Journal of Artificial Intelligence Research 26 (2006) 127-151 Submitted 8/05; published 6/06

Admissible and Restrained Revision

Richard Booth RICHARD.B@MSU.AC.TH
Faculty of Informatics
Mahasarakham University
Mahasarakham 44150, Thailand

Thomas Meyer THOMAS.MEYER@NICTA.COM.AU
National ICT Australia and
University of New South Wales
223 Anzac Parade
Kensington, NSW 2052, Australia

Abstract

As partial justification of their framework for iterated belief revision Darwiche and Pearl convincingly argued against Boutilier's natural revision and provided a prototypical revision operator that fits into their scheme. We show that the Darwiche-Pearl arguments lead naturally to the acceptance of a smaller class of operators which we refer to as *admissible.* Admissible revision ensures that the penultimate input is not ignored completely, thereby eliminating natural revision, but includes the Darwiche-Pearl operator, Nayak's lexicographic revision operator, and a newly introduced operator called *restrained revision.* We demonstrate that restrained revision is the most conservative of admissible revision operators, effecting as few changes as possible, while lexicographic revision is the least conservative, and point out that restrained revision can also be viewed as a composite operator, consisting of natural revision preceded by an application of a "backwards revision" operator previously studied by Papini. Finally, we propose the establishment of a principled approach for choosing an appropriate revision operator in different contexts and discuss future work.

1. Introduction

The ability to rationally change one's knowledge base in the face of new information which possibly contradicts the currently held beliefs is a basic characteristic of intelligent behaviour. Thus the question of *belief revision* is of crucial importance in Artificial Intelligence. In the last twenty years this question has received considerable attention, starting from the work of Alchourrón, Gärdenfors, and Makinson (1985) – usually abbreviated to just AGM – who proposed a set of rationality postulates which any reasonable revision operator should satisfy. A semantic construction of revision operators was later provided by Katsuno and Mendelzon (1991), according to which an agent has in its mind a plausibility ordering – a total preorder – over the set of possible worlds, with the knowledge base associated to this ordering being identified with the set of sentences true in all the most plausible worlds. This approach dates back to the work of Lewis (1973) on counterfactuals. It was introduced into the belief revision literature by Grove (1988) and Spohn (1988). Given a new sentence – or *epistemic input* – α, the revised knowledge base is set as the set of sentences true in all the most plausible worlds in which α holds. As was shown by Katsuno and Mendelzon (1991), the family of operators defined by this construction coincides exactly with the family of operators satisfying the AGM postulates. Due to its intuitive appeal, this construction came to be widely used in the area. However, researchers soon began to notice a deficiency with it – although it prescribes how to obtain a new *knowledge base*, it remains silent on how to obtain a new *plausibility ordering* which can then serve as a target for the next epistemic input. Thus it is not rich enough to deal adequately with the problem of *iterated* belief revision. This paper is a contribution to the study of this problem.

Most iterated revision schemes are sensitive to the history of belief changes[1], based on a version of the "most recent is best" argument, where the newest information is of higher priority than anything else in the knowledge base. Arguably the most extreme case of this is Nayak's lexicographic revision (Nayak, 1994; Nayak, Pagnucco, & Peppas, 2003). However, there are operators where, once *admitted* to the knowledge base, it rapidly becomes as much of a candidate for removal as anything else in the set when another, newer, piece of information comes along, Boutilier's natural revision (1993, 1996) being a case in point. A dual to this is what Rott (2003) terms *radical* revision where the new information is accepted with maximal, irremediable entrenchment – see also Segerberg (1998). Another issue to consider is the problem termed *temporal incoherence* (Rott, 2003):

> the comparative recency of information should translate systematically into comparative importance, strength or entrenchment

In an influential paper Darwiche and Pearl (1997) proposed a framework for iterated revision. Their proposal is characterised in terms of sets of syntactic and semantic postulates, but can also be viewed from the perspective of conditional beliefs. It is an extension of the formulation by Katsuno and Mendelzon (1991) of AGM revision (Alchourrón et al., 1985). To justify their proposal Darwiche and Pearl mount a comprehensive argument. The argument includes a critique of natural revision, which is shown to admit too few changes. In addition, they provide a concrete revision operator which is shown to satisfy their postulates. In many ways this can be seen as the prototypical Darwiche-Pearl operator. It is instructive to observe that the two best-known operators satisfying the Darwiche-Pearl postulates, natural revision and lexicographic revision, form the opposite extremes of the Darwiche-Pearl framework: Natural revision is the most conservative Darwiche-Pearl operator, in the sense that it effects as few changes as possible, while lexicographic revision is the least conservative.

In this paper we show that the Darwiche-Pearl arguments lead naturally to the acceptance of a smaller class of operators which we refer to as *admissible*. We provide characterisations of admissible revision, in terms of syntactic as well as semantic postulates. Admissible revision ensures that the penultimate input is not ignored completely. A consequence of this is that natural revision is eliminated. On the other hand, admissible revision includes the prototypical Darwiche-Pearl operator as well as lexicographic revision, the latter result also showing that lexicographic revision is the least conservative of the admissible operators. The removal of natural revision from the scene leaves a gap which is filled by the introduction of a new operator we refer to as *restrained revision*. It is the most conservative of admissible revision operators, and can thus be seen as an appropriate replacement of natural revision. We give a syntactic and a semantic characterisation of restrained revision,

1. An *external* revision scheme like that of Areces and Becher (2001) and Freund and Lehmann (1994) is not.

and demonstrate that it satisfies desirable properties. In particular, and unlike lexicographic revision, it ensures that older information is not discarded unnecessarily, and it shows that the problem of temporal incoherence can be dealt with.

Although natural revision does not feature in the class of admissible revision operators, we show that it still has a role to play in iterated revision, provided it is first tempered appropriately. We show that restrained revision can also be viewed as a composite operator, consisting of natural revision preceded by an application of a "backwards revision" operator previously studied by Papini (2001).

The paper is organised as follows. After outlining some notation, we review the Darwiche-Pearl framework in Section 2. This is followed by a discussion of admissible revision in Section 3. In Section 4 we introduce restrained revision, and in Section 5 we show how it can be defined as a composite operator. Section 6 discusses the possibility of enriching epistemic states as a way of determining the appropriate admissible revision operator in a particular context. In this section we also conclude and briefly discuss some future work.

1.1 Notation

We assume a finitely generated propositional language L which includes the constants $\top$ and $\perp$, is closed under the usual propositional connectives, and is equipped with a classical model-theoretic semantics. V is the set of valuations of L and $[\alpha]$ (or $[B]$) is the set of models of $\alpha \in L$ (or $B \subseteq L$). Classical entailment is denoted by $\models$ and logical equivalence by $\equiv$. We also use Cn to denote the operation of closure under classical entailment. Greek letters $\alpha, \beta, \ldots$ stand for arbitrary sentences. In our examples we sometimes use the lower case letters p, q, and r as propositional atoms, and sequences of 0s and 1s to denote the valuations of the language. For example, 01 denotes the valuation, in a language generated by p and q, in which p is assigned the value 0 and q the value 1, while 011 denotes the valuation, in a language generated by p, q and r, in which p is assigned the value 0 and both q and r the value 1. Whenever we use the term *knowledge base* we will always mean a set of sentences X which is *deductively closed*, i.e., $X = Cn(X)$.

2. Darwiche-Pearl Revision

Darwiche and Pearl (1997) reformulated the AGM postulates (Alchourrón et al., 1985) to be compatible with their suggested approach to iterated revision. This necessitated a move from knowledge bases to *epistemic states*. An epistemic state contains, in addition to a knowledge base, all the information needed for coherent reasoning including, in particular, the strategy for belief revision which the agent wishes to employ at a given time. Darwiche and Pearl consider epistemic states as abstract entities, and do not provide a single formal representation. It is thus possible to talk about two epistemic states $\mathbb{E}$ and $\mathbb{F}$ being identical (denoted by $\mathbb{E} = \mathbb{F}$) , but yet syntactically different.[2] This has to be borne in mind below, particularly when considering postulate $(\mathbb{E} * 5)$. In Darwiche and Pearl's reformulated postulates $*$ is a belief change operator on epistemic states, not knowledge bases. We denote by $B(\mathbb{E})$ the knowledge base extracted from an epistemic state $\mathbb{E}$.

($\mathbb{E}$∗1) $B(\mathbb{E} * \alpha) = Cn(B(\mathbb{E} * \alpha))$

2. Personal communication with Adnan Darwiche.

($\mathbb{E}$∗2) $\alpha \in B(\mathbb{E} * \alpha)$

($\mathbb{E}$∗3) $B(\mathbb{E} * \alpha) \subseteq B(\mathbb{E}) + \alpha$

($\mathbb{E}$∗4) If $\neg\alpha \notin B(\mathbb{E})$ then $B(\mathbb{E}) + \alpha \subseteq B(\mathbb{E} * \alpha)$

($\mathbb{E}$∗5) If $\mathbb{E} = \mathbb{F}$ and $\alpha \equiv \beta$ then $B(\mathbb{E} * \alpha) = B(\mathbb{F} * \beta)$

($\mathbb{E}$∗6) $\perp \in B(\mathbb{E} * \alpha)$ iff $\models \neg\alpha$

($\mathbb{E}$∗7) $B(\mathbb{E} * (\alpha \wedge \beta)) \subseteq B(\mathbb{E} * \alpha) + \beta$

($\mathbb{E}$∗8) If $\neg\beta \notin B(\mathbb{E} * \alpha)$ then $B(\mathbb{E} * \alpha) + \beta \subseteq B(\mathbb{E} * (\alpha \wedge \beta))$

Darwiche and Pearl then show, via a representation result similar to that of Katsuno and Mendelzon (1991), that revision on epistemic states can be represented in terms of plausibility orderings associated with epistemic states.[3] More specifically, every epistemic state $\mathbb{E}$ has associated with it a total preorder $\preceq_{\mathbb{E}}$ on all valuations, with elements lower down in the ordering deemed more plausible. Moreover, for any two epistemic states $\mathbb{E}$ and $\mathbb{F}$ which are identical (but may be syntactically different), it has to be the case that $\preceq_{\mathbb{E}}=\preceq_{\mathbb{F}}$. Let $\min(\alpha, \preceq_{\mathbb{E}})$ denote the minimal models of α under $\preceq_{\mathbb{E}}$. The knowledge base associated with the epistemic state is obtained by considering the minimal models in $\preceq_{\mathbb{E}}$ i.e., $[B(\mathbb{E})] = \min(\top, \preceq_{\mathbb{E}})$. Observe that this means that $B(\mathbb{E})$ has to be consistent. This requirement enables us to obtain a unique knowledge base from the total preorder $\preceq_{\mathbb{E}}$. Preservation of the results in this paper when this requirement is relaxed is possible, but technically messy.

The observant reader will note that our assumption of a consistent $B(\mathbb{E})$ is incompatible with a successful revision by $\perp$. This requires that we jettison ($\mathbb{E}$∗6) and insist on consistent epistemic inputs only. (The left-to-right direction of ($\mathbb{E}$∗6) is rendered superfluous by ($\mathbb{E}$∗1) and the assumption that knowledge bases extracted from all epistemic states have to be consistent.) The other difference between the original AGM postulates and the Darwiche-Pearl reformulation – first inspired by a critical observation by Freund and Lehmann (1994) – occurs in ($\mathbb{E}$∗5), which states that revising by logically equivalent sentences results in epistemic states with identical associated knowledge bases. This is a weakening of the original AGM postulate, phrased in our notation as follows:

(B∗5) If $B(\mathbb{E}) = B(\mathbb{F})$ and $\alpha \equiv \beta$ then $B(\mathbb{E} * \alpha) = B(\mathbb{F} * \beta)$

(B∗5) states that two epistemic states with identical associated *knowledge bases* will, after having been revised by equivalent inputs, produce two epistemic states with identical associated knowledge bases. This is stronger than ($\mathbb{E}$∗5) which requires equivalent associated knowledge bases only if the original *epistemic states* were identical. We shall refer to the reformulated AGM postulates, with ($\mathbb{E}$∗6) removed, as DP-AGM.

DP-AGM guarantees a unique extracted knowledge base when revision by α is performed. It sets $[B(\mathbb{E} * \alpha)]$ equal to $\min(\alpha, \preceq_{\mathbb{E}})$ and thereby fixes the most plausible valuations in $\preceq_{\mathbb{E}*\alpha}$. However, it places no restriction on the rest of the ordering. The purpose

3. Alternative frameworks for studying iterated revision, both based on using sequences of sentences rather than plausibility orderings, are those of Lehmann (1995) and Konieczny and Pino-Pérez (2000).

of the Darwiche-Pearl framework is to constrain this remaining part of the new ordering. It is done by way of a set of postulates for iterated revision (Darwiche & Pearl, 1997). (Throughout the paper we follow the convention that $*$ is left associative.)

(C1) If $\beta \vDash \alpha$ then $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * \beta)$

(C2) If $\beta \vDash \neg\alpha$ then $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * \beta)$

(C3) If $\alpha \in B(\mathbb{E} * \beta)$ then $\alpha \in B(\mathbb{E} * \alpha * \beta)$

(C4) If $\neg\alpha \notin B(\mathbb{E} * \beta)$ then $\neg\alpha \notin B(\mathbb{E} * \alpha * \beta)$

The postulate (C1) states that when two pieces of information—one more specific than the other—arrive, the first is made redundant by the second. (C2) says that when two contradictory epistemic inputs arrive, the second one prevails; the second evidence alone yields the same knowledge base. (C3) says that a piece of evidence α should be retained after accommodating more recent evidence β that entails α given the current knowledge base. (C4) simply says that no epistemic input can act as its own defeater. We shall refer to the class of belief revision operators satisfying DP-AGM and (C1) to (C4) as *DP-revision*. The following are the corresponding semantic versions (with $v, w \in V$):

(CR1) If $v \in [\alpha], w \in [\alpha]$ then $v \preceq_{\mathbb{E}} w$ iff $v \preceq_{\mathbb{E}*\alpha} w$

(CR2) If $v \in [\neg\alpha], w \in [\neg\alpha]$ then $v \preceq_{\mathbb{E}} w$ iff $v \preceq_{\mathbb{E}*\alpha} w$

(CR3) If $v \in [\alpha], w \in [\neg\alpha]$ then $v \prec_{\mathbb{E}} w$ only if $v \prec_{\mathbb{E}*\alpha} w$

(CR4) If $v \in [\alpha], w \in [\neg\alpha]$ then $v \preceq_{\mathbb{E}} w$ only if $v \preceq_{\mathbb{E}*\alpha} w$

(CR1) states that the relative ordering between α-worlds remain unchanged following an α-revision, while (CR2) requires the same for $\neg\alpha$-worlds. (CR3) requires that, for an α-world strictly more plausible than a $\neg\alpha$-world, this relationship be retained after an α-revision, and (CR4) requires the same for weak plausibility. Darwiche and Pearl showed that, given DP-AGM, a precise correspondence obtains between (Ci) and (CRi) above ($i = 1, 2, 3, 4$).

One of the guiding principles of belief revision is the principle of minimal change: changes to a belief state ought to be kept to a minimum. What is not always clear is what ought to be minimised. In AGM theory the prevailing wisdom is that minimal change refers to the sets of sentences corresponding to knowledge bases. But there are other interpretations. With the move from knowledge bases to epistemic states, minimal change can be defined in terms of the fewest possible changes to the associated plausibility ordering $\preceq_{\mathbb{E}}$. In what follows we will frequently have the opportunity to refer to the latter interpretation of minimal change. See also the discussion of this principle by Rott (2000).

3. Admissible Revision

In this section we consider two of the best-known DP-operators, and propose three postulates to be added to the Darwiche-Pearl framework. The first is more of a correction than a strengthening. We show that the Darwiche-Pearl representation of the principle of the irrelevance of syntax is too weak and suggest an appropriate strengthened postulate.

The second is suggested by some of the arguments advanced by Darwiche and Pearl themselves. It eliminates one of the operators they criticise, and is satisfied by the sole operator they provide as an instance of their framework. The addition of these two postulates to the Darwiche-Pearl framework leads to the definition of the class of *admissible* revision operators. Finally, we point out a problem with Nayak's well-known lexicographic revision operator and propose a third postulate to be added. The consequences of insisting on the addition of this third postulate are discussed in detail in Section 4.

As mentioned in Section 2, Darwiche and Pearl replaced the original AGM postulate $(B * 5)$ with $(\mathbb{E} * 5)$. Both are attempts at an appropriate formulation of the principle of the irrelevance of syntax, popularised by Dalal (1988). But whereas $(B * 5)$ has been shown to be too strong, as shown by Darwiche and Pearl (1997), closer inspection reveals that $(\mathbb{E} * 5)$ is too *weak*. To be more precise, it fails as an adequate formulation of syntax irrelevance for *iterated* revision. It specifies that revision by two equivalent sentences should produce epistemic states with identical associated knowledge bases, but does not require that these epistemic states, after another revision by two equivalent sentences, also have to produce epistemic states with identical associated knowledge bases. So, as can be seen from the following example, under DP-AGM (and indeed, even if (C1) to (C4) are added) it is possible for $B(E * \alpha * \gamma)$ to differ from $B(E * \beta * \delta)$ even if α is equivalent to β and γ is equivalent to δ.

Example 1 *Consider a propositional language generated by the two atoms p and q and let $\mathbb{E}$ be an epistemic state such that $B(\mathbb{E}) = Cn(p \vee q)$. Now consider the two epistemic states $\mathbb{E}'$ and $\mathbb{E}''$ such that $B(\mathbb{E}') = B(\mathbb{E}'') = Cn(p)$, $01 \prec_{\mathbb{E}'} 00$ and $00 \prec_{\mathbb{E}''} 01$. Observe that this gives complete descriptions of $\preceq_{\mathbb{E}'}$ and $\preceq_{\mathbb{E}''}$. It is tedious, but not difficult, to verify that setting $\preceq_{\mathbb{E}*p} = \preceq_{\mathbb{E}'}$ and $\preceq_{\mathbb{E}*\neg\neg p} = \preceq_{\mathbb{E}''}$ is compatible with DP-AGM. But observe then that $B(\mathbb{E} * p * \neg p) = Cn(\neg p \wedge q)$, while $B(\mathbb{E} * \neg\neg p * \neg p) = Cn(\neg p \wedge \neg q)$.*

As a consequence of this, we propose that $(\mathbb{E}*5)$ be replaced by the following postulate:

$(\mathbb{E}*5')$ If $\mathbb{E} = \mathbb{F}$, $\alpha \equiv \beta$ and $\delta \equiv \gamma$ then $B(\mathbb{E} * \alpha * \gamma) = B(\mathbb{F} * \beta * \delta)$

The semantic equivalent of $(\mathbb{E}*5')$ looks like this:

$(\mathbb{E}R*5')$ If $\mathbb{E} = \mathbb{F}$ and $\alpha \equiv \beta$ then $\preceq_{\mathbb{E}*\alpha} = \preceq_{\mathbb{F}*\beta}$

$(\mathbb{E}R*5')$ states that the revision of two identical epistemic states by two equivalent sentences has to result in epistemic states with identical associated total preorders, not just in epistemic states with identical associated knowledge bases.

Proposition 1 *$(\mathbb{E}*5')$ and $(\mathbb{E}R*5')$ are equivalent, given DP-AGM.*

Proof: The proof that $(\mathbb{E}*5')$ follows from $(\mathbb{E}R*5')$ is straightfoward. For the converse, suppose that $(\mathbb{E}R*5')$ does not hold; i.e. $\alpha \equiv \beta$ and $\preceq_{\mathbb{F}*\alpha} \neq \preceq_{\mathbb{E}*\beta}$ for some α and β. This means there exist $x, y \in V$ such that $x \preceq_{\mathbb{E}*\alpha} y$ but $y \prec_{\mathbb{F}*\beta} x$. Now let γ be such that $[\gamma] = \{x, y\}$. Then $x \in [B(\mathbb{E} * \alpha * \gamma)]$, but $[B(\mathbb{F} * \beta * \gamma)] = \{y\}$, and so $B(\mathbb{E} * \alpha * \gamma) \neq B(\mathbb{F} * \beta * \gamma)$; a violation of $(\mathbb{E}*5')$. $\square$

From this it should already be clear that ($\mathbb{E}$∗5′) is a desirable property. This view is bolstered further by observing that all the well-known iterated revision operators satisfy it; natural revision, the Darwiche-Pearl operator $\bullet$, and Nayak's lexicographic revision, the first and third of which are to be discussed in detail below. In fact, we conjecture that Darwiche and Pearl's intention was to replace (B∗5) with ($\mathbb{E}$∗5′), not with ($\mathbb{E}$∗5) and propose this as a permanent replacement.

Definition 1 *The set of postulates obtained by replacing ($\mathbb{E}$∗5) with ($\mathbb{E}$∗5′) in DP-AGM is defined as RAGM.*

Observe that RAGM, like DP-AGM, guarantees that $[B(\mathbb{E} * \alpha)] = \min(\alpha, \preceq_{\mathbb{E}})$.

Rule ($\mathbb{E}$∗5′) is the first of the new postulates we want to add to the Darwiche-Pearl framework. We now lead up to the second. One of the oldest known DP-operators is *natural revision*, usually credited to Boutilier (1993, 1996), although the idea can also be found in (Spohn, 1988). Its main feature is the application of the principle of minimal change to epistemic states. It is characterised by DP-AGM plus the following postulate:

(CB) If $\neg\beta \in B(\mathbb{E} * \alpha)$ then $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * \beta)$

(CB) requires that, whenever $B(\mathbb{E} * \alpha)$ is inconsistent with β, revising $\mathbb{E} * \alpha$ with β will completely ignore the revision by α. Its semantic counterpart is as follows:

(CBR) For $v, w \notin [B(\mathbb{E} * \alpha)]$, $v \preceq_{\mathbb{E}*\alpha} w$ iff $v \preceq_{\mathbb{E}} w$

As shown by Darwiche and Pearl (1997), natural revision minimises changes in *conditional beliefs*, with $\beta \mid \alpha$ being a conditional belief of an epistemic state $\mathbb{E}$ iff $\beta \in B(\mathbb{E} * \alpha)$. In fact, Darwiche and Pearl show (Lemma 1, p. 7), that keeping $\preceq_{\mathbb{E}}$ and $\preceq_{\mathbb{E}*\alpha}$ as similar as possible has the effect of minimising the changes in conditional beliefs to a revision. So, from (CBR) it is clear that natural revision is an application of minimal change to epistemic states. It requires that, barring the changes mandated by DP-AGM, the relative ordering of valuations remains unchanged, thus keeping $\preceq_{\mathbb{E}*\alpha}$ as similar as possible to $\preceq_{\mathbb{E}}$. In that sense then, natural revision is the most conservative of *all* DP-operators. Such a strict adherence to minimal change is inadvisable and needs to be tempered appropriately, an issue that will be addressed in Section 5. Darwiche and Pearl have shown that (CB) is too strong, and that natural revision is not all that natural, sometimes yielding counterintuitive results.

Example 2 *(Darwiche & Pearl, 1997) We encounter a strange animal and it appears to be a bird, so we believe it is one. As it comes closer, we see clearly that the animal is red, so we believe it is a red bird. To remove further doubts we call in a bird expert who examines it and concludes that it is not a bird, but some sort of animal. Should we still believe the animal is red? (CB) tells us we should no longer believe it is red. This can be seen by substituting $B(\mathbb{E}) = Cn(\neg\beta) = Cn(\texttt{bird})$ and $\alpha \equiv \texttt{red}$ in (CB), instructing us to totally ignore the observation α as if it had never taken place.*

Given Example 2, it is perhaps surprising that Darwiche and Pearl never considered postulate (P) below. In this example, the argument for retaining the belief that the creature is red hinges upon the assumption that being red is not in conflict with the newly obtained information that it is a kind of animal. That is, because learning that the creature is an animal will not automatically disqualify it from being red, it is reasonable to retain the belief that it is red. More generally then, whenever α is consistent with a revision by β, it should be retained if an α-revision is inserted just before the β-revision.

(P) If $\neg\alpha \notin B(\mathbb{E} * \beta)$ then $\alpha \in B(\mathbb{E} * \alpha * \beta)$

Applying (P) to Example 2 we see that, if `red` is consistent with $B(\mathbb{E} * \neg\mathtt{bird})$, we have $red \in B(\mathbb{E} * \mathtt{red} * \neg\mathtt{bird})$. Put differently, (P) requires that you retain your belief in the animal's redness, provided this would not have been precluded if the observation about it being red had never occurred. (P) was also proposed independently of the present paper by Jin and Thielscher (2005) where it is named *Independence*. The semantic counterpart of (P) looks like this:

(PR) For $v \in [\alpha]$ and $w \in [\neg\alpha]$, if $v \preceq_{\mathbb{E}} w$ then $v \prec_{\mathbb{E}*\alpha} w$

(PR) requires an α-world v that is at least as plausible as a $\neg\alpha$-world w to be strictly more plausible than w after an α-revision. The following result was also proved independently by Jin and Thielscher (2005).

Proposition 2 *If $*$ satisfies DP-AGM, then it satisfies (P) iff it also satisfies (PR).*

Proof: For (P)$\Rightarrow$(PR), let $v \in [\alpha]$, $w \in [\neg\alpha]$, $v \preceq_{\mathbb{E}} w$, and let β be such that $[\beta] = \{v, w\}$. This means that $\neg\alpha \notin B(\mathbb{E} * \beta)$ (since $[B(\mathbb{E} * \beta)]$ is either equal to $\{v\}$ or to $\{v, w\}$), and so, by (P), $\alpha \in B(\mathbb{E} * \alpha * \beta)$. And therefore $v \prec_{\mathbb{E}*\alpha} w$, for if not, we would have that $w \preceq_{\mathbb{E}*\alpha} v$, from which it follows that $w \in [B(\mathbb{E} * \alpha * \beta)]$, and so $\alpha \notin B(\mathbb{E} * \alpha * \beta)]$.

For (P)$\Leftarrow$(PR), suppose that $\neg\alpha \notin B(\mathbb{E} * \beta)$. This means there is a $v \in [\alpha] \cap [B(\mathbb{E} * \beta)]$; that is, $v \preceq_{\mathbb{E}} w$ for every $w \in [\beta]$. And this means that $\alpha \in B(\mathbb{E} * \alpha * \beta)$. For if not, it means there is an x in $[\neg\alpha] \cap [B(\mathbb{E} * \alpha * \beta)]$. Now, since $x \in [B(\mathbb{E} * \alpha * \beta)]$, it follows from DP-AGM that $x \preceq_{\mathbb{E}*\alpha} w$ for every $w \in [\beta]$, and so $x \preceq_{\mathbb{E}*\alpha} v$ (since $v \in [B(\mathbb{E} * \beta)] \subseteq [\beta]$). But it also follows from DP-AGM that $x \in [\beta]$, and therefore that $v \preceq_{\mathbb{E}} x$, and by (PR) it then follows that $v \prec_{\mathbb{E}*\alpha} x$; a contradiction. $\Box$

Rule (PR) *enforces* certain changes in the ordering $\preceq_{\mathbb{E}}$ after receipt of α. In fact as soon as there exist an α-world v and a $\neg\alpha$-world w on the same plausibility level somewhere in $\preceq_{\mathbb{E}}$ (in that both $v \preceq_{\mathbb{E}} w$ and $w \preceq_{\mathbb{E}} v$), (PR) implies $\preceq_{\mathbb{E}*\alpha} \neq \preceq_{\mathbb{E}}$. Furthermore these changes must also occur even when α is already believed in $\mathbb{E}$ to begin with, i.e., $\alpha \in B(\mathbb{E})$. (Although of course if $\alpha \in B(\mathbb{E})$ then $B(\mathbb{E} * \alpha) = B(\mathbb{E})$, i.e., the *knowledge base* associated to $\mathbb{E}$ will remain unchanged – this follows from DP-AGM.) The rules (P)/(PR) ensure input α is believed with a certain *minimal strength* of belief – enough to help it survive the *next* revision. The point that being informed of α can lead to an increase in the strength of an agent's belief in α, even in cases where the agent already believes α to begin with, has been made before, e.g., by Friedman and Halpern (1999, p.405). Note that (P) has the antecedent of (C4) and the consequent of (C3). In fact, (P) is stronger than (C3) and (C4) combined. This is easily seen from the semantic counterparts of these postulates. It also follows that the only concrete example of an iterated revision operator provided by Darwiche and Pearl, the operator they refer to as $\bullet$ and which employs a form of Spohnian conditioning (Spohn, 1988), satisfies (PR), and therefore (P) as well. Furthermore, by adopting (P) we explicitly exclude natural revision as a permissible operator. So accepting (P) is a move towards the viewpoint that information obtained before the latest input ought not to be discarded unnecessarily.

Based on the analysis of this section we propose a strengthening of the Darwiche-Pearl framework in which ($\mathbb{E}$∗5) is replaced by ($\mathbb{E}$∗5′) and (C3) and (C4) are replaced by (P).

Definition 2 *A revision operator is* admissible *iff it satisfies RAGM, (C1), (C2), and (P).*

Inasmuch as the Darwiche-Pearl framework can be visualised as one in which α-worlds slide "downwards" relative to $\neg\alpha$-worlds, admissible revision ensures, via (PR), that this "downwards" slide is a strict one.

We now pave the way for the third postulate we would like to add in this paper to the Darwiche-Pearl framework. To begin with, note that another view of (P) is that it is a significant weakening of the following property, first introduced by Nayak et al. (1996):

(Recalcitrance) If $\neg\alpha \notin Cn(\beta)$ then $\alpha \in B(\mathbb{E} * \alpha * \beta)$

Semantically, (Recalcitrance) corresponds to the following property, as was pointed out by Booth (2005) and implicitly contained in the work of Nayak et al. (2003):

(R) For $v \in [\alpha]$, $w \in [\neg\alpha]$, $v \prec_{\mathbb{E}*\alpha} w$

(Recalcitrance) is a property of the *lexicographic revision* operator, the second of the well-known DP-operators we consider, and one that is just as old as natural revision. It was first introduced by Nayak (1993) and has been studied most notably by Nayak et al. (1994, 2003), although, as with natural revision, the idea actually dates back to Spohn (1988). In fact, lexicographic revision is characterised by DP-AGM (and also RAGM) together with (C1), (C2) and (Recalcitrance), a result that is easily proved from the semantic counterparts of these properties and Nayak et al.'s semantic characterisation of lexicographic revision in (2003). Informally, lexicographic revision takes the assumption of "most recent is best", on which the Success postulate ($\mathbb{E} * 2$) is based, and adds to it the assumption of temporal coherence. In combination, this leads to the stronger assumption that "*more* recent is better".

An analysis of the semantic characterisation of lexicographic revision shows that it is the *least* conservative of the DP-operators, in the sense that it effects the most changes in the relative ordering of valuations permitted by DP-AGM (or RAGM for that matter) and the Darwiche-Pearl postulates. Since it is also an admissible revision operator, it follows that it is also the least conservative admissible operator.

The problem with (Recalcitrance) is that the decision of whether to accept β after a subsequent revision by α is completely determined by the logical relationship between β and α – the epistemic state $\mathbb{E}$ is robbed of all influence. The replacement of (Recalcitrance) by the weaker (P) already gives $\mathbb{E}$ more influence in the outcome. What we will do shortly is constrain matters further by giving $\mathbb{E}$ as much influence as allowed by the postulates for admissible revision. Such a move ensures greater sensitivity to the agent's epistemic record in making further changes.

Note that lexicographic revision assumes that more recent information takes complete precedence over information obtained previously. Thus, when applied to Example 2, it requires us to believe that the animal, previously assumed to be a bird, is indeed red, because $\mathtt{r}ed$ is a recent input which does not conflict with the most recently obtained input. While this is a reasonable approach in many circumstances, a dogmatic adherence to it can be problematic, as the following example shows.

Example 3 *While holidaying in a wildlife park we observe a creature which is clearly red, but we are too far away to determine whether it is a bird or a land animal. So we adopt the knowledge base* $B(\mathbb{E}) = Cn(\mathtt{red})$. *Next to us is a person with knowledge of the local area who declares that, since the creature is red, it is a a bird. We have no reason to doubt him, and so we adopt the belief* $\mathtt{red} \rightarrow \mathtt{bird}$. *Now the creature moves closer and it becomes clear that it is not a bird. The question is, should we continue believing that it is red? Under the circumstances described above we want our initial observation to take precedence, and believe that the animal is red. But lexicographic revision does not allow us to do so.*

Other examples along similar lines speaking against a rigid acceptance of (Recalcitrance) are those of Glaister (1998, p.31) and Jin and Thielscher (2005, p.482).

While (P) allows for the possibility of retaining the belief that the animal is red, it does not *enforce* this belief. The rest of this section is devoted to the discussion of a property which does so. To help us express this property, we introduce an extra piece of terminology and notation.

Definition 3 α *and* β counteract *with respect to an epistemic state* $\mathbb{E}$, *written* $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$, *iff* $\neg\beta \in B(\mathbb{E} * \alpha)$ *and* $\neg\alpha \in B(\mathbb{E} * \beta)$.

The use of the term *counteract* to describe this relation is taken from Nayak et al. (2003). $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ means that, from the viewpoint of $\mathbb{E}$, α and β tend to "exclude" each other. We will now discuss a few properties of this relation. First note that $\leftrightsquigarrow_{\mathbb{E}}$ depends only on the total preorder $\preceq_{\mathbb{E}}$ obtained from $\mathbb{E}$. Indeed we have $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ iff both $\min(\alpha, \preceq_{\mathbb{E}}) \subseteq [\neg\beta]$ and $\min(\beta, \preceq_{\mathbb{E}}) \subseteq [\neg\alpha]$. This in turn can be reformulated in the following way, which provides a useful aid to visualise a counteracts relation:

Proposition 3 $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ *iff there exist* $v \in [\alpha]$, $w \in [\beta]$ *such that both* $v \prec_{\mathbb{E}} x$ *and* $w \prec_{\mathbb{E}} x$ *for all* $x \in \min(\alpha \wedge \beta, \preceq_{\mathbb{E}})$.

Proof: First note that, since obviously $\min(\alpha, \preceq_{\mathbb{E}}) \subseteq [\alpha]$, $\min(\alpha, \preceq_{\mathbb{E}}) \subseteq [\neg\beta]$ may be rewritten as $\min(\alpha, \preceq_{\mathbb{E}}) \subseteq [\neg(\alpha \wedge \beta)]$. Using the fact that $\preceq_{\mathbb{E}}$ is a total preorder, it is easy to see that this can hold iff there exists $v \in [\alpha]$ such that $v \prec_{\mathbb{E}} x$ for all $x \in \min(\alpha \wedge \beta, \preceq_{\mathbb{E}})$. In the same way we may rewrite $\min(\beta, \preceq_{\mathbb{E}}) \subseteq [\neg\alpha]$ as $\min(\beta, \preceq_{\mathbb{E}}) \subseteq [\neg(\alpha \wedge \beta)]$, which is then equivalent to saying there exists $w \in [\neg\beta]$ such that $w \prec_{\mathbb{E}} x$ for all $x \in \min(\alpha \wedge \beta, \preceq_{\mathbb{E}})$. $\square$

In other words, Proposition 3 says $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ iff there exist *both* an α-world and a β-world which are strictly more plausible than the most plausible $(\alpha \wedge \beta)$-worlds. Other immediate things to note about $\leftrightsquigarrow_{\mathbb{E}}$ are that it is symmetric, and that it is syntax-independent, i.e., if $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ and $\beta \equiv \beta'$ then $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta'$. Furthermore if α and β are logically inconsistent with each other then $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$, but the converse need not hold (see the short example after the next proposition for confirmation). Thus $\leftrightsquigarrow_{\mathbb{E}}$ can be seen as a weak form of inconsistency. The next result gives two more properties of $\leftrightsquigarrow_{\mathbb{E}}$:

Proposition 4 *Given RAGM, the following properties hold for* $\leftrightsquigarrow_{\mathbb{E}}$*:*
(i) *If* $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ *and* $\gamma \leftrightsquigarrow_{\mathbb{E}} \beta$ *then* $(\alpha \vee \gamma) \leftrightsquigarrow_{\mathbb{E}} \beta$
(ii) *If* $\alpha \not\leftrightsquigarrow_{\mathbb{E}} \beta$ *and* $\gamma \not\leftrightsquigarrow_{\mathbb{E}} \beta$ *then* $(\alpha \vee \gamma) \not\leftrightsquigarrow_{\mathbb{E}} \beta$

Proof: *(i)* Suppose $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ and $\gamma \leftrightsquigarrow_{\mathbb{E}} \beta$. To show $(\alpha \vee \gamma) \leftrightsquigarrow_{\mathbb{E}} \beta$ we need to show both $\neg\beta \in B(\mathbb{E} * (\alpha \vee \gamma))$ and $\neg(\alpha \vee \gamma) \in B(\mathbb{E} * \beta)$. For the former we already have both $\neg\beta \in B(\mathbb{E} * \alpha)$ and $\neg\beta \in B(\mathbb{E} * \gamma)$ from $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ and $\gamma \leftrightsquigarrow_{\mathbb{E}} \beta$ respectively. Since it follows from RAGM that $B(\mathbb{E} * \lambda) \cap B(\mathbb{E} * \chi) \subseteq B(\mathbb{E} * (\lambda \vee \chi))$ for *any* $\lambda, \chi \in L$, we can conclude from this $\neg\beta \in B(\mathbb{E} * (\alpha \vee \gamma))$. For the latter we already have both $\neg\alpha \in B(\mathbb{E} * \beta)$ and $\neg\gamma \in B(\mathbb{E} * \beta)$ from $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ and $\gamma \leftrightsquigarrow_{\mathbb{E}} \beta$ respectively. And from this we can conclude $\neg(\alpha \vee \gamma) \in B(\mathbb{E} * \beta)$, again using RAGM (specifically $(\mathbb{E} * 1)$).
(ii) Suppose $\alpha \not\leftrightsquigarrow_{\mathbb{E}} \beta$ and $\gamma \not\leftrightsquigarrow_{\mathbb{E}} \beta$. Firstly, if either $\neg\alpha \notin B(\mathbb{E} * \beta)$ or $\neg\gamma \notin B(\mathbb{E} * \beta)$ then we must have $\neg(\alpha \vee \gamma) \notin B(\mathbb{E} * \beta)$ by RAGM and so $(\alpha \vee \gamma) \not\leftrightsquigarrow_{\mathbb{E}} \beta$ as required. So suppose both $\neg\alpha \in B(\mathbb{E} * \beta)$ and $\neg\gamma \in B(\mathbb{E} * \beta)$. Then, since $\alpha \not\leftrightsquigarrow_{\mathbb{E}} \beta$ and $\gamma \not\leftrightsquigarrow_{\mathbb{E}} \beta$, this means we have both $\neg\beta \notin B(\mathbb{E} * \alpha)$ and $\neg\beta \notin B(\mathbb{E} * \gamma)$. Since it follows from RAGM that $B(\mathbb{E} * (\lambda \vee \chi)) \subseteq B(\mathbb{E} * \lambda) \cup B(\mathbb{E} * \chi)$ for *any* $\lambda, \chi \in L$, it follows from these two that $\neg\beta \notin B(\mathbb{E} * (\alpha \vee \gamma))$ and so also in this case $(\alpha \vee \gamma) \not\leftrightsquigarrow_{\mathbb{E}} \beta$ as required. □

The first property above says that if β counteracts with two sentences separately, then it counteracts with their disjunction, while the second says that it cannot counteract with a disjunction without counteracting with *at least one* of the disjuncts. Obviously these properties also hold for the binary relation of logical inconsistency. However one departure from the inconsistency relation is that it is possible to have both $\gamma \not\leftrightsquigarrow_{\mathbb{E}} \beta$ and $(\alpha \vee \gamma) \leftrightsquigarrow_{\mathbb{E}} \beta$. To see this assume for the moment L is generated by just three propositional atoms $\{p, q, r\}$ and take $\alpha = p$, $\beta = q$ and $\gamma = r$. Then take $\preceq_{\mathbb{E}}$ to be such that its lowest plausibility level contains only the two valuations 010 and 100, and the next plausibility level only the valuation 111.

We are now ready to introduce our third postulate. It is the following:

(D) If $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ then $\neg\alpha \in B(\mathbb{E} * \alpha * \beta)$

(D) requires that, whenever α and β counteract with respect to $\mathbb{E}$, α should be *disallowed* when an α-revision is followed by a β-revision. That is, when the β-revision of $\mathbb{E} * \alpha$ takes place, the information encoded in $\mathbb{E}$ takes precedence over the information contained in $\mathbb{E} * \alpha$. Darwiche and Pearl (1997) considered this property (it is their rule (C6)) but argued against it, citing the following example.

Example 4 *(Darwiche & Pearl, 1997) We believe that exactly one of John and Mary committed a murder. Now we get persuasive evidence indicating that John is the murderer. This is followed by persuasive information indicating that Mary is the murderer. Let* α *represent that John committed the murder and* β *that Mary committed the murder. Then (D) forces us to conclude that Mary, but not John, was involved in the murder. This, according to Darwiche and Pearl, is counterintuitive, since we should conclude that both were involved in committing the murder.*

Darwiche and Pearl's argument against (D) rests upon the assumption that more recent information ought to take precedence over information previously obtained. But as we have seen in Example 3, this is not always a valid assumption. In fact, the application of (D) to Example 3, with $\alpha = \texttt{red} \rightarrow \texttt{bird}$ and $\beta = \neg\texttt{bird}$, produces the intuitively correct result of a belief in the observed animal being red: $red \in B(\mathbb{E} * (\texttt{red} \rightarrow \texttt{bird}) * \neg\texttt{bird})$.

Another way to gain insight into the significance of (D) is to consider its semantic counterpart:

(DR) For $v \in [\neg\alpha]$, $w \in [\alpha]$, and $w \notin [B(\mathbb{E} * \alpha)]$, if $v \prec_{\mathbb{E}} w$ then $v \prec_{\mathbb{E}*\alpha} w$

(DR) curtails the rise in plausiblity of α-worlds after an α-revision. It ensures that, with the exception of the most plausible α-worlds, the relative ordering between an α-world and the $\neg\alpha$-worlds more plausible than it remains unchanged.

Proposition 5 *Whenever a revision operator $*$ satisfies RAGM, then $*$ satisfies (D) iff it satisfies (DR).*

Proof: For (D)$\Rightarrow$(DR), suppose that $v \in [\neg\alpha]$, $w \in [\alpha]$, $w \notin [B(\mathbb{E} * \alpha)]$, $v \prec_{\mathbb{E}} w$, and let β be such that $[\beta] = \{v, w\}$. Then $\neg\alpha \in B(\mathbb{E} * \beta)$ and $\neg\beta \in B(\mathbb{E} * \alpha)$, and so, by (D). $\neg\alpha \in B(\mathbb{E} * \alpha * \beta)$. From this it follows that $v \prec_{\mathbb{E}*\alpha} w$. For if not, we would have that $w \preceq_{\mathbb{E}*\alpha} v$, which means that $w \in [B(\mathbb{E} * \alpha * \beta)]$, and therefore that $\neg\alpha \notin B(\mathbb{E} * \alpha * \beta)$; a contradiction.

For (D)$\Leftarrow$(DR), suppose that $\neg\beta \in B(\mathbb{E} * \alpha)$ and that $\neg\alpha \in B(\mathbb{E} * \beta)$, but assume that $\neg\alpha \notin B(\mathbb{E} * \alpha * \beta)$. This means there is a $w \in [\alpha]$ that is also in $[B(\mathbb{E} * \alpha * \beta)]$. Now observe that $w \notin [B(\mathbb{E} * \beta)]$ since w is an α-model. Also, since $w \in [B(\mathbb{E} * \alpha * \beta)]$, it follows from RAGM that w is a β-model, and therefore $w \notin [B(\mathbb{E} * \alpha)]$. Our supposition of $\neg\alpha \in B(\mathbb{E} * \beta)$ means that $\min(\beta, \preceq_{\mathbb{E}}) \subseteq [\neg\alpha]$. Since $w \in [\alpha] \cap [\beta]$ it thus follows that there is a $v \in [\beta] \cap [\neg\alpha]$ such that $v \prec_{\mathbb{E}} w$. By (DR) it then follows that $v \prec_{\mathbb{E}*\alpha} w$. But then w cannot be a model of $B(\mathbb{E} * \alpha * \beta)$; a contradiction. □

4. Restrained Revision

We now strengthen the requirements on admissible revision (those operators satisfying RAGM, (C1), (C2) and (P)) by insisting that (D) is satisfied as well. To do so, let us first consider the semantic definition of an interesting admissible revision operator. Recall that RAGM fixes the set of $(\preceq_{\mathbb{E}*\alpha})$-minimal models, setting them equal to $\min(\alpha, \preceq_{\mathbb{E}})$, but places no restriction on how the remaining valuations should be ordered. The following property provides a *unique* relative ordering of the remaining valuations.

(RR) $\forall v, w \notin [B(\mathbb{E} * \alpha)]$, $v \preceq_{\mathbb{E}*\alpha} w$ iff $\begin{cases} v \prec_{\mathbb{E}} w \text{ or,} \\ v \preceq_{\mathbb{E}} w \text{ and } (v \in [\alpha] \text{ or } w \in [\neg\alpha]) \end{cases}$

(RR) says that the relative ordering of the valuations that are not $(\preceq_{\mathbb{E}*\alpha})$-minimal remains unchanged, except for α-worlds and $\neg\alpha$-worlds on the same plausibility level; those are split into two levels with the α-worlds more plausible than the $\neg\alpha$-worlds. So RAGM combined with (RR) fixes a unique ordering of valuations.

Definition 4 *The revision operator satisfying RAGM and (RR) is called* restrained revision.

It turns out that within the framework provided by admissible revision, it is only restrained revision that satisfies (D). We prove this with the help of the following lemma, asserting the equivalence of (RR) to (CR1), (CR2), (PR) and (DR) in the presence of RAGM.

Lemma 1 *Whenever a revision operator $*$ satisfies RAGM, then $*$ satisfies (RR) iff it satisfies (CR1), (CR2), (PR), and (DR).*

Proof: For (CR1)$\Leftarrow$(RR), pick $v, w \in [\alpha]$. If $v \in [B(\mathbb{E} * \alpha)]$ then (CR1) follows from RAGM. If not, it follows from RAGM that $w \notin [B(\mathbb{E} * \alpha)]$, and then (CR1) follows from a direct application of (RR). For (CR2)$\Leftarrow$(RR), pick $v, w \in [\neg\alpha]$. From RAGM it follows that $v, w \notin [B(\mathbb{E} * \alpha)]$, and then we obtain (CR2) from a direct application of (RR).

Observe that (RR) can be rewritten as

$$\textbf{(RR}'\textbf{)}\ \forall v, w \notin [B(\mathbb{E} * \alpha)],\ v \prec_{\mathbb{E}*\alpha} w \text{ iff } \begin{cases} v \preceq_{\mathbb{E}} w \text{ and,} \\ v \prec_{\mathbb{E}} w \text{ or } (v \in [\alpha] \text{ and } w \in [\neg\alpha]) \end{cases}$$

Now, for (PR)$\Leftarrow$(RR), pick $v \in [\alpha]$ and $w \in [\neg\alpha]$. If $v \in [B(\mathbb{E} * \alpha)]$ then (PR) follows from RAGM. If not, it follows from a direct application of (RR$'$). For (DR)$\Leftarrow$(RR), pick $v \in [\neg\alpha]$, $w \in [\alpha]$, and $w \notin [B(\mathbb{E} * \alpha)]$. Then (PR) follows from a direct application of (RR$'$).

For (CR1), (CR2), (PR), (DR)$\Rightarrow$(RR), let $v, w \notin [B(\mathbb{E} * \alpha)]$ and suppose that $v \preceq_{\mathbb{E}*\alpha} w$ and $v \not\prec_{\mathbb{E}} w$ (i.e. $w \preceq_{\mathbb{E}} v$). We have to show that $v \preceq_{\mathbb{E}} w$ and either $v \in [\alpha]$ or $w \in [\neg\alpha]$. Assume this is not the case. Then $w \prec_{\mathbb{E}} v$ or both $v \in [\neg\alpha]$ and $w \in [\alpha]$. Now, the second case is impossible because, together with $w \preceq_{\mathbb{E}} v$ and (PR) it implies that $w \prec_{\mathbb{E}*\alpha} v$; a contradiction. But the first case is also impossible. To see why, observe that by (CR1) it implies that v and w cannot both be α-models, by (CR2) v and w cannot both be $\neg\alpha$-models, by (DR) it cannot be the case that $w \in [\neg\alpha]$ and $v \in [\alpha]$. And by (PR) it cannot be the case that $w \in [\alpha]$ and $v \in [\neg\alpha]$. This concludes the first part of the proof of (CR1), (CR2), (PR), (DR)$\Rightarrow$(RR). For the second part, let $v, w \notin [B(\mathbb{E} * \alpha)]$ and suppose first that $v \prec_{\mathbb{E}} w$. If $v, w \in [\alpha]$ then $v \preceq_{\mathbb{E}*\alpha} w$ follows from (CR1). If $v, w \in [\neg\alpha]$ then $v \preceq_{\mathbb{E}*\alpha} w$ follows from (CR2). If $v \in [\alpha]$ and $w \in [\neg\alpha]$ then $v \preceq_{\mathbb{E}*\alpha} w$ follows from (PR). If $v \in [\neg\alpha]$ and $w \in [\alpha]$ then $v \preceq_{\mathbb{E}*\alpha} w$ follows from (DR). Now suppose that $v \preceq_{\mathbb{E}} w$ and either $v \in [\alpha]$ or $w \in [\neg\alpha]$. If $v \in [\alpha]$ then $v \preceq_{\mathbb{E}*\alpha} w$ follows either from (CR1) or (PR), depending on whether $w \in [\alpha]$ or $w \in [\neg\alpha]$. And similarly, if $w \in [\neg\alpha]$ then $v \preceq_{\mathbb{E}*\alpha} w$ follows either from (CR2) or (PR), depending on whether $v \in [\neg\alpha]$ or $v \in [\alpha]$. $\square$

Theorem 2 *RAGM, (C1), (C2), (P) and (D) provide an exact characterisation of restrained revision.*

Proof: The proof follows from Lemma 1, Proposition 2, Proposition 5, and the correspondence between (C1) and (CR1), and (C2) and (CR2). $\square$

Another interpretation of (RR) is that it maintains the relative ordering of the valuations that are not ($\preceq_{\mathbb{E}*\alpha}$)-minimal, except for the changes mandated by (PR). From this it can be seen that restrained revision is the most conservative of all admissible revision operators, in the sense that effects the least changes in the relative ordering of valuations permitted by admissible revision. So, in the context of admissible revision, restrained revision takes on the role played by natural revision in the Darwiche-Pearl framework.

In the rest of this section we examine some further properties of restrained revision. Firstly, Examples 3 and 4 share some interesting structural properties. In both, the initial knowledge base $B(\mathbb{E})$ is pairwise consistent with each of the subsequent sentences in the revision sequence, while the sentences in each revision sequence are pairwise inconsistent. And in both examples the information contained in the initial knowledge base $B(\mathbb{E})$ is retained after the revision sequence. These commonalities are instances of an important general result. Let Γ denote the non-empty sequence of inputs $\gamma_1, \ldots, \gamma_n$, and let $\mathbb{E} * \Gamma$ denote the revision sequence $\mathbb{E} * \gamma_1 * \ldots * \gamma_n$. Furthermore we shall refer to an epistemic state $\mathbb{E}$ as Γ-*compatible* provided that $\neg\gamma_i \notin B(\mathbb{E})$ for every i in $\{1, \ldots, n\}$.

(O) If $\mathbb{E}$ is Γ-compatible then $B(\mathbb{E}) \subseteq B(\mathbb{E} * \Gamma)$

(O) says that as long as $B(\mathbb{E})$ is not in direct conflict with *any* of the inputs in the sequence $\gamma_1, \ldots, \gamma_n$, the entire $B(\mathbb{E})$ *has* to be propagated to the knowledge base obtained from the revision sequence $\mathbb{E} * \gamma_1 * \ldots * \gamma_n$. This is a preservation property that is satisfied by restrained revision.

Proposition 6 *Restrained revision satisfies (O).*

Proof: We denote by $\mathbb{E} * \Gamma_i$, for $i = 0, \ldots, n$, the revision sequence $\mathbb{E} * \gamma_1, \ldots, \gamma_i$ (with $\mathbb{E} * \Gamma_0 = \mathbb{E}$). We give an inductive proof that, for $\forall v \in [B(\mathbb{E})]$ and $\forall w \notin [B(\mathbb{E})]$, $v \prec_{\mathbb{E}*\Gamma_i} w$ for $i = 0, \ldots, n$. In other words, every $B(\mathbb{E})$-world is always strictly below every non $B(\mathbb{E})$-world. From this the result follows immediately. For $i = 0$ this amounts to showing that $v \prec_{\mathbb{E}} w$ which follows immediately from the definition of $\preceq_{\mathbb{E}}$ and $B(\mathbb{E})$. Now pick any $i = 1, \ldots, n$ and assume that $v \prec_{\mathbb{E}*\Gamma_{i-1}} w$. We consider four cases. If $v, w \in [\gamma_i]$ then it follows by (CR1) that $v \prec_{\mathbb{E}*\Gamma_i} w$. If $v, w \in [\neg\gamma_i]$ then it follows by (CR2) that $v \prec_{\mathbb{E}*\Gamma_i} w$. If $v \in [\gamma_i]$ and $w \in [\neg\gamma_i]$ then it follows by (PR) that $v \prec_{\mathbb{E}*\Gamma_i} w$. And finally, suppose $v \in [\neg\gamma_i]$ and $w \in [\gamma_i]$. By Γ_i-compatibility there is an $x \in [B(E)] \cap [\gamma_i]$, and by the inductive hypothesis, $x \prec_{\mathbb{E}*\Gamma_{i-1}} w$. So $w \notin [B(\mathbb{E} * \Gamma_i)]$, and then it follows by (DR) that $v \prec_{\mathbb{E}*\Gamma_i} w$. □

Although restrained revision preserves information which has not been directly contradicted, it is not dogmatically wedded to older information. If neither of two successive, but incompatible, epistemic states are in conflict with any of the inputs of a sequence $\Gamma = \gamma_1, \ldots, \gamma_n$, it prefers the latter epistemic state when revising by Γ.

Proposition 7 *Restrained revision satisfies the following property:*

(Q) *If $\mathbb{E}$ and $\mathbb{E}*\alpha$ are both Γ-compatible but $B(\mathbb{E}) \cup B(\mathbb{E}*\alpha) \vDash \bot$, then $B(\mathbb{E}*\alpha) \subseteq B(\mathbb{E}*\alpha*\Gamma)$ and $B(\mathbb{E}) \nsubseteq B(\mathbb{E} * \alpha * \Gamma)$*

Proof: It follows immediately from Proposition 6 that $B(\mathbb{E} * \alpha) \subseteq B(\mathbb{E} * \alpha * \Gamma)$. And $B(\mathbb{E}) \nsubseteq B(\mathbb{E} * \alpha * \Gamma)$ then follows from the consistency of $B(\mathbb{E} * \alpha * \Gamma)$. □

Next we consider another preservation property, but this time, unlike the case for (O) and (Q), we look at circumstances where $B(\mathbb{E})$ is incompatible with some of the inputs in a revision sequence.

(S) If $\neg\beta \in B(\mathbb{E} * \alpha)$ and $\neg\beta \in B(\mathbb{E} * \neg\alpha)$ then $B(\mathbb{E} * \alpha * \neg\alpha * \beta) = B(\mathbb{E} * \alpha * \beta)$

Note that, given RAGM, the antecedent of (S) implies that $\neg\beta \in B(\mathbb{E})$. Thus (S) states that if $\neg\beta$ is believed initially, and that a subsequent commitment to either α or its negation would not change this fact, then after the sequence of inputs in which β is preceded by α and $\neg\alpha$, the *second* input concerning α is nullified, and the older input regarding α is retained.

Proposition 8 *Restrained revision satisfies (S).*

Proof: Suppose the antecedent holds. If $\neg\alpha \leftrightsquigarrow_{\mathbb{E}*\alpha} \beta$ then the consequent holds. In fact this can be seen from the property (T) in Proposition 10 below. So suppose $\neg\alpha \not\leftrightsquigarrow_{\mathbb{E}*\alpha} \beta$. Then either $\alpha \notin B(\mathbb{E} * \alpha * \beta)$ or $\neg\beta \notin B(\mathbb{E} * \alpha * \neg\alpha)$. This latter doesn't hold by one of the assumptions together with (C2), so the former must hold. This implies $\neg\alpha \in B(\mathbb{E} * \beta)$ by (P). Combining this with the other assumption we get $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$. In this case we get $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * \beta)$ (again using (T)), while (since $\neg\alpha \not\leftrightsquigarrow_{\mathbb{E}*\alpha} \beta$) $B(\mathbb{E} * \alpha * \neg\alpha * \beta) = B(\mathbb{E} * \alpha * (\neg\alpha \wedge \beta))$ ((T) once more), which in turn equals $B(\mathbb{E} * (\neg\alpha \wedge \beta))$ by (C2). Since $\neg\alpha \in B(\mathbb{E} * \beta)$ this is in turn equal to $B(\mathbb{E} * \beta)$ by RAGM as required. $\square$

We now provide a more compact syntactic representation of restrained revision. First we show that (C1) and (P) can be combined into a single property, and so can (C2) and (D).

Proposition 9 *Given RAGM,*

1. *(C1) and (P) are together equivalent to the single rule*

   (C1P) *If* $\neg\alpha \notin B(\mathbb{E} * \beta)$ *then* $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * (\alpha \wedge \beta))$

2. *(C2) and (D) are together equivalent to the single rule*

   (C2D) *If* $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ *then* $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * \beta)$.

Proof: For (C1),(P)$\Rightarrow$(C1P), suppose $\neg\alpha \notin B(\mathbb{E}*\beta)$. By (P) it follows that $\alpha \in B(\mathbb{E}*\alpha*\beta)$ which means, by RAGM, that $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * \alpha * (\alpha \wedge \beta))$. By (C1) it follows that $B(\mathbb{E}*\alpha*(\alpha\wedge\beta)) = B(\mathbb{E}*(\alpha\wedge\beta))$, and thus that $B(\mathbb{E}*\alpha*\beta) = B(\mathbb{E}*(\alpha\wedge\beta))$. For (C1)$\Leftarrow$(C1P), suppose that $\beta \vDash \alpha$. Then $\neg\alpha \notin B(\mathbb{E} * \beta)$ by RAGM, and so $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * (\alpha \wedge \beta))$ by (C1P). But since $\beta \equiv \alpha \wedge \beta$ it follows that $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * \beta)$. For (P)$\Leftarrow$(C1P), suppose that $\neg\alpha \notin B(\mathbb{E} * \beta)$. Then $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * (\alpha \wedge \beta))$ by (C1P) which means, by RAGM, that $\alpha \in B(\mathbb{E} * \alpha * \beta)$.

For (C2),(D)$\Rightarrow$(C2D), suppose that $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$. By (D), $\neg\alpha \in B(\mathbb{E} * \alpha * \beta)$. By RAGM this means that $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * \alpha * (\neg\alpha \wedge \beta))$. Now, by (C2) it follows that $B(\mathbb{E} * \alpha * (\neg\alpha \wedge \beta)) = B(\mathbb{E} * (\neg\alpha \wedge \beta))$. So $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * (\neg\alpha \wedge \beta))$. But since $\neg\alpha \in B(\mathbb{E}*\beta)$, we get by RAGM that $B(\mathbb{E}*(\neg\alpha\wedge\beta)) = B(\mathbb{E}*\beta)$, from which it follows that $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * \beta)$. For (C2)$\Leftarrow$(C2D), suppose that $\beta \vDash \neg\alpha$. Then $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ for any $\mathbb{E}$ and by $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * \beta)$ by (C2D). For (D)$\Leftarrow$(C2D), suppose that $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$. Then $B(\mathbb{E}*\alpha*\beta) = B(\mathbb{E}*\beta)$ by (C2D) and since $\neg\alpha \in B(\mathbb{E}*\beta)$, it follows that $\neg\alpha \in B(\mathbb{E}*\alpha*\beta)$. $\square$

Both (C1P) and (C2D) provide conditions for the reduction of the two-step revision sequence $\mathbb{E} * \alpha * \beta$ to a single-step revision (if only as regards the resulting knowledge base). (C1P) reduces it to an $(\alpha \wedge \beta)$-revision when α is consistent with a β-revision. (C2D) reduces it to a β-revision, ignoring α completely, when α and β counteract with respect to $\mathbb{E}$. Now, it follows from RAGM that the consequent of (C1P) also obtains when $\neg\beta \notin B(\mathbb{E} * \alpha)$. Putting this together we get a most succinct characterisation of restrained revision.

Proposition 10 *Only restrained revision satisfies RAGM and:*

$$\textbf{(T)}\ \ B(\mathbb{E} * \alpha * \beta) = \begin{cases} B(\mathbb{E} * \beta) & \text{if } \alpha \leftrightsquigarrow_{\mathbb{E}} \beta \\ B(\mathbb{E} * (\alpha \wedge \beta)) & \text{otherwise.} \end{cases}$$

Proof: From Theorem 2 and Proposition 9 it is sufficient to show that RAGM, (C1P) and (C2D) hold iff RAGM and (T) hold. So, suppose that $*$ satisfies RAGM and (T). (C1P) follows from the bottom part of (T), while (C2D) follows from the top part. Conversely, suppose that $*$ satisfies RAGM, (C1P) and (C2D). If $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ it follows from (C2D) that $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * \beta)$. If not, we consider two cases. If $\neg\alpha \notin B(\mathbb{E} * \beta)$ it follows from (C1P) that $B(\mathbb{E}*\alpha*\beta) = B(\mathbb{E}*(\alpha\wedge\beta))$. Otherwise it has to be the case that $\neg\beta \notin B(\mathbb{E}*\alpha)$. But then it follows from RAGM that $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * (\alpha \wedge \beta))$. $\square$

If we were to replace "$\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$" in the first clause in (T) by the stronger "α and β are logically inconsistent", we would obtain instead the characterisation of lexicographic revision given by Nayak et al. (2003).

Proposition 10 allows us to see clearly another significant property of restrained revision. For if $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ then we know $\neg\alpha \in B(\mathbb{E} * \alpha * \beta)$ directly from (D), while if $\alpha \not\leftrightsquigarrow_{\mathbb{E}} \beta$ then Proposition 10 tells us $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * (\alpha \wedge \beta))$ and so $\alpha \in B(\mathbb{E} * \alpha * \beta)$ by RAGM. Thus we see in the state $\mathbb{E} * \alpha * \beta$ the epistemic status of α (either accepted or rejected) is *always* completely determined, i.e., we have proved:

Proposition 11 *Restrained revision satisfies the following property:*

(U) *If* $\neg\alpha \notin B(\mathbb{E} * \alpha * \beta)$ *then* $\alpha \in B(\mathbb{E} * \alpha * \beta)$

(Given its similar characterisation just mentioned above, it is easy to see lexicographic revision satisfies (U) too.) Like (P), property (U) can be read as providing conditions under which the penultimate revision input α should be believed. Its antecedent is simply saying $B(\mathbb{E}*\alpha*\beta)$ is *consistent* with α. Thus (U) is saying the penultimate input should be believed *as long as it is consistent to do so.* By chaining (U) together with (C4), we easily see that (U) actually *implies* (P) in the presence of (C4). As a consequence, we obtain the following alternative axiomatic characterisation of restrained revision.

Theorem 3 *RAGM, (C1), (C2), (C4), (U) and (D) provide an exact characterisation of restrained revision.*

For (U), we are also able to provide a simple semantic counterpart property. It corresponds to a *separating* of all the α-worlds from all the $\neg\alpha$-worlds in the total preorder $\preceq_{\mathbb{E}*\alpha}$ following an α-revision, in that each plausibility level in $\preceq_{\mathbb{E}*\alpha}$ either contains only α-worlds or contains only $\neg\alpha$-worlds:

Proposition 12 *Whenever a revision operator $*$ satisfies RAGM, then $*$ satisfies (U) iff it satisfies the following property:*

(UR) *For $v \in [\alpha]$ and $w \in [\neg\alpha]$, either $v \prec_{\mathbb{E}*\alpha} w$ or $w \prec_{\mathbb{E}*\alpha} v$*

Proof: For (U)⇒(UR) suppose (UR) doesn't hold, i.e., there exist α, $v \in [\alpha]$ and $w \in [\neg\alpha]$ such that both $v \preceq_{\mathbb{E}*\alpha} w$ and $w \preceq_{\mathbb{E}*\alpha} v$. Letting β be such that $[\beta] = \{v, w\}$ we get $[B(\mathbb{E} * \alpha * \beta)] = \{v, w\}$ from RAGM and thus both $\neg\alpha, \alpha \notin B(\mathbb{E} * \alpha * \beta)$ (because of $v, w \in [B(\mathbb{E} * \alpha * \beta)]$ respectively). Hence (U) doesn't hold.

For (U)⇐(UR) suppose (U) doesn't hold, i.e., there exist α, β such that both $\neg\alpha, \alpha \notin B(\mathbb{E} * \alpha * \beta)$. Then there exist $v \in [\alpha]$ and $w \in [\neg\alpha]$ such that $v, w \in [B(\mathbb{E} * \alpha * \beta)] =$ (by RAGM) $\min(\beta, \preceq_{\mathbb{E}*\alpha})$. Since both v and w are $(\preceq_{\mathbb{E}*\alpha})$-minimal β-worlds we must have both $v \preceq_{\mathbb{E}*\alpha} w$ and $w \preceq_{\mathbb{E}*\alpha} v$. Hence α, v, w give a counterexample to (UR). □

Finally in this section we turn to two properties first mentioned (as far as we know) by Schlecta et al. (1996) (see also the work of Lehmann et al. (2001)):

(Disj1) $B(\mathbb{E} * \alpha * \beta) \cap B(\mathbb{E} * \gamma * \beta) \subseteq B(\mathbb{E} * (\alpha \vee \gamma) * \beta)$

(Disj2) $B(\mathbb{E} * (\alpha \vee \gamma) * \beta) \subseteq B(\mathbb{E} * \alpha * \beta) \cup B(\mathbb{E} * \gamma * \beta)$

(Disj1) says that if a sentence is believed after any one of two sequences of revisions that differ only at step i (step i being α in one case and γ in the other), then the sentence should also be believed after that sequence which differs from both only in that step i is a revision by the disjunction $\alpha \vee \gamma$. Similarly, (Disj2) says that every sentence believed after an $(\alpha \vee \gamma)$-β-revision should be believed after at least one of (α-β) and (γ-β). Both conditions are reasonable properties to expect of revision operators.

Proposition 13 *Restrained revision satisfies (Disj1) and (Disj2).*

To prove this result we will make use of the properties of the counteracts relation given in Proposition 4, along with the following lemma.

Lemma 4 *If $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ and $(\alpha \vee \gamma) \not\leftrightsquigarrow_{\mathbb{E}} \beta$ then $B(\mathbb{E} * ((\alpha \vee \gamma) \wedge \beta)) = B(\mathbb{E} * (\gamma \wedge \beta))$*

Proof: Suppose $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ and $(\alpha \vee \gamma) \not\leftrightsquigarrow_{\mathbb{E}} \beta$. We will first show that this implies $\neg\alpha \in B(\mathbb{E} * ((\alpha \vee \gamma) \wedge \beta))$. We will then be able to conclude the required $B(\mathbb{E} * ((\alpha \vee \gamma) \wedge \beta)) = B(\mathbb{E} * (\gamma \wedge \beta))$ using RAGM. So suppose on the contrary $\neg\alpha \notin B(\mathbb{E} * ((\alpha \vee \gamma) \wedge \beta))$. Then there exists some α-world $w \in \min((\alpha \vee \gamma) \wedge \beta, \preceq_{\mathbb{E}})$. Then also $w \in \min(\alpha \wedge \beta, \preceq_{\mathbb{E}})$. Since $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ we know from Proposition 3 there exist an α-world w_1 and a β-world w_2 such that $w_i \prec_{\mathbb{E}} w$ for $i = 1, 2$. Clearly w_1 is also a $(\alpha \vee \gamma)$-world, so we infer $(\alpha \vee \gamma) \not\leftrightsquigarrow_{\mathbb{E}} \beta$ – contradiction. Hence $\neg\alpha \in B(\mathbb{E} * ((\alpha \vee \gamma) \wedge \beta))$ as required. □

Proof:[of Proposition 13] We prove both properties simultaneously by looking at two cases:
Case *(i)*: $(\alpha \vee \gamma) \leftrightsquigarrow_{\mathbb{E}} \beta$. In this case $B(\mathbb{E} * (\alpha \vee \gamma) * \beta) = B(\mathbb{E} * \beta)$ by property (T) in Proposition 10. Meanwhile we know from Proposition 4*(ii)* that either $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ or $\gamma \leftrightsquigarrow_{\mathbb{E}} \beta$, and so using (T) again we know at least one of $B(\mathbb{E} * \alpha * \beta)$ and $B(\mathbb{E} * \gamma * \beta)$ must also be equal to $B(\mathbb{E} * \beta)$. Hence we see both (Disj1) and (Disj2) hold in this case.
Case *(ii)*: $(\alpha \vee \gamma) \not\leftrightsquigarrow_{\mathbb{E}} \beta$. In this case (T) tells us $B(\mathbb{E} * (\alpha \vee \gamma) * \beta) = B(\mathbb{E} * ((\alpha \vee \gamma) \wedge \beta)) = B(\mathbb{E} * ((\alpha \wedge \beta) \vee (\gamma \wedge \beta))$. Meanwhile Proposition 4*(i)* tells us at least one of $\alpha \not\leftrightsquigarrow_{\mathbb{E}} \beta$ and $\gamma \not\leftrightsquigarrow_{\mathbb{E}} \beta$ holds. We now consider two subcases according to which either both these

hold, or only one holds. If both these hold then $B(\mathbb{E} * \alpha * \beta) = B(\mathbb{E} * (\alpha \wedge \beta))$ and $B(\mathbb{E} * \gamma * \beta) = B(\mathbb{E} * (\gamma \wedge \beta))$, so (Disj1) and (Disj2) reduce to

(Disj1′) $B(\mathbb{E} * (\alpha \wedge \beta)) \cap B(\mathbb{E} * (\gamma \wedge \beta)) \subseteq B(\mathbb{E} * ((\alpha \wedge \beta) \vee (\gamma \wedge \beta)))$

(Disj2′) $B(\mathbb{E} * ((\alpha \wedge \beta) \vee (\gamma \wedge \beta))) \subseteq B(\mathbb{E} * (\alpha \wedge \beta)) \cup B(\mathbb{E} * (\gamma \wedge \beta))$

respectively. Now it is a consequence of RAGM that for *any* sentences θ, ϕ we have both (1) $B(\mathbb{E} * \theta) \cap B(\mathbb{E} * \phi) \subseteq B(\mathbb{E} * (\theta \vee \phi))$ and (2) $B(\mathbb{E} * (\theta \vee \phi)) \subseteq B(\mathbb{E} * \theta) \cup B(\mathbb{E} * \phi)$. Substituting $\alpha \wedge \beta$ for θ and $\gamma \wedge \beta$ for ϕ here gives us the required (Disj1′) (from (1)) and (Disj2′) (from (2)).

Now let's consider the subcase where $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ and $\gamma \not\leftrightsquigarrow_{\mathbb{E}} \beta$. (A symmetric argument will work for the other subcase $\alpha \not\leftrightsquigarrow_{\mathbb{E}} \beta$ and $\gamma \leftrightsquigarrow_{\mathbb{E}} \beta$.) Then from $\gamma \not\leftrightsquigarrow_{\mathbb{E}} \beta$ we get $B(\mathbb{E} * \gamma * \beta) = B(\mathbb{E} * (\gamma \wedge \beta))$, while from $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ together with $(\alpha \vee \gamma) \not\leftrightsquigarrow_{\mathbb{E}} \beta$ we get also $B(\mathbb{E}*(\alpha \vee \gamma)*\beta) = B(\mathbb{E}*(\gamma \wedge \beta))$ using Lemma 4. So in this case $B(\mathbb{E}*(\alpha \vee \gamma)*\beta) = B(\mathbb{E}*\gamma*\beta)$, from which both (Disj1) and (Disj2) follow immediately. $\square$

We end this section by remarking that it can be shown that lexicographic revision also satisfies (Disj1) and (Disj2).

5. Restrained Revision as a Composite Operator

As we saw in Section 3, Boutilier's natural revision operator – let us denote it in this section by $\oplus$ – is vulnerable to damaging counterexamples such as the red bird Example 2, and fails to satisfy the very reasonable postulate (P). Although a new input α is accepted in the very next epistemic state $\mathbb{E} \oplus \alpha$, $\oplus$ does not in any way provide for the preservation of α after subsequent revisions. As Hans Rott (2003, p.128) describes it, "[t]he most recent input sentence is always embraced without reservation, the last but one input sentence, however, is treated with utter disrespect". Thus, there seem to be convincing reasons to reject $\oplus$ as a viable operator for performing iterated revision. However, the literature on epistemic state change constantly reminds us that keeping changes *minimal* should be a major concern, and when judged from a purely *minimal change* viewpoint, it is clear that $\oplus$ can't be beaten! How can we find our way out of this apparent quandary? In this section we show that the use of $\oplus$ can be retained, *provided* its application is preceded by an *intermediate* operation in which, rather than revising $\mathbb{E}$ *by* new input α, essentially α *is revised by* $\mathbb{E}$.

Given an epistemic state $\mathbb{E}$ and sentence α, let us denote by $\mathbb{E} \triangleleft \alpha$ the result of this intermediate operation. $\mathbb{E} \triangleleft \alpha$ is an epistemic state. The idea is that when forming $\mathbb{E} \triangleleft \alpha$, the information in $\mathbb{E}$ should be maintained. That is, the total preorder $\preceq_{\mathbb{E} \triangleleft \alpha}$ should satisfy

$$v \prec_{\mathbb{E}} w \text{ implies } v \prec_{\mathbb{E} \triangleleft \alpha} w. \tag{1}$$

But rather than leaving behind α *entirely* in favour of $\mathbb{E}$, as much of the informational content of α should be preserved in $\mathbb{E} \triangleleft \alpha$ as possible. This is formalised by saying that for any $v \in [\alpha], w \in [\neg\alpha]$, we should take $v \prec_{\mathbb{E} \triangleleft \alpha} w$ as long as this does not conflict with (1) above. It is this second requirement which will guarantee α enough of a "presence" in the revised epistemic state $\mathbb{E} * \alpha$ to help it survive subsequent revisions and allow (P) to be captured. Taken together, the above two requirements are enough to specify $\preceq_{\mathbb{E} \triangleleft \alpha}$ uniquely:

$$v \preceq_{\mathbb{E} \triangleleft \alpha} w \text{ iff } \begin{cases} v \prec_{\mathbb{E}} w, \text{ or} \\ v \preceq_{\mathbb{E}} w \text{ and } (v \in [\alpha] \text{ or } w \in [\neg\alpha]). \end{cases} \tag{2}$$

Thus, $\preceq_{\mathbb{E} \triangleleft \alpha}$ is just the lexicographic refinement of $\preceq_{\mathbb{E}}$ by the "two-level" total preorder $\preceq_\alpha$ defined by $v \preceq_\alpha w$ iff $v \in [\alpha]$ or $w \in [\neg\alpha]$. This "backwards revision" operator is not new. It has been studied by Papini (2001). It can also be viewed as just a "backwards" version of Nayak's lexicographic revision operator. We do not necessarily have $\alpha \in B(\mathbb{E} \triangleleft \alpha)$ (this will hold only if $\neg\alpha \notin B(\mathbb{E})$), and so $\triangleleft$ does not satisfy RAGM.

Given $\triangleleft$, we can define the composite revision operator $*_\triangleleft$ by setting

$$\mathbb{E} *_\triangleleft \alpha = (\mathbb{E} \triangleleft \alpha) \oplus \alpha \tag{3}$$

This is reminiscent of the Levi Identity (Gärdenfors, 1988), used in AGM theory as a recipe for reducing the operation of revision on *knowledge bases* to a composite operation consisting of contraction plus expansion. In (3), $\oplus$ is playing the role of expansion. The operator $*_\triangleleft$ *does* satisfy RAGM. In fact, as can easily be seen by comparing (2) above with condition (RR) at the start of Section 4, $*_\triangleleft$ coincides with restrained revision.

Proposition 14 *Let $*_{\mathrm{R}}$ denote the restrained revision operator. Then $*_{\mathrm{R}} = *_\triangleleft$.*

Thus we have proved that restrained revision can be viewed as a *combination* of two existing operators.

6. How to Choose a Revision Operator

The contribution of this paper so far can be summarised as follows. We have argued for the replacement of the Darwiche-Pearl framework by the class of admissible revision operators, arguing that the former needs to be strengthened. In doing so we have eliminated natural revision, but retained lexicographic revision and the $\bullet$ operator of Darwiche and Pearl as admissible operators. We have also introduced a new admissible revision operator, restrained revision, and argued for its plausibility. But this is not an argument that restrained revision is somehow unique, or more preferred than other revision operators. The contention is merely that, for epistemic state revision, the Darwiche-Pearl framework is too weak and should be replaced by admissible revision. And restrained revision, being an admissible revision operator, is therefore only one of many revision operators deemed to be rational. The question of which admissible revision operator to use in a particular situation is one which depends on a number of issues, such as context, the strength with which certain beliefs are held, the source of the information, and so on. This point has essentially also been made by Friedman and Halpern (1999). For example, Example 3 formed part of an argument for the use of restrained revision, and against the use of lexicographic revision. In effect we used the example to argue that $\texttt{red}$ ought to be in $B(\mathbb{E} * \texttt{red} \rightarrow \texttt{bird} * \neg\texttt{bird})$, where $B(\mathbb{E}) = Cn(\texttt{red})$. But if we change the context slightly, it becomes an example in *favour* of the use of lexicographic revision, and *against* restrained revision.

Example 5 *We observe a creature which seems to be red, but we are too far away to determine whether it is a bird or a land animal. So we adopt the knowledge base* $B(\mathbb{E}) =$

$Cn(\texttt{red})$. Next to us is an expert on birds who remarks that, if the creature is indeed red, it must be a bird. So we adopt the belief $\texttt{red} \to \texttt{bird}$. Then we get information from someone standing closer to the creature that it is not a bird. Given this context, that is, the reliability of the expert combined with the statement that the creature initially seemed to be red, it is reasonable to adopt the lexicographic approach of "more recent is best" and conclude that the bird is not red. Formally, $\neg\texttt{red} \in B(\mathbb{E} * \texttt{red} \to \texttt{bird} * \neg\texttt{bird})$, where $B(\mathbb{E}) = Cn(\texttt{red})$.

For a case where the source of the information dramatically affects the outcome, consider the following example.

Example 6 *Consider the sequence of inputs where p is followed by a finite number, say n, of instances of the pair $p \to q, \neg q$. (To make this more concrete, the reader might wish to substitute p with* red *and q with* bird*.) Since p is not in direct conflict with any sentences in the sequence succeeding it, any revision operator satisfying the property O (and this includes restrained revision) will require that p be contained in the knowledge base obtained from this revision sequence. Now, if each pair $p \to q$ and $\neg q$ is obtained from a* different *source, such a conclusion is clearly unreasonable. After all, such a sequence amounts to being told that p is the case, followed by n different sources essentially telling you that $\neg p$ is the case. On the other hand, if the pairs $p \to q$ and $\neg q$ all come from the same source, the case is not so clear cut anymore. In fact, in this case one would expect the result to be the same as that obtained from the sequence $p, p \to q, \neg q$, a sequence with the same formal structure as that employed in Example 3, where restrained revision was seen to be a reasonable approach.*

Another example in which restrained revision fares less well is the following:[4]

Example 7 *Suppose we are teaching a class of students consisting of n boys and m girls, and suppose the class takes part in a mathematics competition. For each $i = 1, \ldots, n$ and $j = 1, \ldots, m$ let the propositional variables p_i and q_j stand for "boy i won the competition" and "girl j won the competition" respectively, and suppose initially we believe one of the boys won the competition, i.e., $B(\mathbb{E}) = Cn(\phi \wedge \Sigma)$ where $\phi = \bigvee_i p_i$ and Σ is just some sentence expressing the uniqueness of the competition winner. Now suppose we interview each of the boys one after the other, and each of them tells us that either one of the girls or he himself won, i.e., we obtain the sequence of inputs $(\neg\phi \vee p_i)_i$. Suppose we are willing to accept the boys' testimony. Using a revision operator which satisfies O will lead us to believe boy n won the competition, which seems implausible. Lexicographic revision gives the desired result that one of the girls won the competition.*

From these examples it is clear that an agent need not, and in most cases, *ought* not to stick to the same revision operator every time that it has to perform a revision. This means that the agent will keep on switching from one revision operator to another during the process of iterated revision. Of course, this leads to the question of how to choose among the available (admissible) revision operators at any particular point. A comprehensive answer to this question is beyond the scope of this paper, but we do provide some clues on how to address the problem. In brief, we contend that epistemic states have to be enriched, with a more detailed specification of their internal structure. Looking back at the history of belief revision, we can see that this is exactly how the field has progressed. In the initial papers,

4. We are grateful to one of the anonymous referees for suggesting this example.

such as those on AGM revision, an epistemic state was taken to contain nothing more than a knowledge base. So, for example, basic AGM revision as characterised by the first six AGM postulates imposes no structure on epistemic states at all. We shall refer to these as *simple* epistemic states. With full AGM belief revision as characterised by all eight AGM postulates, the view is still one of the revision of knowledge bases, but now every revision operator for knowledge base B is uniquely associated with a B-faithful total preorder; i.e., a total preorder on valuations with the models of B as its minimal elements. From here it is a small step to *define* epistemic states to include such an ordering, i.e. to include the total preorder $\preceq_{\mathbb{E}}$ associated with an epistemic state $\mathbb{E}$ as part of the definition of $\mathbb{E}$. We shall refer to these as *complex* epistemic states.

This leads to two different views of the same revision process. If we view revision as an operator on simple epistemic states we have many different revision operators; one corresponding to each of the B-faithful total preorders, but with no way to distinguish between them when having to choose a revision operator. Viewed as such, iterated revision is a process in which a (possibly) different revision operator is employed at every revision step. This is the principal view adopted by Nayak et al. (2003). However, if we view revision as an operator on complex epistemic states, every epistemic state contains enough information to determine *uniquely* the knowledge base, but not the faithful total preorder, resulting from the revision. In other words, we now have enough information encoded in an epistemic state to uniquely determine the knowledge base resulting from a revision, but we lack the information to uniquely determine the full epistemic state. The Darwiche-Pearl framework, and now also admissible revision, place some constraints on the resulting epistemic state, but do not impose any additional structure on the complex epistemic state. In our view admissible revision for complex epistemic states is analogous to basic AGM revision for simple epistemic states. The next step would thus be to impose additional structure on complex epistemic states. This could possibly involve the addition of a second ordering on valuations as was done, for example, by Booth et al. (2004). In the case of simple epistemic states the effect of adding the two supplementary postulates is to constrain basic revision to the extent that each revision operator can be uniquely associated with a B-faithful total preorder. In that sense, the addition of the supplementary postulates allowed for the imposition of additional structure on simple epistemic states. Recall that one way of interpreting the two supplementary postulates is that they explain the interaction between revision by two sentences and revision by their conjunction, something the basic postulates do not address.

So, as we have seen, the addition of the supplementary postulates leads to the definition of revision as operators on complex epistemic states. We conjecture that giving additional structure to complex epistemic states might involve the provision of postulates analogous to the two AGM supplementary postulates. In particular, we conjecture that such postulates might be such that they explain the interaction between two sentences and their conjunction, or disjunction, for *iterated revision*. Observe that none of the Darwiche-Pearl postulates, or the additional postulates for admissible revision for that matter, address this issue. In fact the only postulates to have been suggested (of which we are aware) which so far *do* address it are (Disj1) and (Disj2). We speculate that the appropriate set of supplementary postulates for iterated revision (which may or may not include the two just mentioned) will lead to the definition of extra structure on complex epistemic states, which can then be incorporated into an enriched version of complex epistemic states, with revision then being seen as operators on these enriched entities. Let us refer to them as *enriched* epistemic states. Enriched epistemic states will enable us to determine uniquely the complex epistemic state resulting from a revision, thereby solving the question we started off with; that of determining which revision on complex epistemic states to use at every particular point during a process of iterated revision. Below we shall briefly discuss a possible way of enriching complex epistemic states. But note also that a recent proposal for doing so is that of Booth et al. (2006). It is instructive to observe that (Disj1) and (Disj2) both hold in their framework.

The proposed outline above is not without its pitfalls. The most obvious problem with such an approach is that it leaves us with a meta-version of the dilemma that we started off with. Using enriched epistemic states we are now able to uniquely determine the complex epistemic states resulting from a revision, but not the resulting *enriched* epistemic state. We can lessen the problem by constraining the permissible resulting enriched epistemic states in the same way that admissible revision constrains the permissible complex epistemic states, but chances are that this whittling down will not produce a single permissible enriched epistemic state. And, of course, this is bound to occur over and over again. That is, whenever we solve the problem of uniquely determining an epistemic state with a certain structure by a process of further enrichment, we will be saddled with the question of how to uniquely determine the further enriched epistemic state resulting from a revision. Our conjecture is that at some level a point will be reached where constraining further enriched epistemic states, *á la* admissible revision, will eventually lead to a unique further enriched epistemic state associated with every revision. Only further research will determine whether our conjecture holds water.

In conclusion, we have shown that the Darwiche-Pearl arguments lead to the acceptance of the admissible revision operators as a class worthy of study. The restrained revision operator, in particular, exhibits quite desirable properties. Besides taking the place of natural revision as the operator adhering most closely to the principle of minimal change, its satisfaction of the properties (O), (Q) and (U) shows that it does not unnecessarily remove previously obtained information.

For future work we would also like to explore more thoroughly the class of admissible revision operators. In this paper we saw that restrained revision and lexicographic revision lie at opposite ends of the spectrum of admissible operators. They represent respectively the most conservative and the least conservative admissible operators in the sense that they effect the most changes and the least changes, respectively, in the relative ordering of valuations permitted by admissible revision. A natural question is whether there exists an axiomatisable class of admissible operators which represents the "middle ground". One clue for finding such a class can be found in the counteracts relation $\leftrightsquigarrow_{\mathbb{E}}$ which can be derived from an epistemic state $\mathbb{E}$. As we said, this relation depends only on the preorder $\preceq_{\mathbb{E}}$ associated to $\mathbb{E}$. In fact, given *any* total preorder $\preceq$ over V we can define the relation $\leftrightsquigarrow_{\preceq}$ by

$$\alpha \leftrightsquigarrow_{\preceq} \beta \text{ iff } \min(\alpha, \preceq) \subseteq [\neg\beta] \text{ and } \min(\beta, \preceq) \subseteq [\neg\alpha].$$

Then clearly $\leftrightsquigarrow_{\mathbb{E}} = \leftrightsquigarrow_{\preceq_{\mathbb{E}}}$. Furthermore if $\preceq$ is the full relation $V \times V$ then $\leftrightsquigarrow_{\preceq}$ reduces to logical inconsistency. A counteracts relation *stronger* than $\leftrightsquigarrow_{\mathbb{E}}$, but still *weaker* than logical inconsistency can be found by setting $\leftrightsquigarrow = \leftrightsquigarrow_{\preceq'}$, where $\preceq'$ lies somewhere *in between* $\preceq_{\mathbb{E}}$

and $V \times V$. Hence one avenue worth exploring might be to assume that from each epistemic state $\mathbb{E}$ we can extract not one but *two* preorders $\preceq_{\mathbb{E}}$ and $\preceq'_{\mathbb{E}}$ such that $\preceq_{\mathbb{E}} \subseteq \preceq'_{\mathbb{E}}$. Then, instead of only requiring $\alpha \leftrightsquigarrow_{\mathbb{E}} \beta$ to deduce $\neg\alpha \in B(\mathbb{E} * \alpha * \beta)$, as is done with restrained revision (the postulate (D)), we could require the stronger condition $\alpha \leftrightsquigarrow_{\preceq'_{\mathbb{E}}} \beta$ for this to hold. We are currently experimenting with strategies for using the second preorder to *guide* the manipulation of $\preceq_{\mathbb{E}}$ to enable this property to be satisfied. The use of a second preorder can be seen as a way of enriching the epistemic state, and might thus contribute to the solution of the choice of revision operators discussed in Section 6.

Some more future work relates also to the two extreme cases revision, but looked at from a different angle. As mentioned earlier, lexicographic revision is a formalisation of the "most recent is best" approach to revision taken to its logical extreme. This approach is exemplified by the ($\mathbb{E}$∗2) postulate, also known as Success, which requires a revision to be successful, in the sense that the epistemic input provided always has to be contained in the resulting knowledge base. Given that ($\mathbb{E}$∗2) is one of the postulates for admissible revision, this requirement carries over even to restrained revision, which is on the opposite end of the spectrum for admissible revision. But this means that the admissible revision operator which differs the most from lexicographic revision still adheres to the dictum of "most recent is best", which raises the question of why the most recent input is given such prominence. The relaxation of this requirement would imply giving up ($\mathbb{E}$∗2) and venturing into the area known as *non-prioritised* revision (Booth, 2001; Chopra, Ghose, & Meyer, 2003; Hansson, 1999). We speculate that an appropriate relaxation of admissible revision, with ($\mathbb{E}$∗2) removed as a requirement, will lead to a class of (non-prioritised) revision operators strictly containing admissible revision, and with lexicographic revision still at one end of the spectrum, but with the other end of the spectrum occupied by the operator studied by Papini (2001) which was used as a sub-operation of restrained revision in Section 5. This operator is formalised by the extreme version of "most recent is worst"; in other words, "the older the better".

Acknowledgements

Much of the first author's work was done during stints as a researcher at Wollongong University and Macquarie University, Sydney. He wishes to thank Aditya Ghose and Abhaya Nayak both for making it possible to enjoy the great working environments there, and also for some interesting comments on this work. Thanks are also due to Samir Chopra who contributed to a preliminary version of the paper, Adnan Darwiche for clearing up some misconceptions on the definition of epistemic states, and three anonymous referees for their valuable and insightful comments. National ICT Australia is funded by the Australia Government's Department of Communications, Information and Technology and the Arts and the Australian Research Council through Backing Australia's Ability and the ICT Centre of Excellence program. It is supported by its members the Australian National University, University of NSW, ACT Government, NSW Government and affiliate partner University of Sydney.

References

Alchourrón, C. E., Gärdenfors, P., & Makinson, D. (1985). On the logic of theory change: Partial meet functions for contraction and revision. *Journal of Symbolic Logic*, *50*, 510–530.

Areces, C., & Becher, V. (2001). Iterable AGM functions. In *Frontiers in belief revision*, pp. 261–277. Kluwer, Dordrecht.

Booth, R. (2001). A negotiation-style framework for non-prioritised revision. In van Benthem, J. (Ed.), *Theoretical Aspects of Rationality and Knowledge: Proceedings of the Eighth Conference (TARK 2001)*, pp. 137–150, San Francisco, California. Morgan Kaufmann.

Booth, R. (2005). On the logic of iterated non-prioritised revision. In *Conditionals, Information and Inference – Selected papers from the Workshop on Conditionals, Information and Inference, 2002*, Vol. 3301 of *LNAI*, pp. 86–107. Springer-Verlag, Berlin.

Booth, R., Chopra, S., Ghose, A., & Meyer, T. (2004). A unifying semantics for belief change. In Mantaras, R. L. D., & Saitta, L. (Eds.), *Sixteenth European Conference on Artificial Intelligence: ECAI2004*, pp. 793–797. IOS Press.

Booth, R., Meyer, T., & Wong, K.-S. (2006). A bad day surfing is better than a good day working: How to revise a total preorder. In *Proceedings of KR2006, Tenth International Conference on the Principles of Knowledge Representation and Reasoning*.

Boutilier, C. (1993). Revision sequences and nested conditionals. In Bajcsy, R. (Ed.), *IJCAI-93. Proceedings of the 13th International Joint Conference on Artificial Intelligence held in Chambery, France, August 28 to September 3, 1993*, Vol. 1, pp. 519–525, San Mateo, CA. Morgan Kaufmann.

Boutilier, C. (1996). Iterated revision and minimal changes of conditional beliefs. *Journal of Philosophical Logic*, *25*(3), 263–305.

Chopra, S., Ghose, A., & Meyer, T. (2003). Non-prioritized ranked belief change. *Journal of Philosophical Logic*, *32*(3), 417–443.

Dalal, M. (1988). Investigations into a theory of knowledge base revision. In *Proceedings of the 7th National Conference of the American Association for Artificial Intelligence, Saint Paul, Minnesota*, pp. 475–479.

Darwiche, A., & Pearl, J. (1997). On the logic of iterated belief revision. *Artificial Intelligence*, *89*, 1–29.

Freund, M., & Lehmann, D. (1994). Belief revision and rational inference. Tech. rep. TR 94-16, The Leibniz Centre for Research in Computer Science, Institute of Computer Science, Hebrew University of Jerusalem.

Friedman, N., & Halpern, J. Y. (1999). Belief revision: A critique. *Journal of Logic, Language and Information*, *8*, 401–420.

Gärdenfors, P. (1988). *Knowledge in Flux : Modeling the Dynamics of Epistemic States.* The MIT Press, Cambridge, Massachusetts.

Glaister, S. M. (1998). Symmetry and belief revision. *Erkenntnis*, *49*, 21–56.

Grove, A. (1988). Two modellings for theory change. *Journal of Philosophical Logic*, *17*, 157–170.

Hansson, S. O. (1999). A survey of non-prioritized belief revision. *Erkenntnis*, *50*, 413–427.

Jin, Y., & Thielscher, M. (2005). Iterated belief revision, revised. In *Proceedings of the Nineteenth International Joint Conference on Artificial Intelligence (IJCAI 05)*, pp. 478–483.

Katsuno, H., & Mendelzon, A. O. (1991). Propositional knowledge base revision and minimal change. *Artificial Intelligence*, *52*, 263–294.

Konieczny, S., & Pino Pérez, R. (2000). A framework for iterated revision. *Journal of Applied Non-Classical Logics*, *10(3-4)*, 339–367.

Lehmann, D. (1995). Belief revision, revised. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI'95)*, pp. 1534–1540.

Lehmann, D., Magidor, M., & Schlechta, K. (2001). Distance semantics for belief revision. *Journal of Symbolic Logic*, *66*, 295–317.

Lewis, D. K. (1973). Counterfactuals. *Journal of Philosophy*, *70*, 556–567.

Nayak, A. C. (1993). *Studies in Belief Change*. Ph.D. thesis, University of Rochester.

Nayak, A. C. (1994). Iterated belief change based on epistemic entrenchment. *Erkenntnis*, *41*, 353–390.

Nayak, A. C., Foo, N. Y., Pagnucco, M., & Sattar, A. (1996). Changing Conditional Belief Unconditionally. In Shoham, Y. (Ed.), *Theoretical Aspects of Rationality and Knowledge: Proceedings of the Sixth Conference (TARK 1996)*, pp. 119–136, San Francisco, California. Morgan Kaufmann.

Nayak, A. C., Pagnucco, M., & Peppas, P. (2003). Dynamic belief change operators. *Artificial Intelligence*, *146*, 193–228.

Papini, O. (2001). Iterated revision operations stemming from the history of an agent's observations. In *Frontiers in belief revision*, pp. 281–303. Kluwer, Dordrecht.

Rott, H. (2000). Two dogmas of belief revision. *Journal of Philosophy*, *97*, 503–522.

Rott, H. (2003). Coherence and conservatism in the dynamics of belief II: Iterated belief change without dispositional coherence. *Journal of Logic and Computation*, *13*(1), 111–145.

Schlechta, K., Lehmann, D., & Magidor, M. (1996). Distance semantics for belief revision. In Shoham, Y. (Ed.), *Proceedings of the Sixth Conference on Theoretical Aspects of Rationality and Knowledge*, pp. 137–145. Morgan Kaufmann.

Segerberg, K. (1998). Irrevocable belief revision in dynamic doxastic logic. *Notre Dame Journal of Formal Logic*, *39*, 287–306.

Spohn, W. (1988). Ordinal conditional functions: A dynamic theory of epistemic states. In Harper, W. L., & Skyrms, B. (Eds.), *Causation in Decision: Belief, Change and Statistics: Proceedings of the Irvine Conference on Probability and Causation: Volume II*, Vol. 42 of *The University of Western Ontario Series in Philosophy of Science*, pp. 105–134, Dordrecht. Kluwer Academic Publishers.