfies the axioms $(IC1) - (IC6)$, one can construct an entire ranking function. Thus, to the extent that Hild and Spohn are correct that all other attempts to deal with iterated contraction in the belief revision literature are unsatisfactory, they have provided a strong argument for both ranking theory and their particular account of iterated contraction.

There is a lacuna in Hild and Spohn's approach. Hild and Spohn claim that one can take a ranking function, construct the corresponding iterated contraction function, and then reconstruct the original ranking function uniquely up to multiplication by a constant. Recall that a ranking function determines a unique iterated contraction function. Additionally, Hild and Spohn demonstrate that an iterated contraction function satisfying $(IC1)-(IC6)$ can also induce an entire ranking function, without any need to refer to an original ranking function. However, Hild and Spohn never prove that these two methods coincide; namely, that the iterated contraction function induced by a ranking function will *itself* always satisfy $(IC1)-(IC6)$. Thankfully, this lacuna can be filled, and I have done so in the appendix.

What about the Problem of Old Evidence? Let's reconstruct the problem in a Ranking-Theoretic framework. Since H entails E, $\kappa(\bar{E} \mid H) = \infty$ (E is maximally believed given H). Since E is already contained in the belief set $\mathcal{K}$, $\kappa(\bar{E}) = y$ for some large y.

I propose that we translate Howson's intuitive but flawed resolution of the Problem of Old Evidence into our Ranking-Theoretic framework. So consider the contraction of κ by E. Since E is believed, $\kappa_{\div E}$ is simply the $E \to 0$ conditionalization of κ. In particular, H will not be believed after contraction by E ($\kappa_{E\to 0}(\bar{H}) = 0$). Furthermore, once we re-conditionalize on E with firmness n for some large n, $(\kappa_{E\to 0})_{E\to n}(\bar{H}) = \min\{\kappa_{E\to 0}(\bar{H} \mid E), \kappa_{E\to 0}(\bar{H} \mid \bar{E}) + n\} = \min\{\kappa_{E\to 0}(\bar{H} \cap E), n\} = \min\{\kappa(\bar{H} \cap E), n\}$, which will equal $\kappa(\bar{H} \cap E)$ provided that n is not too small. So H will be believed after re-conditionalization with the same degree of firmness as the agent had in the proposition '$H \vee \bar{E}$' prior to contraction (which in particular must be greater than or equal to the degree of firmness the agent had in H prior to contraction and re-conditionalization).

So in the case considered above the agent could contract E from her belief set and then re-conditionalize on E, and her degree of belief in H would indeed increase significantly once she re-conditionalized. Thus E does indeed confirm H relative to $\mathcal{K}$ contracted by E. Also, contraction is uniquely defined with respect to the initial doxastic state of the agent (there aren't multiple ways to contract E), and provided that E is conditionalized with a high degree of firmness, the resulting degree of belief in H will also be uniquely defined with respect to the initial doxastic state. Hence by adopting Ranking Theory, we can avoid *all* of the problems that plagued Howson's resolution of the Problem of Old Evidence.

6 Concluding Remarks

So what have we learned? Bayesian Epistemology seems unable to account for dichotomous belief and flounders when confronted with the problem of confirming newly-formulated hypotheses with evidence that is already known. The AGM Belief Revision literature offered some hope of resolution, but AGM contraction functions are not uniquely defined with respect to an agent's degrees of belief and furthermore can't help us formulate new probabilities for E and H. However, Ranking Theory allows us to define contraction uniquely in terms of the agent's belief set and degrees of firmness in the beliefs therein. One can then re-conditionalize on E, which will indeed result in an increase in the firmness of belief in H, and furthermore this firmness will also be *uniquely* defined provided E is itself re-conditionalized upon with a high degree of firmness.

E may not raise the probability of H when H is a newly-formulated theory and E is already known, but it will often still be true that if E wasn't already a member of our background information, then learning E *would* raise the probability of H. By abandoning talk of probability in favor of talk of dichotomous belief with differing degrees of firmness, Ranking Theory can capture this intuition successfully while it seems that Bayesian Epistemology cannot. So, coupled with the resolution of long-standing problems in AGM Belief Revision, it seems that we have a strong reason for choosing Ranking Theory as an alternative to both Bayesian Epistemology and AGM Belief Revision. There is clearly much more that can be said about this interesting topic, especially regarding the Problem of New Theories and the order in which an agent conditionalizes on a sequence of evidential propositions, but such topics are beyond the intended scope of this essay.

A Appendix

A.1 Theorem and Proof

Theorem. Let W be a nonempty set of possibilities. Let $\mathcal{A}$ be a Boolean algebra of subsets of W. Let κ be a negative ranking function for $\mathcal{A}$. Let $\mathcal{N} = \{A \in \mathcal{A} \mid \kappa(A) = \infty\}$ and $\mathcal{N}^c = \{\bar{A} \in \mathcal{A} \mid A \in \mathcal{N}\} = \{\bar{A} \in \mathcal{A} \mid \kappa(A) = \infty\}$. Let $\div_\kappa$ be the iterated contraction induced by κ. Let $\mathcal{A}_\mathcal{N}$ denote the set of all finite sequences of elements of $\mathcal{A} - \mathcal{N}^c$.

Then $\div_\kappa$ is an iterated contraction (IC) for $(\mathcal{A},\ \mathcal{N})$; i.e. $\div_\kappa$ is a potential IC for $(\mathcal{A},\ \mathcal{N})$ such that for all $A, B, C \in \mathcal{A} - \mathcal{N}^c$ and $S \in \mathcal{A}_\mathcal{N}$:

(IC1) the function $A \mapsto \div_\kappa\langle A\rangle$ is a single contraction (as specified in definition 2.10),

(IC2) if $A \notin \div_\kappa\langle\rangle$, then $\div_\kappa\langle A, S\rangle = \div_\kappa\langle S\rangle$,

(IC3) if $\bar{A} \cap \bar{B} = \emptyset$, then $\div_\kappa\langle A, B, S\rangle = \div_\kappa\langle B, A, S\rangle$,

(IC4) if $A \subseteq B$ and $A \cup \bar{B} \notin \div_\kappa\langle A\rangle$, then $\div_\kappa\langle A \cup \bar{B}, B, S\rangle = \div_\kappa\langle A, B, S\rangle$,

(IC5) if both $A \subseteq \bar{C}$ or $A, B \subseteq C$ and $A \trianglelefteq_{\div_\kappa} B$, then $A \trianglelefteq_{\div_{\kappa\langle C\rangle}} B$, and if the inequality in the antecedent is strict, that of the consequent is strict, too,

(IC6) $\div_{\kappa\langle S\rangle}$ is an IC.

Proof. $\div_\kappa : S \longmapsto \mathcal{K}(\kappa_{\div\langle S\rangle})$ is of course a potential IC for $(\mathcal{A},\ \mathcal{N})$. It remains to show that for all $A, B, C \in \mathcal{A} - \mathcal{N}^c$ and for all $S \in \mathcal{A}_\mathcal{N}$, (IC1) - (IC6) hold.

Choose $A, B, C \in \mathcal{A} - \mathcal{N}^c$ and $S \in \mathcal{A}_\mathcal{N}$. So $\kappa(\bar{A}), \kappa(\bar{B}), \kappa(\bar{C}) < \infty$. Furthermore, if $S = \langle S_1, \ldots, S_n\rangle$ for some positive $n \in \mathbb{N}$, $\kappa(\bar{S_i}) < \infty$ $(i = 1, \ldots, n)$. Otherwise, $S = \langle\rangle$.

(IC1): $\div_\kappa\langle A\rangle = \div_\kappa(A) = \mathcal{K}(\kappa_{\div A})$. So the function $A \mapsto \div_\kappa\langle A\rangle$ is a single contraction by Definition 2.14 and Corollary 2.15.[12]

[12] Wolfgang Spohn has pointed out to me in personal correspondence that this observation is already contained in (Spohn, 1988).

(IC2): Assume $A \notin \div_\kappa\langle\rangle$. So $A \notin \mathcal{K}(\kappa) = \{D \in \mathcal{A} \mid \kappa(\bar{D}) > 0\}$. So $\kappa(\bar{A}) = 0$. So $\div_\kappa\langle A, S\rangle = \mathcal{K}(\kappa_{\div\langle A,S\rangle}) = \mathcal{K}((\ldots(\kappa_{\div A})\ldots)_{\div S_n}) = \mathcal{K}((\ldots(\kappa)\ldots)_{\div S_n}) = \mathcal{K}(\kappa_{\div\langle S\rangle}) = \div_\kappa\langle S\rangle$.

(IC3): Assume $\bar{A} \cap \bar{B} = \emptyset$. So $A \cup B = W$. So $\kappa(\bar{A} \cap \bar{B}) = \kappa(\bar{A}) + \kappa(\bar{B} \mid \bar{A}) = \infty$ and $\kappa(A \cup B) = 0$.

If $A \notin \div_\kappa\langle\rangle$ or $B \notin \div_\kappa\langle\rangle$, then either $\kappa(\bar{A}) = 0$ or $\kappa(\bar{B}) = 0$, respectively. In either case we have $\div_\kappa\langle A, B, S\rangle = \mathcal{K}(\kappa_{\div\langle A,B,S\rangle}) = \mathcal{K}((\ldots((\kappa_{\div A})_{\div B})\ldots)_{\div S_n}) = \mathcal{K}((\ldots((\kappa_{\div B})_{\div A})\ldots)_{\div S_n}) = \div_\kappa\langle B, A, S\rangle$, since $(\kappa_{\div A})_{\div B} = (\kappa_{\div B})_{\div A}$.

So assume $A, B \in \div_\kappa\langle\rangle = \mathcal{K}(\kappa)$. So $\kappa(\bar{A}), \kappa(\bar{B}) > 0$ and $\kappa(A) = \kappa(B) = 0$. Also, since $\bar{A} \cap \bar{B} = \emptyset$, $\bar{A} \subseteq B$ and $\bar{B} \subseteq A$. We shall prove that $(\kappa_{\div A})_{\div B} = (\kappa_{\div B})_{\div A}$, which of course completes our proof that $\div_\kappa\langle A, B, S\rangle = \div_\kappa\langle B, A, S\rangle$.

Since $\kappa(\bar{A}) > 0$, $(\kappa_{\div A})_{\div B} = (\kappa_{A\to 0})_{\div B}$. Since $\kappa_{A\to 0}(\bar{B}) = \min\{\kappa(\bar{B} \mid A), \kappa(\bar{B} \mid \bar{A})\} = \min\{\kappa(A \cap \bar{B}), \infty\} = \kappa(A \cap \bar{B}) = \kappa(\bar{B}) > 0$, $(\kappa_{\div A})_{\div B} = (\kappa_{A\to 0})_{B\to 0}$. Similarly, $(\kappa_{\div B})_{\div A} = (\kappa_{B\to 0})_{A\to 0}$.

Choose $D \in \mathcal{A}$ such that $\kappa(\bar{D}) < \infty$. $(\kappa_{A\to 0})_{B\to 0}(D) = \min\{\kappa_{A\to 0}(D \mid B), \kappa_{A\to 0}(D \mid \bar{B})\} = \min\{\kappa_{A\to 0}(B \cap D) - \kappa_{A\to 0}(B), \kappa_{A\to 0}(\bar{B} \cap D) - \kappa_{A\to 0}(\bar{B})\} =$
$\min\Big\{\min\{\kappa(B\cap D \mid A), \kappa(B\cap D \mid \bar{A})\} - \min\{\kappa(B \mid A), \kappa(B \mid \bar{A})\}, \min\{\kappa(\bar{B}\cap D \mid A), \kappa(\bar{B}\cap D \mid \bar{A})\} - \min\{\kappa(\bar{B} \mid A), \kappa(\bar{B} \mid \bar{A})\}\Big\} = \min\Big\{\min\{\kappa(A \cap B \cap D), \kappa(D \mid \bar{A})\} - \min\{\kappa(A \mid B), 0\}, \min\{\kappa(\bar{B} \cap D), \infty\} - \min\{\kappa(\bar{B}), \infty\}\Big\} = \min\{\kappa(A \cap B \cap D), \kappa(D \mid \bar{A}), \kappa(D \mid \bar{B})\}$.
Similarly, $(\kappa_{B\to 0})_{A\to 0}(D) = \min\{\kappa(A\cap B\cap D), \kappa(D \mid \bar{A}), \kappa(D \mid \bar{B})\}$ and hence $(\kappa_{\div A})_{\div B} = (\kappa_{A\to 0})_{B\to 0} = (\kappa_{B\to 0})_{A\to 0} = (\kappa_{\div B})_{\div A}$. Thus $\div_\kappa\langle A, B, S\rangle = \div_\kappa\langle B, A, S\rangle$.

(IC4): Assume $A \subseteq B$ and $A \cup \bar{B} \notin \div_\kappa\langle A\rangle = \{D \in \mathcal{A} \mid \kappa_{\div A}(\bar{D}) > 0\}$. So $\kappa_{\div A}(\bar{A} \cap B) = 0$. So either $\kappa(\bar{A}) = 0 = \kappa(\bar{A} \cap B)$ or both $\kappa(\bar{A}) > 0$ and $\min\{\kappa(\bar{A} \cap B \mid A), \kappa(\bar{A} \cap B \mid \bar{A})\} = \min\{\infty, \kappa(B \mid \bar{A})\} = \kappa(B \mid \bar{A}) = \kappa(\bar{A}\cap B) - \kappa(\bar{A}) = 0$. In either case $\kappa(\bar{A}\cap B) = \kappa(\bar{A}) < \infty$. So $\bar{A} \cap B \neq \emptyset$ and $A \subsetneq B$.

If $A \notin \mathcal{K}(\kappa)$, then $\kappa(\bar{A}) = \kappa(\bar{A} \cap B) = 0$, so $\div_\kappa\langle A \cup \bar{B}, B, S\rangle = \mathcal{K}(\kappa_{\div\langle A\cup\bar{B},B,S\rangle}) = \mathcal{K}((\ldots((\kappa_{\div A\cup\bar{B}})_{\div B})\ldots)_{\div S_n}) = \mathcal{K}((\ldots(\kappa_{\div B})\ldots)_{\div S_n}) = \mathcal{K}((\ldots((\kappa_{\div A})_{\div B})\ldots)_{\div S_n}) =$

$\div_\kappa\langle A, B, S\rangle$.

So assume $A \in \mathcal{K}(\kappa)$. So $\kappa(\bar{A}) = \kappa(\bar{A} \cap B) > 0$. Since $A \subseteq B$, $\bar{B} \subseteq \bar{A}$ and hence $\kappa(\bar{B}) = \kappa(\bar{B} \cap \bar{A}) \geq \kappa(\bar{A}) = \kappa(\bar{A} \cap B) > 0$, since $\kappa(\bar{A}) = \min\{\kappa(\bar{A} \cap B), \kappa(\bar{A} \cap \bar{B})\}$. Thus $(\kappa_{\div A\cup\bar{B}})_{\div B} = (\kappa_{A\cup\bar{B}\to 0})_{\div B}$ and $(\kappa_{\div A})_{\div B} = (\kappa_{A\to 0})_{\div B}$.

$\kappa_{A\cup\bar{B}\to 0}(\bar{B}) = \min\{\kappa(\bar{B} \mid A \cup \bar{B}), \kappa(\bar{B} \mid \bar{A} \cap B)\} = \kappa(\bar{B} \mid A \cup \bar{B}) = \min\{\kappa(A \cap \bar{B}), \kappa(\bar{B})\} - \min\{\kappa(A), \kappa(\bar{B})\} = \min\{\infty, \kappa(\bar{B})\} - \min\{0, \kappa(\bar{B})\} = \kappa(\bar{B}) > 0$. So $(\kappa_{\div A\cup\bar{B}})_{\div B} = (\kappa_{A\cup\bar{B}\to 0})_{B\to 0}$. $\kappa_{A\to 0}(\bar{B}) = \min\{\kappa(\bar{B} \mid A), \kappa(\bar{B} \mid \bar{A})\} = \kappa(\bar{B} \mid \bar{A}) = \kappa(\bar{A} \cap \bar{B}) - \kappa(\bar{A}) = \kappa(\bar{B}) - \kappa(\bar{A})$. We will show that $(\kappa_{A\cup\bar{B}\to 0})_{B\to 0} = (\kappa_{A\to 0})_{\div B}$.

Choose $D \in \mathcal{A}$ such that $\kappa(\bar{D}) < \infty$. $(\kappa_{A\cup\bar{B}\to 0})_{B\to 0}(D) = \min\{\kappa_{A\cup\bar{B}\to 0}(D \mid B), \kappa_{A\cup\bar{B}\to 0}(D \mid \bar{B})\} = \min\{\kappa_{A\cup\bar{B}\to 0}(B \cap D) - \kappa_{A\cup\bar{B}\to 0}(B), \kappa_{A\cup\bar{B}\to 0}(\bar{B} \cap D) - \kappa_{A\cup\bar{B}\to 0}(\bar{B})\} = \min\Big\{\min\{\kappa(B \cap D \mid A \cup \bar{B}), \kappa(B \cap D \mid \bar{A} \cap B)\} - \min\{\kappa(B \mid A \cup \bar{B}), \kappa(B \mid \bar{A} \cap B)\}, \min\{\kappa(\bar{B} \cap D \mid A \cup \bar{B}), \kappa(\bar{B} \cap D \mid \bar{A} \cap B)\} - \min\{\kappa(\bar{B} \mid A \cup \bar{B}), \kappa(\bar{B} \mid \bar{A} \cap B)\}\Big\} = \min\Big\{\min\{\kappa(A \cap D), \kappa(B \cap D \mid \bar{A})\} - \min\{\kappa(A), 0\}, \min\{\kappa(\bar{B} \cap D), \infty\} - \min\{\kappa(\bar{B}), \infty\}\Big\} = \min\{\kappa(A \cap D), \kappa(B \cap D \mid \bar{A}), \kappa(D \mid \bar{B})\}$.

Since $\kappa_{A\to 0}(\bar{B}) = \kappa(\bar{B}) - \kappa(\bar{A})$ and $\kappa(\bar{B}) = \kappa(\bar{A} \cap \bar{B}) \geq \kappa(\bar{A})$, there are two possible cases:

Case 1: $\kappa(\bar{B}) > \kappa(\bar{A})$. So $(\kappa_{A\to 0})_{\div B} = (\kappa_{A\to 0})_{B\to 0}$. Now $(\kappa_{A\to 0})_{B\to 0}(D) = \min\Big\{\min\{\kappa(B \cap D \mid A), \kappa(B \cap D \mid \bar{A})\} - \min\{\kappa(B \mid A), \kappa(B \mid \bar{A})\}, \min\{\kappa(\bar{B} \cap D \mid A), \kappa(\bar{B} \cap D \mid \bar{A})\} - \min\{\kappa(\bar{B} \mid A), \kappa(\bar{B} \mid \bar{A})\}\Big\} = \min\Big\{\min\{\kappa(A \cap D), \kappa(B \cap D \mid \bar{A})\} - \min\{\kappa(A \cap B), 0\}, \min\{\infty, \kappa(\bar{B} \cap D) - \kappa(\bar{A})\} - \min\{\infty, \kappa(\bar{B}) - \kappa(\bar{A})\}\Big\} = \min\{\kappa(A \cap D), \kappa(B \cap D \mid \bar{A}), \kappa(D \mid \bar{B})\} = (\kappa_{A\cup\bar{B}\to 0})_{B\to 0}(D)$.

Case 2: $\kappa(\bar{B}) = \kappa(\bar{A})$. So $(\kappa_{A\to 0})_{\div B} = \kappa_{A\to 0}$. So $\kappa(\bar{B}) = \kappa(\bar{A}) = \kappa(\bar{A} \cap B) = \kappa(\bar{A} \cap \bar{B}) > 0$, $\kappa(A \cap \bar{B}) = \infty$, and $\kappa(A \cap B) = \kappa(A) = \kappa(B) = 0$. In this case, $(\kappa_{A\cup\bar{B}\to 0})_{B\to 0}(D) = \min\{\kappa(A \cap D), \kappa(\bar{A} \cap B \cap D) - \kappa(\bar{A}), \kappa(\bar{A} \cap \bar{B} \cap D) - \kappa(\bar{A})\} = \min\{\kappa(D \mid A), \kappa(D \mid \bar{A})\}$, since $\kappa(\bar{A} \cap D) = \min\{\kappa(\bar{A} \cap B \cap D), \kappa(\bar{A} \cap \bar{B} \cap D)\}$. So $(\kappa_{A\cup\bar{B}\to 0})_{B\to 0}(D) = \kappa_{A\to 0}(D) = (\kappa_{A\to 0})_{\div B}(D)$.

So in either case, $(\kappa_{\div A\cup\bar{B}})_{\div B} = (\kappa_{A\cup\bar{B}\to 0})_{B\to 0} = (\kappa_{A\to 0})_{\div B} = (\kappa_{\div A})_{\div B}$. Thus $\div_\kappa\langle A \cup \bar{B}, B, S\rangle = \div_\kappa\langle A, B, S\rangle$.

(IC5): Assume $A \trianglelefteq_{\div_\kappa} B$ and either $A \subseteq \bar{C}$ or $A, B \subseteq C$. So either $B \in \mathcal{N}$ or $\bar{A} \notin \div_\kappa\langle \bar{A} \cap \bar{B}\rangle$.

If $B \in \mathcal{N}$ then trivially $A \trianglelefteq_{\div_{\kappa\langle C\rangle}} B$. Also, if $B \in \mathcal{N}$ and $A \triangleleft_{\div_\kappa} B$, then $\bar{B} \in \div_\kappa\langle \bar{A} \cap \bar{B}\rangle$. So $\kappa_{\div \bar{A}\cap\bar{B}}(B) > 0$. In particular, $\kappa_{\div \bar{A}\cap\bar{B}}$ is well-defined, so $\kappa(A \cup B) < \infty$. So since $\kappa(B) = \infty$, $\kappa(A) < \infty$. Hence if $\kappa(\bar{C}) = 0$, then $(\kappa_{\div C})_{\div \bar{A}\cap\bar{B}}(B) = \kappa_{\div \bar{A}\cap\bar{B}}(B) > 0$ and hence $\bar{B} \in \div_\kappa\langle C, \bar{A} \cap \bar{B}\rangle$. If $\kappa(\bar{C}) > 0$, then $\kappa_{\div C}(B) = \kappa_{C\to 0}(B) = \min\{\kappa(B \mid C), \kappa(B \mid \bar{C})\} = \infty$, since $\infty > \kappa(\bar{C}) > 0$ and $\kappa(B) = \infty$. So either $\kappa_{C\to 0}(A \cup B) = 0$ and hence $(\kappa_{\div C})_{\div \bar{A}\cup\bar{B}}(B) = \kappa_{C\to 0}(B) = \infty$ or $\kappa_{C\to 0}(A \cup B) > 0$ and hence $(\kappa_{\div C})_{\div \bar{A}\cap\bar{B}}(B) = \min\{\kappa_{C\to 0}(B \mid \bar{A} \cap \bar{B}), \kappa_{C\to 0}(B \mid A \cup B)\} = \infty$, since both $\kappa_{C\to 0}(\bar{A} \cap \bar{B}) = 0$ and $\kappa_{C\to 0}(A \cup B) = \kappa_{C\to 0}(A) = \min\{\kappa(A \mid C), \kappa(A \mid \bar{C})\} = \min\{\kappa(A \cap C), \kappa(A \cap \bar{C}) - \kappa(\bar{C})\} < \infty$ since $\min\{\kappa(A \cap C), \kappa(A \cap \bar{C})\} = \kappa(A) < \infty$. So $\bar{B} \in \div_\kappa\langle C, \bar{A} \cap \bar{B}\rangle$. Thus $A \triangleleft_{\div_{\kappa\langle C\rangle}} B$.

So assume $B \notin \mathcal{N}$ and $\bar{A} \notin \div_\kappa\langle \bar{A} \cap \bar{B}\rangle$. So $\bar{A} \notin \mathcal{K}(\kappa_{\div\langle \bar{A}\cap\bar{B}\rangle}) = \{D \in \mathcal{A} \mid \kappa_{\div \bar{A}\cap\bar{B}}(\bar{D}) > 0\}$. So $\kappa_{\div \bar{A}\cap\bar{B}}(A) = 0$. So either $\min\{\kappa(A), \kappa(B)\} = 0 = \kappa(A)$ or $\min\{\kappa(A), \kappa(B)\} > 0$ and $\min\{\kappa(A \mid \bar{A} \cap \bar{B}), \kappa(A \mid A \cup B)\} = \kappa(A \mid A \cup B) = \min\{\kappa(A), \kappa(A \cap B)\} - \min\{\kappa(A), \kappa(B)\} = \kappa(A) - \min\{\kappa(A), \kappa(B)\} = 0$. So either $\kappa(A) = 0$ or $\kappa(A) = \kappa(B) > 0$ or $\kappa(B) > \kappa(A) > 0$. So $\kappa(A) \leq \kappa(B)$.

If $A \triangleleft_{\div_\kappa} B$, then furthermore $\bar{B} \in \div_\kappa\langle \bar{A} \cap \bar{B}\rangle$, so $\kappa_{\div \bar{A}\cap\bar{B}}(B) > 0$ and hence either $\kappa(A \cup B) = 0$ and hence $0 = \kappa(A) < \kappa(B)$ or $\kappa(A \cup B) > 0$ and hence $\kappa(A), \kappa(B) > 0$ and $\min\{\kappa(B \mid \bar{A} \cap \bar{B}), \kappa(B \mid A \cup B)\} = \min\{\kappa(B), \kappa(A \cap B)\} - \min\{\kappa(A), \kappa(B)\} = \kappa(B) - \kappa(A) > 0$. So if $A \triangleleft_{\div_\kappa} B$, then $\kappa(A) < \kappa(B)$.

It suffices to show that $\bar{A} \notin \div_\kappa\langle C, \bar{A} \cap \bar{B}\rangle = \mathcal{K}(\kappa_{\div\langle C, \bar{A}\cap\bar{B}\rangle}) = \{D \in \mathcal{A} \mid (\kappa_{\div C})_{\div \bar{A}\cap\bar{B}}(\bar{D}) > 0\}$. So it suffices to show that $(\kappa_{\div C})_{\div \bar{A}\cap\bar{B}}(A) = 0$.

If $\kappa(\bar{C}) = 0$, then $(\kappa_{\div C})_{\div \bar{A}\cap\bar{B}} = \kappa_{\div \bar{A}\cap\bar{B}}$ and $\bar{A} \notin \div_\kappa\langle \bar{A} \cap \bar{B}\rangle = \div_\kappa\langle C, \bar{A} \cap \bar{B}\rangle$. (Also, if $\kappa(\bar{C}) = 0$ and $A \triangleleft_{\div_\kappa} B$, then $\bar{B} \in \div_\kappa\langle \bar{A} \cap \bar{B}\rangle = \div_\kappa\langle C, \bar{A} \cap \bar{B}\rangle$.) So assume $\kappa(\bar{C}) > 0$. So $(\kappa_{\div C})_{\div \bar{A}\cap\bar{B}} = (\kappa_{C\to 0})_{\div \bar{A}\cap\bar{B}}$ and $\kappa(C) = 0$.

$\kappa_{C\to 0}(A) = \min\{\kappa(A \mid C), \kappa(A \mid \bar{C})\}$ and $\kappa_{C\to 0}(B) = \min\{\kappa(B \mid C), \kappa(B \mid \bar{C})\}$. So $\kappa_{C\to 0}(A \cup B) = \min\{\kappa_{C\to 0}(A), \kappa_{C\to 0}(B)\} = \min\{\kappa(A \mid C), \kappa(A \mid \bar{C}), \kappa(B \mid C), \kappa(B \mid \bar{C})\}$. Also, $(\kappa_{C\to 0})_{\bar{A}\cap\bar{B}\to 0}(A) = \min\{\kappa_{C\to 0}(A \mid \bar{A} \cap \bar{B}), \kappa_{C\to 0}(A \mid A \cup B)\} = \kappa_{C\to 0}(A \mid A \cup B) = \min\{\kappa_{C\to 0}(A), \kappa_{C\to 0}(A \cap B)\} - \kappa_{C\to 0}(A \cup B) = \kappa_{C\to 0}(A) - \kappa_{C\to 0}(A \cup B)$. Similarly,

$(\kappa_{C\to 0})_{\bar{A}\cap\bar{B}\to 0}(B) = \kappa_{C\to 0}(B) - \kappa_{C\to 0}(A\cup B)$.

Assume, for contradiction, that $(\kappa_{\div C})_{\div\bar{A}\cap\bar{B}}(A) > 0$. So either $\kappa_{C\to 0}(A\cup B) = 0$ and thus $(\kappa_{\div C})_{\div\bar{A}\cap\bar{B}}(A) = \kappa_{C\to 0}(A) > 0$ and hence $0 = \kappa_{C\to 0}(A\cup B) = \min\{\kappa(B \mid C), \kappa(B \mid \bar{C})\} < \min\{\kappa(A \mid C), \kappa(A \mid \bar{C})\}$ or $\kappa_{C\to 0}(A\cup B) > 0$ and thus $0 < (\kappa_{\div C})_{\div\bar{A}\cap\bar{B}}(A) = (\kappa_{C\to 0})_{\bar{A}\cap\bar{B}\to 0}(A) = \kappa_{C\to 0}(A) - \kappa_{C\to 0}(A\cup B)$. In either case $\min\{\kappa(B \mid C), \kappa(B \mid \bar{C})\} < \min\{\kappa(A \mid C), \kappa(A \mid \bar{C})\}$.

Case 1: $A \subseteq \bar{C}$. So $\min\{\kappa(A \mid C), \kappa(A \mid \bar{C})\} = \kappa(A \mid \bar{C}) = \kappa(A) - \kappa(\bar{C})$ and either $\kappa(B\cap C) < \kappa(A) - \kappa(\bar{C})$ or $\kappa(B\cap\bar{C}) - \kappa(\bar{C}) < \kappa(A) - \kappa(\bar{C})$. Since $\kappa(B) = \min\{\kappa(B\cap C), \kappa(B\cap\bar{C})\}$, in either case $\kappa(B) < \kappa(A)$, contradicting $\kappa(A) \leq \kappa(B)$.

Case 2: $A, B \subseteq C$. So $\kappa(B) - \kappa(C) < \kappa(A) - \kappa(C)$ and hence $\kappa(B) < \kappa(A)$, again contradicting $\kappa(A) \leq \kappa(B)$.

So $(\kappa_{\div C})_{\div\bar{A}\cap\bar{B}}(A) = 0$. So $\bar{A} \notin \div_\kappa\langle C, \bar{A}\cap\bar{B}\rangle$. Thus $A \trianglelefteq_{\div_{\kappa\langle C\rangle}} B$.

Furthermore, if $A \triangleleft_{\div_\kappa} B$ and $(\kappa_{\div C})_{\div\bar{A}\cap\bar{B}}(B) = 0$, then either $\kappa_{C\to 0}(A\cup B) = 0$ and $\kappa_{C\to 0}(B) = \min\{\kappa(B \mid C), \kappa(B \mid \bar{C})\} = 0$ or $\kappa_{C\to 0}(A\cup B) > 0$ and $(\kappa_{C\to 0})_{\bar{A}\cap\bar{B}\to 0}(B) = \kappa_{C\to 0}(B) - \kappa_{C\to 0}(A\cup B) = 0$ so $\min\{\kappa(B \mid C), \kappa(B \mid \bar{C})\} \leq \min\{\kappa(A \mid C), \kappa(A \mid \bar{C})\}$ (which of course also holds in the first case). Hence by a similar argument (namely, by substituting '$\leq$' for '$<$' in the two cases above), $\kappa(B) \leq \kappa(A)$, contradicting $\kappa(A) < \kappa(B)$. So $(\kappa_{\div C})_{\div\bar{A}\cap\bar{B}}(B) > 0$. So $\bar{B} \in \div_\kappa\langle C, \bar{A}\cap\bar{B}\rangle$. Thus $A \triangleleft_{\div_{\kappa\langle C\rangle}} B$.

(IC6): Since $\kappa_{\div_{\langle S\rangle}}$ is a negative ranking function, $\div_{\kappa\langle S\rangle}$, the iterated contraction induced by $\kappa_{\div_{\langle S\rangle}}$, must satisfy (IC1)–(IC5) (as we have just proven). Thus $\div_{\kappa\langle S\rangle}$ is an IC.

Therefore, $\div_\kappa$ is an iterated contraction (IC) for $(\mathcal{A}, \mathcal{N})$. $\square$

A.2 Definitions and Corollaries

Let W be a nonempty set of possibilities. Let $\mathcal{A}$ be a Boolean algebra of subsets of W. The elements of $\mathcal{A}$ are called *propositions*.

Definition 2.1[13]. κ is a *negative ranking function* for $\mathcal{A}$ iff κ is a function from $\mathcal{A}$ into $\mathbb{R}^+ = \mathbb{R} \cup \{\infty\}$ such that for all $A, B \in \mathcal{A}$:

(a) $\kappa(A) \geq 0$, $\kappa(W) = 0$, and $\kappa(\emptyset) = \infty$,

(b) $\kappa(A \cup B) = \min\{\kappa(A), \kappa(B)\}$.

Corollary 2.2. Either $\kappa(A) = 0$ or $\kappa(\bar{A}) = 0$ or both.

Proof. $\min\{\kappa(A), \kappa(\bar{A})\} = \kappa(A \cup \bar{A}) = \kappa(W) = 0$. □

Corollary. Let κ be a negative ranking function for $\mathcal{A}$ and $A \in \mathcal{A}$. If $B \in \mathcal{A}$, then $\kappa(B) = \min\{\kappa(A \cap B), \kappa(\bar{A} \cap B)\}$.

Proof. $\min\{\kappa(A \cap B), \kappa(\bar{A} \cap B)\} = \kappa((A \cap B) \cup (\bar{A} \cap B)) = \kappa((A \cup \bar{A}) \cap B) = \kappa(W \cap B) = \kappa(B)$. □

Definition 2.5. Let κ be a negative ranking function for $\mathcal{A}$ and $A \in \mathcal{A}$ such that $\kappa(A) < \infty$. Then, for any $B \in \mathcal{A}$, the *conditional rank of B given A* is defined as $\kappa(B \mid A) = \kappa(A \cap B) - \kappa(A)$.

Definition 2.8. Let κ be a negative ranking function for $\mathcal{A}$ and $A \in \mathcal{A}$ such that $\kappa(A), \kappa(\bar{A}) < \infty$, and $x \in \mathbb{R}^+$. Then the *$A \to x$ conditionalization* $\kappa_{A \to x}$ of κ is defined as $\kappa_{A \to x}(B) = \min\{\kappa(B \mid A), \kappa(B \mid \bar{A}) + x\}$.

$\mathcal{K}$ is a (consistent) *belief set* iff $\mathcal{K}$ is a subset of $\mathcal{A}$ such that $W \in \mathcal{K}$; $\emptyset \notin \mathcal{K}$; if $A, B \in \mathcal{K}$ then $A \cap B \in \mathcal{K}$; and if $A \in \mathcal{K}$ and $A \subseteq B \in \mathcal{K}$, then $B \in \mathcal{K}$. Let $\mathcal{F}(\mathcal{A})$ denote the set of belief sets in $\mathcal{A}$.

$\mathcal{I}$ is an *ideal* iff $\mathcal{I}$ is a subset of $\mathcal{A}$ such that $\mathcal{I}^c = \{\bar{A} \in \mathcal{A} \mid A \in \mathcal{I}\}$ is a belief set. Let $\mathcal{I}(\mathcal{A})$ denote the set of ideals in $\mathcal{A}$.

[13] My numbering of definitions and corollaries follows (Hild and Spohn, 2008).

Definition 2.10. Let $\mathcal{N} \in \mathcal{I}(\mathcal{A})$ be an ideal in $\mathcal{A}$. Then $\div$ is a *single contraction* for $\mathcal{A} - \mathcal{N}^c$ iff $\div$ is a function assigning to each proposition $A \in \mathcal{A} - \mathcal{N}^c$ a belief set $\div(A) \in \mathcal{F}(\mathcal{A})$ such that:

(a) $A \notin \div(A) \subseteq \div(\emptyset)$,

(b) if $A \notin \div(A \cap B)$, then $\div(A) \cap \div(B) \subseteq \div(A \cap B) \subseteq \div(A)$.

The *belief set associated with a negative ranking function* κ for $\mathcal{A}$ is defined as $\mathcal{K}(\kappa) = \{A \in \mathcal{A} \mid \kappa(\bar{A}) > 0\}$.

Definition 2.14. Let κ be a negative ranking function for $\mathcal{A}$ and $A \in \mathcal{A}$ such that $\kappa(\bar{A}) < \infty$. Then the *contraction* $\kappa_{\div A}$ *of* κ *by* A is defined as

$$\kappa_{\div A} = \left\{ \begin{array}{ll} \kappa, & \text{if } \kappa(\bar{A}) = 0 \\ \kappa_{A \to 0}, & \text{if } \kappa(\bar{A}) > 0 \end{array} \right\}.$$

And the *single contraction* $\div_\kappa$ *induced by* κ is defined as the function assigning to each $A \in \mathcal{A}$ such that $\kappa(\bar{A}) < \infty$ the belief set $\div_\kappa(A) = \mathcal{K}(\kappa_{\div A})$.

Corollary 2.15. $\div_\kappa$ is a single contraction for $\mathcal{A} - \mathcal{N}^c$, where $\mathcal{N} = \{A \in \mathcal{A} \mid \kappa(A) = \infty\}$.

Proof. $\div_\kappa(A) = \mathcal{K}(\kappa_{\div A}) = \{D \in \mathcal{A} \mid \kappa_{\div A}(\bar{D}) > 0\}$. Since $\kappa_{\div A}(\bar{A}) = 0$, $A \notin \div_\kappa(A)$. Assume $D \in \div_\kappa(A)$. So $\kappa_{\div A}(\bar{D}) > 0$ and hence $\kappa(\bar{D}) > 0$. So $D \in \{D \in \mathcal{A} \mid \kappa(\bar{D}) > 0\} = \{D \in \mathcal{A} \mid \kappa_{\div \emptyset}(\bar{D}) > 0\} = \div_\kappa(\emptyset)$. So $\div_\kappa(A) \subseteq \div_\kappa(\emptyset)$. So condition (a) holds.

Assume $A \notin \div_\kappa(A \cap B)$. So $\kappa_{\div A \cap B}(\bar{A}) = 0$. So either $\kappa(\bar{A}) = 0$ or $\kappa(\bar{A} \cup \bar{B}) = \min\{\kappa(\bar{A}), \kappa(\bar{B})\} > 0$ and $\min\{\kappa(\bar{A} \mid A \cap B), \kappa(\bar{A} \mid \bar{A} \cup \bar{B})\} = \kappa(\bar{A} \mid \bar{A} \cup \bar{B}) = \kappa(\bar{A}) - \min\{\kappa(\bar{A}), \kappa(\bar{B})\} = 0$. So $\kappa(\bar{A}) \leq \kappa(\bar{B})$.

Assume $C \in \div_\kappa(A) \cap \div_\kappa(B)$. So $\kappa_{\div A}(\bar{C}) > 0$ and $\kappa_{\div B}(\bar{C}) > 0$. If $\kappa(\bar{A} \cup \bar{B}) = 0$, then $\kappa_{\div A \cap B}(\bar{C}) = \kappa(\bar{C}) > 0$ and thus $C \in \div_\kappa(A \cap B)$. So assume $\kappa(\bar{A} \cup \bar{B}) = \min\{\kappa(\bar{A}), \kappa(\bar{B})\} > 0$. So $\min\{\kappa(\bar{C} \mid A), \kappa(\bar{C} \mid \bar{A})\} > 0$ and $\min\{\kappa(\bar{C} \mid B), \kappa(\bar{C} \mid \bar{B})\} > 0$. So $\kappa_{\div A \cap B}(\bar{C}) = \min\{\kappa(\bar{C} \mid A \cap B), \kappa(\bar{C} \mid \bar{A} \cup \bar{B})\} = \min\Big\{\kappa(A \cap B \cap \bar{C}), \min\{\kappa(\bar{A} \cap \bar{C}), \kappa(\bar{B} \cap \bar{C})\} - \min\{\kappa(\bar{A}), \kappa(\bar{B})\}\Big\}$. Since $\kappa(A \cap \bar{C}) > 0$, $\kappa(A \cap B \cap \bar{C}) > 0$. Thus, since $\kappa(\bar{A}) \leq \kappa(\bar{B})$, $\kappa(\bar{A} \cap \bar{C}) - \kappa(\bar{A}) > 0$, and $\kappa(\bar{B} \cap \bar{C}) - \kappa(\bar{B}) > 0$, $\kappa_{\div A \cap B}(\bar{C}) > 0$. Hence $C \in \div_\kappa(A \cap B)$. Thus $\div_\kappa(A) \cap \div_\kappa(B) \subseteq \div_\kappa(A \cap B)$.

Assume $D \in \div_\kappa(A \cap B)$. So $\kappa_{\div A \cap B}(\bar{D}) > 0$. If $\kappa(\bar{A}) = 0$, then $\kappa_{\div A}(\bar{D}) = \kappa(\bar{D}) > 0$ and hence $D \in \div_\kappa(A)$. So assume $0 < \kappa(\bar{A}) \leq \kappa(\bar{B})$. Since $\kappa_{\div A \cap B}(\bar{D}) = \min\Big\{\kappa(A \cap B \cap \bar{D}), \min\{\kappa(\bar{A} \cap \bar{D}), \kappa(\bar{B} \cap \bar{D})\} - \kappa(\bar{A})\Big\} > 0$, both $\kappa(\bar{D} \mid \bar{A}) = \kappa(\bar{A} \cap \bar{D}) - \kappa(\bar{A}) > 0$ and $\kappa(\bar{D} \mid A) = \kappa(A \cap \bar{D}) \geq \kappa(\bar{D}) > 0$. So $\kappa_{\div A}(\bar{D}) = \min\{\kappa(\bar{D} \mid A), \kappa(\bar{D} \mid \bar{A})\} > 0$ and hence $D \in \div_\kappa(A)$. Thus $\div_\kappa(A \cap B) \subseteq \div_\kappa(A)$. So condition (b) holds.

Therefore, $\div_\kappa$ is a single contraction for $\mathcal{A} - \mathcal{N}^c$. □

Definition 4.1. Let κ be a negative ranking function for $\mathcal{A}$ and $A_1, \ldots, A_n \in \mathcal{A}$ ($n \geq 0$) such that $\kappa(\bar{A}_i) < \infty$ ($i = 1, \ldots, n$). Then the *iterated contraction* $\kappa_{\div\langle A_1, \ldots, A_n\rangle}$ of κ by $\langle A_1, \ldots, A_n\rangle$ is defined as $\kappa_{\div\langle A_1, \ldots, A_n\rangle} = (\ldots(\kappa_{\div A_1})\ldots)_{\div A_n}$. The *iterated contraction* $\div_\kappa$ *induced by* κ is defined as that function which assigns to any finite sequence $\langle A_1, \ldots, A_n\rangle$ of propositions with $\kappa(\bar{A}_i) < \infty$ ($i = 1, \ldots, n$) the belief set $\div_\kappa\langle A_1, \ldots, A_n\rangle = \mathcal{K}(\kappa_{\div\langle A_1, \ldots, A_n\rangle})$. Hence, $\div_\kappa\langle\rangle = \mathcal{K}(\kappa)$.

Definition 4.3. Let $\mathcal{A}$ be an algebra of propositions over W and $\mathcal{N} \in \mathcal{I}(\mathcal{A})$ an ideal in $\mathcal{A}$. Let $\mathcal{A}_\mathcal{N}$ denote the set of all finite sequences of propositions from $\mathcal{A} - \mathcal{N}^c$. Then $\div$ is a *potential iterated contraction*, a *potential IC*, for ($\mathcal{A}$, $\mathcal{N}$) iff $\div$ is a function from the set $\mathcal{A}_\mathcal{N}$ of such finite sequences into the set $\mathcal{F}(\mathcal{A})$ of belief sets.

Definition 4.5. Let $\div$ be a potential IC for ($\mathcal{A}$, $\mathcal{N}$). Then the *potential disbelief comparision* $\trianglelefteq_\div$ *associated with* $\div$ is the binary relation on $\mathcal{A}$ such that for all $A, B \in \mathcal{A}$: $A \trianglelefteq_\div B$ iff $B \in \mathcal{N}$ or $\bar{A} \notin \div\langle\bar{A} \cap \bar{B}\rangle$. The associated *disbelief equivalence* $\triangleq_\div$ and the *strict disbelief comparison* $\triangleleft_\div$ are defined in the usual way.

Definition 5.1. Let $\mathcal{A}$ be an algebra of propositions over W and $\mathcal{N} \in \mathcal{I}(\mathcal{A})$ an ideal in $\mathcal{A}$. Let $\mathcal{A}_\mathcal{N}$ denote the set of all finite sequences of propositions from $\mathcal{A} - \mathcal{N}^c$. Then $\div$ is an *iterated contraction (IC)* for ($\mathcal{A}$, $\mathcal{N}$) iff $\div$ is a potential IC for ($\mathcal{A}$, $\mathcal{N}$) such that for all $A, B, C \in \mathcal{A} - \mathcal{N}^c$ and $S \in \mathcal{A}_\mathcal{N}$:

(IC1) the function $A \mapsto \div\langle A\rangle$ is a single contraction (as specified in Definition 2.10),

(IC2) if $A \notin \div\langle\rangle$, then $\div\langle A, S\rangle = \div\langle S\rangle$,

(IC3) if $\bar{A} \cap \bar{B} = \emptyset$, then $\div\langle A, B, S\rangle = \div\langle B, A, S\rangle$,

(IC4) if $A \subseteq B$ and $A \cup \bar{B} \notin \div\langle A\rangle$, then $\div\langle A \cup \bar{B}, B, S\rangle = \div\langle A, B, S\rangle$,

(IC5) if both $A \subseteq \bar{C}$ or $A, B \subseteq C$ and $A \trianglelefteq_{\div} B$, then $A \trianglelefteq_{\div_{\langle C\rangle}} B$, and if the inequality in the antecedent is strict, that of the consequent is strict, too,

(IC6) $\div_{\langle S\rangle}$ is an IC.

$\div_{\langle S\rangle}$ in (IC6) denotes the function assigning the value $\div\langle S, S'\rangle$ to each seqence S' in $\mathcal{A}_{\mathcal{N}}$. $\div_{\langle C\rangle}$ in (IC5) is similarly defined.

References

[1] Alchourron, Carlos E., Peter Gärdenfors, David Makinson (1985), "On the Logic of Theory Change: Partial Meet Contraction and Revision Functions", *Journal of Symbolic Logic* 50, 510-530.

[2] Darwiche, Adnan, Judea Pearl (1997), "On the Logic of Iterated Belief Revision", *Artificial Intelligence* 89, 1-29.

[3] Earman, John (1989), "Old Evidence, New Theories: Two Unresolved Problems in Bayesian Confirmation Theory", *Pacific Philosophical Quarterly* 70, 323-340.

[4] Fitelson, Branden (2001). *Studies in Bayesian Confirmation Theory*, Ph.D. Dissertation, University of Wisconsin.

[5] Fitelson, Branden (unpublished). "Earman on Old Evidence and Measures of Confirmation", manuscript.

[6] Gärdenfors, Peter (1988), *Knowledge in Flux. Modeling the Dynamics of Epistemic States*. Cambridge, MA: MIT Press.

[7] Glymour, Clark (1980), *Theory and Evidence*. Princeton: Princeton University Press.

[8] Hild, Matthias, Wolfgang Spohn (2008), "The Measurement of Ranks and the Laws of Iterated Contraction". *Artificial Intelligence*, 172: 1195-1218.

[9] Howson, Colin (1991), "The 'Old Evidence' Problem", *British Journal for the Philosophy of Science*, 42: 547-555.

[10] Howson, Colin, Peter Urbach (1993), *Scientific Reasoning: The Bayesian Approach*, 2nd ed. Chicago: Open Court.

[11] Jeffrey, Richard (1983), *The Logic of Decision*. 2nd ed. Chicago: University of Chicago Press.

[12] Kyburg, Henry E. Jr. (1961), *Probability and the Logic of Rational Belief.* Middletown, CT: Wesleyan University Press.

[13] Spohn, Wolfgang (1988), "Ordinal Conditional Functions. A Dynamic Theory of Epistemic States", in: W. L. Harper, B. Skyrms (eds.), *Causation in Decision, Belief Change, and Statistics, vol. II*, Dordrecht: Kluwer, pp. 105-134.

[14] Spohn, Wolfgang (unpublished), *Ranking Theory: A Tool for Epistemology.* manuscript.