

Iterated revision and the axiom of recovery: a unified treatment via epistemic states

Samir Chopra¹ and Aditya Ghose² and Thomas Meyer³

Abstract. The axiom of recovery, while capturing a central intuition regarding belief change, has been the source of much controversy. We argue briefly against putative counterexamples to the axiom—while agreeing that some of their insight deserves to be preserved—and present additional recovery-like axioms in a framework that uses epistemic states, which encode preferences, as the object of revisions. This makes iterated revision possible and makes explicit the connection between iterated belief change and the axiom of recovery. We provide a representation theorem that connects the semantic conditions we impose on iterated revision and the additional syntactical properties mentioned. We also show interesting similarities between our framework and that of Darwiche-Pearl [5]. In particular, we show that the intuitions underlying the controversial (C2) postulate are captured by the recovery axiom and our recovery-like postulates (the latter can be seen as weakenings of (C2)).

1 Introduction

A particularly simple sequence of belief change in reasoning agents is that of giving up and then adopting the same belief (“I believed I had money for the movies, but realized I had lost my wallet. A few minutes later, I discovered a twenty in my pocket and regained my belief that I had enough money for the movies”). The axiom of recovery in the AGM framework [1] places a rationality constraint on the form of such a change. It states that expansion by a belief recovers any beliefs lost by the previous contraction by that belief. The status of the axiom of recovery has been a source of much controversy in belief revision [7, 8, 9, 12]. There are well-known counterexamples to recovery, with the most convincing ones being Hansson’s Cleopatra and George-the-criminal examples [8, 10]. The following is a slightly amended version of the former:

I believe that ‘Cleopatra had a son’ (ϕ) and that ‘Cleopatra had a daughter’ (ψ), and thus also that ‘Cleopatra had a child’ ($\phi \vee \psi$). Then I receive information that Cleopatra had no children, which makes me give up my belief in $\phi \vee \psi$. But then I am told that Cleopatra did have children, and so I add $\phi \vee \psi$. But I should not regain my belief in either ψ or ϕ as a result.

One response to this situation is to isolate a class of belief change operators that do not satisfy recovery i.e., the so-called *withdrawal*

operators [13]. We do not adopt this approach for a couple of reasons. Firstly, withdrawal operators can violate the principle of minimal change [10]. As an example, consider the withdrawal operator — defined as follows (K is a belief set closed under a logical consequence operator Cn , α an arbitrary epistemic input): if $\alpha \notin K$, then $K - \alpha = K$, otherwise, $K - \alpha = Cn(\emptyset)$. Secondly, a fundamental intuition behind minimal contraction is the *principle of core-retainment* which states that if $\beta \in K$ and $\beta \notin K - \alpha$ then there is a set K' such that $K' \subseteq K$ and that $\alpha \notin Cn(K')$ but $\alpha \in Cn(K' \cup \{\beta\})$. It requires of an excluded sentence β that it in some way contribute to the implication of α from K . This is only satisfied by withdrawal operators if they satisfy the recovery axiom as well. This should reinstate our faith in the recovery axiom since it is hard to find a satisfactory alternative formalization of the intuition that beliefs that do not contribute to K implying α should be retained in $K - \alpha$. So, while the counterexamples do tickle our intuitions, it is equally the case that there is an important intuition about rational belief change that the recovery postulate captures. Indeed, the recovery postulate is best thought of as a version of the principle of minimal change: so much of the original belief state is retained on contraction that the original belief state can simply be restored on adopting the same belief. Our opinion is that even if the original postulate is rejected as being too permissive, *some* recovery like postulates must constrain belief revision if the principle of minimal change is to be taken seriously. Furthermore, recovery follows from other highly plausible postulates such as closure, inclusion, vacuity, success, extensionality and core-retainment [10]. Significantly, there is a clear and intimate connection between *iterated revision* and the recovery axiom: we can view the axiom as specifying the form of the iterated revision that should take place when contracting and revising by the same belief. In what follows, we make this connection clearer.

But what about the counterexamples? Surely, they point to counterintuitive scenarios arising from the adoption of the recovery axiom? We argue that, underlying these examples is an assumption that information leading to the specified sequence of contraction and expansion is not received from the same source. Our claim is that recovery should hold when restricted to the case where information is obtained from the same source, but that it need not hold when information is obtained from different sources. Consider the Cleopatra counterexample. The agent believes both ϕ and ψ originally, and as a result is committed to the belief that $\phi \vee \psi$. Now the agent receives information that $\neg(\phi \vee \psi)$. Crucially, what is left out of this example are details about the sources of the epistemic inputs. If source S_1 provides the *reasons* for believing $\neg(\phi \vee \psi)$ and source S_2 provides the reason for believing $\phi \vee \psi$ then it makes sense to think that the agent does not recover its original beliefs in ϕ or ψ . However, if it is the same source that provides information on both $\neg(\phi \vee \psi)$ and $\phi \vee \psi$,

¹ Knowledge Systems Group, School of Computer Science and Engineering, University of New South Wales, NSW 2052, Australia. Email: schopra@cse.unsw.edu.au

² Department of Information Systems, University of Wollongong, Wollongong, NSW 2522, Australia. Email: aditya@uow.edu.au

³ Computer Science Department, University of Pretoria, Pretoria 0002, South Africa. Email: tmeyer@cs.up.ac.za

then why should the agent not regain its belief in ϕ and ψ ? After all, source S_1 provided the reason for the agent dropping its belief in ϕ and ψ in the first place. If it then supplies information to the contrary, the agent's reasons for dropping those beliefs have been negated, and it should regain its original beliefs. To do otherwise would be counterintuitive. If however another source provides the new information, then the agent's original reasons for contracting by ϕ and ψ remain unaffected and there is no reason for it to start believing ϕ or ψ again. (For a similar though crucially different response see [16]). The issue of what happens when information is obtained from different sources is interesting in its own right, and deserves to be treated separately. For the remainder of this paper we will assume that all information is received from one source.

Our proposal considers versions of postulates in the same spirit as recovery. We argue that a shift to belief change on epistemic states i.e., belief states possessing a preferential structure, in the Darwiche-Pearl spirit is necessary, since we need a framework in which to talk about iterated revision. Cantwell [4] also provides recovery-like properties in the context of iterated revision, but these however restate recovery itself in terms of revision—where contracting with α is replaced by a revision with $\neg\alpha$. This is done to show that the counterexamples to recovery are not only a criticism of AGM contraction—as has been argued in the past—but also a criticism of AGM revision. Cantwell provides examples similar to the Cleopatra and George-the-criminal examples for iterated revision as well.

While adopting the representational framework of epistemic states, we do not accept all the Darwiche-Pearl postulates. There is sufficient debate on the appropriateness of these. In principle, we are of the opinion that the 3rd and 4th Darwiche-Pearl postulates are valid. Like others we feel that the 2nd postulate is too strong. The results in this paper provide a weaker and acceptable alternative to the 2nd postulate. We are also of the opinion that the 1st Darwiche-Pearl postulate is too strong ([14] provides examples to back up this claim) but will not provide a weakening here. We adopt the basic setting in which belief change is performed on epistemic states, from which a total preorder on valuations and a knowledge base can be extracted. We provide a set of reformulated AGM postulates for contraction—along with the Darwiche-Pearl reformulations of the AGM revision postulates—on epistemic states and *insist* on these.

We present some recovery-like postulates, as well as restrictions on the way in which the orderings extracted from epistemic states may be modified when revision and contraction take place, and provide a representation theorem that connects the recovery-like postulates and the postulates on orderings. The recovery-like postulates, when combined, can be thought of as a weakened version of the (C2) postulate of Darwiche-Pearl. This is brought out clearly when the postulates on orderings are considered. The link between recovery and the (C2) postulate is interesting and surprising. This lets us think of (C2) as having overstated the case and of the recovery postulate and our weakenings as having addressed its problems.

We assume a finitely generated propositional language L closed under the usual propositional connectives and equipped with a classical model-theoretic semantics; the constants \top, \perp are in L . V is the set of valuations of L and $M(\alpha)$ is the set of models of $\alpha \in L$. Classical entailment is denoted by \models . Roman letters, p, q, r, \dots denote propositional atoms; Greek letters α, β, \dots stand for arbitrary formulas. We reserve the letter Φ to denote epistemic states.

Definition 1 *Associated with an epistemic state Φ is a total preorder on valuations \preceq_Φ , and a knowledge base $K(\Phi)$. $M_{\preceq_\Phi}(\alpha)$ denotes the minimal models of α in the total preorder on valuations. The*

knowledge base associated with the epistemic state is obtained by considering the minimal models in \preceq_Φ i.e., $M(K(\Phi)) = M_{\preceq_\Phi}(\top)$. To be consistent with the AGM postulates, $K(\Phi)$ may be inconsistent. So, for any Φ , $K(\Phi)$ is either the inconsistent belief set, or the theory generated by the minimal models of \preceq_Φ .

Therefore $K(\Phi)$ is, in a limited sense, independent of \preceq_Φ .

2 The reformulated AGM postulates

In the reformulated postulates below, $*$ and $-$ are belief change operations on epistemic states, not knowledge bases. So $*$ takes an epistemic state and a sentence and produces an epistemic state. For $-$ and $*$ to satisfy the AGM postulates means that they satisfy the reformulated AGM postulates which apply to epistemic states, not knowledge bases. The reformulated AGM postulates guarantee a unique extracted knowledge base when revision or contraction is performed i.e., the lowest level of valuations in the resulting epistemic state is fixed. What is not fixed is how to order the remaining valuations. Note that the object of revision is the epistemic state, but in stating the postulates we specify the form of the knowledge base extracted from the epistemic state. Here are the reformulated AGM postulates (done in much the same style as the reformulation by Darwiche-Pearl). First contraction:

- ($\Phi-1$) $K(\Phi - \alpha) = Cn(K(\Phi - \alpha))$
- ($\Phi-2$) $K(\Phi - \alpha) \subseteq K(\Phi)$
- ($\Phi-3$) If $\alpha \notin K(\Phi)$ then $K(\Phi - \alpha) = K(\Phi)$
- ($\Phi-4$) If $\not\models \alpha$ then $\alpha \notin K(\Phi - \alpha)$
- ($\Phi-5$) If $\alpha \in K(\Phi)$ then $K(\Phi) \subseteq (K(\Phi - \alpha)) + \alpha$
- ($\Phi-6$) If $\alpha \equiv \beta$ then $\Phi - \alpha = \Phi - \beta$
- ($\Phi-7$) $K(\Phi - \alpha) \cap K(\Phi - \beta) \subseteq K(\Phi - (\alpha \wedge \beta))$
- ($\Phi-8$) If $\beta \notin K(\Phi - (\alpha \wedge \beta))$ then $K(\Phi - (\alpha \wedge \beta)) \subseteq K(\Phi) - \beta$

In what follows, we will be particularly interested in the relationship between $K(\Phi * \alpha - \alpha)$ and $K(\Phi)$. We will show that equality between the two sides conflicts with the reformulated AGM postulates but does hold under some conditions.

The intuitions corresponding to the postulates are roughly those of the original AGM postulates. For example, ($\Phi-1$) states that the knowledge base associated with the revised epistemic state is closed under logical consequence. ($\Phi-6$) states that contracting by logically equivalent formulas results in the same epistemic state. This particular postulate highlights a difference between the original AGM postulates and the reformulations above. The original AGM postulate requires the belief set after contraction to be the same after contraction by logically equivalent formulas, whereas we require that if two epistemic states are the same, then contraction by logically equivalent formulas should result in the same epistemic state. This is crucially different from merely requiring that the knowledge base associated with the epistemic state be the same (such a reformulation of the AGM axioms by Darwiche-Pearl is responsible for making their axioms compatible with AGM). Note that we include the recovery axiom above. The following are the reformulated AGM postulates for revision:

- ($\Phi*1$) $K(\Phi * \alpha) = Cn(K(\Phi * \alpha))$
- ($\Phi*2$) $\alpha \in K(\Phi * \alpha)$
- ($\Phi*3$) $K(\Phi * \alpha) \subseteq K(\Phi) + \alpha$
- ($\Phi*4$) If $\neg\alpha \notin K(\Phi)$ then $K(\Phi) + \alpha \subseteq K(\Phi * \alpha)$
- ($\Phi*5$) If $\alpha \equiv \beta$ then $\Phi * \alpha = \Phi * \beta$
- ($\Phi*6$) $\perp \in K(\Phi * \alpha)$ iff $\models \neg\alpha$

- (Φ *7) $K(\Phi * (\alpha \wedge \beta)) \subseteq K(\Phi * \alpha) + \beta$
(Φ *8) If $\neg\beta \notin K(\Phi * \alpha)$ then $K(\Phi * \alpha) + \beta \subseteq K(\Phi * (\alpha \wedge \beta))$

As with the contraction postulates, the intuitions corresponding to the postulates are roughly the same as those underlying the original AGM postulates. For example, (Φ *1) states that the knowledge base associated with the revised epistemic state is closed. (Φ *6) states that an inconsistent knowledge base only results when revising by contradictions (note the modified (Φ *5) postulate as well).

We now list the Darwiche-Pearl postulates for iterated revision [5]. In the four postulates below $*$ is the revision operator, α, β , represent new epistemic inputs and Φ represents an epistemic state.

- (C1) If $\alpha \models \beta$, then $K(\Phi * \beta * \alpha) = K(\Phi * \alpha)$.
(C2) If $\alpha \models \neg\beta$, then $K(\Phi * \beta * \alpha) = K(\Phi * \alpha)$.
(C3) If $K(\Phi * \alpha) \models \beta$, then $K(\Phi * \beta * \alpha) \models \beta$.
(C4) If $K(\Phi * \alpha) \not\models \neg\beta$, then $K(\Phi * \beta * \alpha) \models \beta$.

The following are the semantic versions, with $u, v \in V$:

- (CR1) If $u \in M(\alpha), v \in M(\alpha)$ then $u \preceq_{\Phi} v$ iff $u \preceq_{\Phi * \alpha} v$
(CR2) If $u \in M(\neg\alpha), v \in M(\neg\alpha)$ then $u \preceq_{\Phi} v$ iff $u \preceq_{\Phi * \alpha} v$
(CR3) If $u \in M(\alpha), v \in M(\neg\alpha)$ then $u \prec_{\Phi} v$ only if $u \prec_{\Phi * \alpha} v$
(CR4) If $u \in M(\alpha), v \in M(\neg\alpha)$ then $u \preceq_{\Phi} v$ only if $u \preceq_{\Phi * \alpha} v$

Darwiche and Pearl have shown that, given the reformulated postulates for revision, a precise correspondence obtains between (C*i*) and (CR*i*) above ($i = 1, 2, 3, 4$). The postulate (C1) is stronger than (Φ *7) and (Φ *8) (it implies them); it 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 belief state. Here the *prima facie* connection with recovery is obvious; the basic form of the recovery axiom deals with ‘contract by α and then expand by α ’ while (C2) deals with ‘revise by α and then revise by (effectively) $\neg\alpha$ ’. The latter is clearly stronger. (C3) says that a piece of evidence β should be retained after accommodating more recent evidence α that entails β given the current belief state. (C4) simply says that no epistemic input can act as its own defeater. [2, 11, 6] are amongst those to have critically commented on (C2); [6] shows that it is inconsistent with the original AGM axioms for belief sets (as is the weaker axiom, C2’ proposed in Nayak *et al.* [15]). This last objection, as noted above, is no longer a problem when the postulates are reformulated for epistemic states. For the purposes of this paper, we do not dispute (C1), (C3) and (C4).

3 The new recovery postulates

In this section we provide—additional to the reformulated AGM postulates—some recovery-like postulates and then provide a semantic condition for iterated revision. These additional properties are desirable for iterated revision and cover a variety of situations, ranging from sequences of revisions and contractions by the same formula to sequences of revisions and contractions by a formula and its negation. In particular they describe the conditions under which we can expect stability or minimal loss of beliefs in the original epistemic state. Note that the sequence of belief changes specified reverses the original formulation of the recovery axiom where contraction is followed by expansion. Stating the postulates in this form enables the connection with iterated revision to become clear since it is in the

case of revision followed by contraction that a notion of iterated revision is necessary (in the original formulation of the recovery axiom, expansion is equivalent to revision thus obviating the need for a framework that requires iteration). In the postulates—and our framework in general—we make the assumption that information is received from the same source.

- (R1) $K(\Phi * \alpha - \alpha) \subseteq K(\Phi - \alpha)$
(R2) $\alpha, \neg\alpha \notin K(\Phi)$ implies $K(\Phi) \subseteq K(\Phi * \alpha - \alpha)$
(R3) $\alpha \notin K(\Phi)$ implies $K(\Phi) \subseteq K(\Phi * \alpha * \neg\alpha)$
(R4) $\alpha \in K(\Phi)$ implies $K(\Phi - \alpha) \subseteq K(\Phi * \alpha - \alpha)$

(R1) says that the result of revising an epistemic state and then contracting by the same formula is contained in the knowledge base obtained after simply contracting by the same formula. (R2) says that if neither a formula nor its negation are in the knowledge base associated with an epistemic state then the original base will be contained in that obtained after revision and contraction by the same formula. (R3) says that if a piece of information is not contained in the knowledge base associated with an epistemic state, then a revision by that formula followed by its negation will always include the original knowledge base. (R4) says that if a formula is contained in the original knowledge base then contracting by the same formula will produce a knowledge base that is contained in one obtained by revising and contracting by the same formula.

The following *additional* properties further place conditions on recovery like situations since they compare $K(\Phi * \alpha - \alpha)$, $K(\Phi)$ and $K(\Phi - \alpha)$, but with conditions distinct from those of (R1-R4).

- (R5) $K(\Phi * \alpha - \alpha) \subseteq K(\Phi)$
(R6) $\alpha \notin K(\Phi)$ implies $K(\Phi) \subseteq K(\Phi * \alpha - \alpha * \neg\alpha)$
(R7) $\alpha \in K(\Phi)$ implies $K(\Phi) \subseteq K(\Phi * \alpha - \alpha * \alpha)$
(R8) $\alpha, \neg\alpha \notin K(\Phi)$ implies $K(\Phi - \alpha) \subseteq K(\Phi * \alpha - \alpha)$

(R5) says that the knowledge base obtained by revising by an input and then contracting by it is contained in the knowledge base associated with the original epistemic state. (R6) says that if a belief is not contained in the original knowledge base, then the knowledge base is contained in the result of revising by a formula, contracting it and then revising by its negation. (R7) says that if a belief is contained in the original knowledge base, then that belief will be preserved under a sequence of revisions which begin with revision followed by contraction and then revision. (R8) says that if the original knowledge base is agnostic about a particular belief then contracting by that belief will result in a knowledge base that is contained in one obtained by revising and then contracting by that belief.

Observation 1 Consider operators, $*$ and $-$ that satisfy the reformulated postulates for revision and contraction. Then (R5,6,8) follow from R(1-4); (R7) follows from the reformulated AGM postulates.

Consider the following condition R’: $\neg\alpha \in K(\Phi)$ implies $K(\Phi) \subseteq K(\Phi * \alpha - \alpha)$. This states that if a belief is not contained in the original knowledge base, then the original knowledge base is contained in that obtained after revising and contracting by its negation. However, such a condition contradicts ($\Phi - 2$) and ($\Phi * 2$). Note that the conclusion of (R3) cannot hold if α is in $K(\Phi)$. The reformulated AGM postulates for epistemic states and our additional recovery postulates provide a comprehensive framework for iterated revision which does justice to the intuitions expressed in the original recovery axiom. One of our stated aims is to link up $K(\Phi * \alpha - \alpha)$ and $K(\Phi)$. We do this via (R2) and (R5). Another way to put it: if $\alpha, \neg\alpha \notin K(\Phi)$ then $K(\Phi) = K(\Phi * \alpha - \alpha)$.

If $\neg\alpha \in K(\Phi)$ then AGM prevents $K(\Phi) = K(\Phi * \alpha - \alpha)$. If $\alpha \in K(\Phi)$ then, since $\alpha \notin K(\Phi * \alpha - \alpha)$ by AGM, it is AGM that prevents $K(\Phi) = K(\Phi * \alpha - \alpha)$. We think of C1—with a caveat made for possible weakenings in the future—C3, C4 and R1-R4 as a framework for iterated revision.

3.1 Semantic properties

We now provide semantic conditions for revisions of epistemic states and make explicit the connection between (R1-4) and (C2). The following lay conditions on the positions of valuations by revision and may be considered the semantic counterpart to (R1-4).

$$(S1) \quad M_{\leq_{\Phi}}(\neg\alpha) \subseteq M_{\leq_{\Phi * \alpha}}(\neg\alpha)$$

$$(S2) \quad M_{\leq_{\Phi * \alpha}}(\neg\alpha) \subseteq M_{\leq_{\Phi}}(\neg\alpha)$$

The semantic properties taken together state an equality between the minimal models of $\neg\alpha$ in the epistemic state prior to revision and after revision. (S1) and (S2) taken together state that the minimal models of $\neg\alpha$ remain the same relative to the other models of $\neg\alpha$. For ease of statement of Theorem 1 below, we state these properties as two separate containments rather than the implied equality. Consider the minimal models of $\neg\alpha$ in the total preorder associated with the epistemic state; these might or might not be included in the minimal models of the total preorder itself. After revision by α , the minimal models of the ordering cannot contain any $\neg\alpha$ models. So the minimal models of $\neg\alpha$ are either demoted in the ordering or stay where they are. Whatever be the case, no models of $\neg\alpha$ can be promoted in the ordering to join the old minimal models of $\neg\alpha$ and furthermore, none of the minimal models of $\neg\alpha$ are demoted. Revision by α can increase the plausibility of α and decrease that of $\neg\alpha$; it certainly cannot increase the plausibility of $\neg\alpha$. Remarkably, this simple condition provides all the semantic linkage we need with the syntactic properties (R1-4) stated above. It should be clear that the semantic properties stated above are a weaker version of the (CR2) postulate since in the Darwiche-Pearl framework, which relies on a form of Spohnian conditioning [17], the position of *all* $\neg\alpha$ models is determined in the new epistemic state (via pointwise decrease in their plausibility by one rank after revision by α , thus preserving their relative ordering in the new epistemic state) whereas in our condition, we simply specify the minimal models of $\neg\alpha$ in the new epistemic state. Strengthening these postulates is possible, but possibly counterproductive, and in any case, it is not our present concern.

Theorem 1 *Let $*$ and $-$ be belief change operations on epistemic states satisfying the reformulated AGM postulates.*

1. $*$ and $-$ satisfy (R1) iff $*$ satisfies (S1).
2. $*$ and $-$ satisfy (R2)-(R4) iff $*$ satisfies (S2).

Proof:

1. (S1) follows immediately from (R1). Suppose (S1) and pick a $u \in M(K(\Phi - \alpha))$. If $u \in M(\alpha)$ then $u \in M(K(\Phi * \alpha - \alpha))$ by AGM. If $u \in M(\neg\alpha)$ then $u \in M_{\leq_{\Phi}}(\neg\alpha)$. By (S1), $u \in M_{\leq_{\Phi * \alpha}}(\neg\alpha)$. Therefore $u \in M(K(\Phi * \alpha - \alpha))$.
2. Suppose (S2). Now suppose $\alpha, \neg\alpha \notin K(\Phi)$. Pick a $u \in M(K(\Phi * \alpha - \alpha))$. If $u \in M(\alpha)$ then $u \in M(K(\Phi))$ by AGM. Otherwise $u \in M(K(\Phi))$ by (S2). So (R2) holds. Now suppose $\alpha \notin K(\Phi)$. Pick a $u \in M(K(\Phi * \alpha * \neg\alpha))$. Since $u \in M(\neg\alpha)$ it follows that $u \in M(K(\Phi))$ by (S2). So (R3) holds. Now suppose $\alpha \in K(\Phi)$. Pick a $u \in M(K(\Phi * \alpha - \alpha))$. If $u \in M(\alpha)$ then

$u \in M(K(\Phi))$ by AGM. Otherwise $u \in M(K(\Phi))$ by (S2). So (R4) holds. Conversely, suppose (R2)-(R4). If $\alpha, \neg\alpha \notin K(\Phi)$ then (S2) follows from (R2). If $\neg\alpha \in K(\Phi)$ then (S2) follows from (R3). If $\alpha \in K(\Phi)$ then (S2) follows from (R4). ■

The following shows that the case we were interested in, the relationship between $K(\Phi * \alpha - \alpha) = K(\Phi - \alpha)$ is one of equality in the case when $\neg\alpha$ is not contained in the original knowledge base.

Corollary 1 *From (R1)-(R4) it follows that, if $\neg\alpha \notin K(\Phi)$ then $K(\Phi * \alpha - \alpha) = K(\Phi - \alpha)$.*

Proof: Follows from (S1) and (S2), which state together that $M_{\leq_{\Phi * \alpha}}(\neg\alpha) = M_{\leq_{\Phi}}(\neg\alpha)$. ■

Furthermore, note that since $*$ and $-$ above are operations that satisfy the reformulated AGM postulates, it follows that they satisfy (R5), (R6), (R7) and (R8) as well.

3.2 C2 and the new recovery postulates

The connections between (S1), (S2) and (C2) are interesting ((C1), (C3) and (C4) are not entailed by our postulates). Objections to (C2) often rely on the observation that revising a belief state ψ with a sentence of the form $p_1 \wedge p_2 \wedge \dots \wedge p_n \wedge q$ followed by a revision with $\neg q$ reduces to revision with $\neg q$. Thus the—potentially useful—belief in the conjunct $p_1 \wedge p_2 \wedge \dots \wedge p_n$ is discarded (unless it was believed in the first place) even though it does not in itself contradict $\neg q$. It can be argued that these criticisms of the (C2) postulate are somewhat unfair, since this unintuitive outcome does not follow if revision by $p_1 \wedge p_2 \wedge \dots \wedge p_n \wedge q$ is replaced by a sequence of revisions by each of the conjuncts. One would revise with the full conjunction only if these beliefs were somehow implicitly related. One scenario where this behaviour required by the (C2) postulate appears to be fully justified is when a source provides $p_1 \wedge p_2 \wedge \dots \wedge p_n \wedge q$ as an input, and subsequently changes its mind (thus revising by $\neg q$). In a similar vein, if two consecutive sensor readings contradict each other, it makes more sense to believe the more recent reading, even if the previous reading provided additional information. The (C2) postulate has also been criticized from other perspectives. Cantwell [4] uses a version of the George-the-criminal example to criticize the (C2) postulate. We note that it is possible to argue against Cantwell's criticism along similar lines to our arguments against the Cleopatra example (if the inputs come from the same source, then the outcomes are intuitive, while inputs from different sources would appear as distinct sentences, making the example redundant).

The following example, a variation of the George-the-criminal setting, makes clear that (C2) is too strong, and that (S1) and (S2) are useful alternative weakenings. Assume that we start by believing George is an armed robber—based on information from my friend the police detective. Then my friend tells us that this is incorrect, since no criminal records can be found for George. Subsequently, she corrects her original statement—she did find a criminal dossier on George at police headquarters (it had been misplaced) and given its location, it could have only come off the stack of files for people convicted of illegal gun possession or the stack of convicted shoplifters' dossiers. We must now revise our beliefs with the information that George is not an armed robber, but either a shoplifter or a person convicted of illegal gun possession. We construct below a scenario where the (C2) postulate forces us to

believe that George was convicted of illegal gun possession (clearly too strong given the available evidence even though given the source and our initial beliefs this is more plausible than George being guilty of shoplifting). We let r denote ‘George is an armed robber’, g denote ‘George has been convicted of illegal gun possession’ and s denote ‘George is a convicted shoplifter’ and use c as an abbreviation for ‘George is a criminal’ i.e., $r \vee g \vee s$. Given the propositional language $\{r, g, s\}$, we will represent models as sequences of 0’s and 1’s, representing the valuations of r , g and s respectively (thus 100 represents a model in which r is true and g and s are false). We assume for the sake of explanatory convenience that epistemic states map valuations to natural numbers with the minimal models being identified as those assigned the lowest rank (not necessarily 0)—thus inducing a total preorder on valuations. Let the initial epistemic state Φ_1 be defined as follows:

$$\begin{aligned}\Phi_1(100) &= \Phi_1(101) = \Phi_1(110) = \Phi_1(111) = 0 \\ \Phi_1(010) &= \Phi_1(011) = 1 \\ \Phi_1(000) &= \Phi_1(001) = 2\end{aligned}$$

Observe that, next to the models of r , we believe the models of g to be most plausible, reflecting the intuition that if George is not an armed robber, then the next most likely scenario is where George is in illegal possession of firearms. To satisfy (C2) the epistemic state $\Phi_2 = \Phi_1 * \neg c$ must appear as follows:

$$\begin{aligned}\Phi_2(000) &= 0 \\ \Phi_2(100) &= \Phi_2(101) = \Phi_2(110) = \Phi_2(111) = 1 \\ \Phi_2(010) &= \Phi_2(011) = 2 \\ \Phi_2(001) &= 3\end{aligned}$$

Observe that $g \in K(\Phi_2 * \neg r \wedge (g \vee s))$, i.e., we are forced to believe George has been convicted of illegal gun possession. If we relax (C2) with (S1) and (S2), a permissible outcome of revising Φ_1 by $\neg c$ is the epistemic state Φ'_2 where:

$$\begin{aligned}\Phi'_2(000) &= 0 \\ \Phi'_2(100) &= \Phi'_2(101) = \Phi'_2(110) = \Phi'_2(111) = 1 \\ \Phi'_2(010) &= \Phi'_2(011) = \Phi'_2(001) = 2\end{aligned}$$

Revising with $\neg r \wedge (g \vee s)$ we note that $g \notin K(\Phi'_2 * \neg r \wedge (g \vee s))$. Our example shows that (S1-S2) can handle desirable outcomes disallowed by C2. (S1-S2) respect the initial ordering of the epistemic state to some extent but not dogmatically so. An interesting alternative reading of (S1-S2) is think of them as providing the agent with a form of short-term memory. We can think of (S1-S2) as saying, ‘‘If I have to revise by α , there must be something wrong with the $\neg\alpha$ -worlds. But I do not want to throw away all the information about $\neg\alpha$ that I had previously. So, I will compromise by remembering the best $\neg\alpha$ -worlds after revision. This ensures that if I decide to undo my revision by α , I’ll end up with the same $\neg\alpha$ -worlds.’’

4 Conclusion

We have shown how the intuitions underlying the axiom of recovery can be rescued by paying attention to the assumptions underlying putative counterexamples. We argued that the axiom of recovery places an important rationality constraint on iterated revision, a framework that requires that we think of revision as taking place on epistemic states which encode preferences rather than just flat belief sets. We believe the connection between the axiom of recovery and the (C2)

postulate of Darwiche-Pearl to be an interesting one. For future work it might be interesting to try and obtain a weakened version of the (C1) postulate in a way that is similar to what we have done in this paper. Further work with other proposals for iterated revision such as [3, 18] is also necessary for a full evaluation of our proposal.

REFERENCES

- [1] Carlos Alchourron, Peter Gärdenfors, and David Makinson, ‘On the logic of theory change: Partial meet functions for contraction and revision’, *J. of Symbolic Logic*, **50**, 510–530, (1985).
- [2] Horacio Arlo-Costa and Rohit Parikh, ‘Two place probabilities, beliefs and belief revision: on the foundations of iterative belief kinematics’, in *Proceedings of the Twelfth Amsterdam Colloquium*, (1999).
- [3] Craig Boutilier, ‘Revision sequences and nested conditionals’, in *Proceedings of the Thirteenth International Joint Conference on Artificial Intelligence*, pp. 519–525, (1993).
- [4] John Cantwell, ‘Some logics of iterated belief change’, *Studia Logica*, **63**(1), 49–84, (1999).
- [5] Adnan Darwiche and Judea Pearl, ‘On the logic of iterated belief revision’, *Artificial Intelligence*, **89**, 1–29, (1997).
- [6] Michael Freund and Daniel Lehmann, ‘Belief revision and rational inference’, Technical Report TR-ARP-94-16, The Leibniz Center for Research in Computer Science, Institute of Computer Science, Hebrew University, July 1994, (1994).
- [7] André Fuhrmann, ‘On the modal logic of theory change’, in *The Logic of Theory Change: Workshop, Konstanz, FRG, October 1989, Proceedings*, eds., A. Fuhrmann and M. Morreau, volume 465 of *Lecture Notes in Artificial Intelligence*, pp. 259–281, Berlin, (1991). Springer-Verlag.
- [8] Sven Ove Hansson, ‘Belief contraction without recovery’, *Studia Logica*, **50**, 251–260, (1991).
- [9] Sven-Ove Hansson, ‘Changes of disjunctively closed bases’, *Journal of Logic, Language and Information*, **2**(4), 255–284, (1993).
- [10] Sven Ove Hansson, ‘A survey of non-prioritized belief revision’, *Erkenntnis*, (1999). To appear.
- [11] Daniel Lehmann, ‘Belief revision, revised’, in *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*, pp. 1534–1540, (1995).
- [12] Isaac Levi, *The Fixation of Belief and Its Undoing*, Cambridge University Press, Cambridge, 1991.
- [13] David Makinson, ‘On the status of the postulate of recovery in the logic of theory change’, *Journal of Philosophical Logic*, **16**, 383–394, (1987).
- [14] Thomas Meyer, ‘Basic Infobase Change’, in *Advanced Topics in Artificial Intelligence*, ed., Norman Foo, volume 1747 of *Lecture Notes in Artificial Intelligence*, pp. 156–167, Berlin, (1999). Springer-Verlag.
- [15] A. C. Nayak, N.Y. Foo, M. Pagnucco, and A. Sattar, ‘Changing conditional beliefs unconditionally’, in *Proceedings of the Sixth Conference on Theoretical Aspects of Rationality and Knowledge*, pp. 119–135, De Zeeuwse Stroom, The Netherlands, (1996). Morgan Kaufmann.
- [16] Abhaya Nayak, *Studies in Belief Change*, Ph.D. dissertation, University of Rochester, 1993.
- [17] Wolfgang Spohn, ‘Ordinal conditional functions: A dynamic theory of epistemic states’, in *Causation in Decision: Belief, Change and Statistics: Proceedings of the Irvine Conference on Probability and Causation: Volume II*, eds., William L. Harper and Brian Skyrms, volume 42 of *The University of Western Ontario Series in Philosophy of Science*, pp. 105–134, Dordrecht, (1988). Kluwer Academic Publishers.
- [18] Mary-Anne Williams, ‘Transmutations of knowledge systems’, in *Proceedings of Fourth International Conference on Principles of Knowledge Representation and Reasoning*, eds., J. Doyle, E. Sandewall, and P. Torasso, pp. 619–629. Morgan Kaufmann, (1994).