

Observations on Darwiche and Pearl’s Approach for Iterated Belief Revision

Theofanis Aravanis¹, Pavlos Peppas^{1,2} and Mary-Anne Williams²

¹University of Patras, Greece

²University of Technology Sydney, Australia

{taravanis, pavlos}@upatras.gr, Mary-Anne.Williams@uts.edu.au

Abstract

Notwithstanding the extensive work on iterated belief revision, there is, still, no fully satisfactory solution within the classical AGM paradigm. The seminal work of Darwiche and Pearl (DP approach, for short) remains the most dominant, despite its well-documented shortcomings. In this article, we make further observations on the DP approach. Firstly, we prove that the DP postulates are, in a strong sense, inconsistent with Parikh’s relevance-sensitive axiom (P), extending previous initial conflicts. Immediate consequences of this result are that an entire class of intuitive revision operators, which includes Dalal’s operator, violates the DP postulates, as well as that the Independence postulate and Spohn’s conditionalization are inconsistent with (P). Lastly, we show that the DP postulates allow for more revision policies than the ones that can be captured by identifying belief states with total preorders over possible worlds, a fact implying that a preference ordering (over possible worlds) is an insufficient representation for a belief state.

1 Introduction

The celebrated “AGM paper”, [Alchourrón *et al.*, 1985], laid the foundations of what is now called the *AGM paradigm* for belief revision [Gärdenfors, 1988; Peppas, 2008]. Within this framework, the beliefs of an agent are modelled as a logical theory K , also called a *belief set*, epistemic input is represented as a logical sentence φ , and the revision of K by φ , denoted by $K * \varphi$, is modelled as a function $*$ mapping theories and sentences to theories. Revision functions are constrained by eight postulates, the *AGM postulates for revision*, introduced to capture the notion of *rationality* in the context of belief change. The functions that satisfy these postulates are known as *AGM revision functions*.

One of the main shortcomings of the original AGM paradigm is its lack of any guidelines for *iterated* belief revision — see [Peppas, 2014] for a survey on this issue. Perhaps the most influential work addressing the problem of iterated revision, within the confines of the AGM paradigm, is Darwiche and Pearl’s approach; for short, *DP approach* [Darwiche and Pearl, 1994]. In the DP approach, four new postu-

lates were introduced, namely, the *DP postulates*, to regulate the process of iterated revision.

However, as originally introduced in [Darwiche and Pearl, 1994], not all the DP postulates were consistent with the AGM postulates for revision [Lehmann, 1995]. The inconsistency was later resolved in [Darwiche and Pearl, 1997], by defining revision functions to operate on *belief states*, rather than on belief sets. With the switch to belief states, Darwiche and Pearl modified the AGM postulates for revision accordingly; we shall refer to the functions satisfying these *modified* postulates as *modified AGM revision functions*.

Even so, the DP approach remains controversial. The criticism, typically, comes in the form of counter-examples in which the DP postulates produce unintuitive results; see, for example, [Konieczny and Pérez, 2000; Jin and Thielscher, 2007; Stalnaker, 2009]. One of the latest criticism constitutes the result of [Peppas *et al.*, 2008], that points out a first conflict between each one of the DP postulates and Parikh’s relevance-sensitive axiom (P), introduced in [Parikh, 1999] as an attempt to capture *relevance* in belief revision.

In this work, we make further observations on the DP approach. In particular:

- The main incompatibility result of [Peppas *et al.*, 2008] has been proved assuming a set of propositional variables with *fixed* cardinality; hence, it is confined to *only one type* of modified AGM revision functions. Specifically, Peppas *et al.* proved their result for modified AGM revision functions defined over a propositional language built from only three variables, utilizing a particular theory of the language. Herein, we provide a *fully-fledged* incompatibility result between relevance-sensitive and iterated belief revision, extending the result of [Peppas *et al.*, 2008] for *any* finite, non-empty set of propositional variables, as well as for *any splittable or (non-trivially) confined* theory. Hence, the possibility for modified AGM revision functions of *any type* to satisfy both axiom (P) and the DP postulates is excluded.
- The aforementioned incompatibility results imply significant corollaries for existing belief-change proposals. *Parametrized Difference revision operators* [Peppas and Williams, 2016; 2018], and their specialization *Dalal’s operator* [Dalal, 1988], are incompatible with the DP approach. Moreover, the *Independence* postulate [Jin and

Thielscher, 2007] (that remedies problems with the DP postulates) and the well-known Spohn’s *conditionalization* [Spohn, 1988] (restricted to revision scenarios) are incompatible with axiom (P).

- We discuss the relation between belief states and total preorders over possible worlds, and prove that the *one-to-one correspondence* between them is not sufficient to “cover” all possible revision policies aligned with the DP postulates. This result implies that a preference ordering (over worlds) is an *insufficient* structure for representing a belief state. This “non-reductionist” view was also recently supported in [Booth and Chandler, 2017; 2018], where relevant considerations were provided.

The rest of the paper is structured as follows. In the next section, we fix some notation and terminology. Then, assuming basic knowledge of the AGM paradigm, the DP approach is introduced. The subsequent section presents the fully-fledged incompatibility results between the DP postulates and axiom (P), whereas Section 5 presents their direct consequences. Section 6 studies the relation between belief states and total preorders over possible worlds. The last section is devoted to some concluding remarks.

2 Formal Preliminaries

Throughout this paper, we work with a *finite*, non-empty set of propositional variables \mathcal{P} . We define \mathcal{L} to be the propositional language generated from \mathcal{P} , using the standard Boolean connectives, the special symbol \perp (Falsity), and governed by *classical propositional logic*. The logic is identified by its *consequence operation* Cn and/or by its *inference relation* \models ; see [Alchourrón *et al.*, 1985] for their formal properties.

A sentence $\varphi \in \mathcal{L}$ is *contingent* iff $\not\models \varphi$ and $\not\models \neg\varphi$. For a set of sentences Γ of \mathcal{L} , $Cn(\Gamma)$ denotes the set of all logical consequences of Γ ; i.e., $Cn(\Gamma) = \{\varphi \in \mathcal{L} : \Gamma \models \varphi\}$. We shall write $Cn(\varphi_1, \dots, \varphi_n)$ as an abbreviation of $Cn(\{\varphi_1, \dots, \varphi_n\})$. A *theory* (*belief set*) K of \mathcal{L} is any set of sentences of \mathcal{L} closed under Cn ; i.e., $K = Cn(K)$.

A *literal* is a propositional variable $p \in \mathcal{P}$ or its negation. We define a *possible world* (abbrev. *world*) r to be a consistent set of literals, such that, for any propositional variable $p \in \mathcal{P}$, either $p \in r$ or $\neg p \in r$. Sometimes, we treat a set of literals as a sentence, i.e., the *conjunction* of all its literals, leaving it to the context to resolve any ambiguity. The set of all possible worlds is denoted by \mathbb{M} . For a sentence (set of sentences) φ , $[\varphi]$ is the set of worlds at which φ is true. For a set of worlds V , $th(V)$ is the set of sentences (theory) satisfied by all worlds in V . If $V = \emptyset$, then we define $th(V) = \mathcal{L}$.

Let Q be a (strict) subset of \mathcal{P} . We denote by \mathcal{L}^Q the *sublanguage* of \mathcal{L} defined over Q . If $Q = \emptyset$, we take \mathcal{L}^Q to be the language generated by \perp , and the Boolean connectives. For a sentence x of \mathcal{L} , we denote by \mathcal{L}_x the unique *minimal* sublanguage of \mathcal{L} , within which we can write a sentence that is logically equivalent to x . If x is inconsistent or a tautology, we take \mathcal{L}_x to be \mathcal{L}^\emptyset . Moreover, $\overline{\mathcal{L}_x}$ denotes the *complement* language of \mathcal{L}_x ; that is, the language built from the propositional variables that do not appear in \mathcal{L}_x .

Lastly, some definitions on preorders. A *preorder* over a set V is any reflexive, transitive binary relation in V . The

preorder \preceq is called *total* iff, for all $r, r' \in V$, $r \preceq r'$ or $r' \preceq r$. We shall write $r \prec r'$ iff $r \preceq r'$ and $r' \not\preceq r$. We shall also write $r \approx r'$ iff $r \preceq r'$ and $r' \preceq r$. In addition, for any $X \subseteq V$, by $min(X, \preceq)$ we denote the set $\{r \in X : \text{for all } r' \in X, \text{ if } r' \preceq r, \text{ then } r \preceq r'\}$.

For ease of presentation, we shall consider only *consistent* belief sets, and *contingent* epistemic input.

3 The DP Approach

One of the most influential proposals addressing the problem of iterated belief revision, that comply with the assumptions of the AGM paradigm, is the work of Darwiche and Pearl [Darwiche and Pearl, 1994]. This proposal was later modified in [Darwiche and Pearl, 1997], partially in response to a problem identified in [Lehmann, 1995].¹ In the modified version, Darwiche and Pearl defined revision functions to operate on *belief states* (rather than on belief sets), and reformulated accordingly the AGM postulates for revision — hence, they slightly deviated from the AGM paradigm.² As stated, we shall call a function \circ a *modified AGM revision function* iff it satisfies the *modified AGM postulates for revision*, to differentiate it from the one in the original AGM paradigm. Clearly, \circ is a function that maps a belief state S and a sentence φ of \mathcal{L} to a new belief state $S \circ \varphi$.

Darwiche and Pearl did not fully specify what *exactly* they consider a belief state to be. What was clear from their exposition is that, to each belief state S , corresponds a belief set $B(S)$, and, moreover, it is possible to assign the *same* belief set to *different* belief states. This latter feature of belief states is crucial for resolving the original inconsistency identified by Lehmann. Herein, although this is not required by the DP approach, we shall identify a belief state with a special kind of total preorder over all possible worlds, called *faithful preorder*, unless explicitly stated otherwise.

Definition 1 (Faithfulness, [Katsuno and Mendelzon, 1991; Darwiche and Pearl, 1997]). *For a belief state S , a preorder \preceq_S over the possible worlds of \mathbb{M} is faithful to $B(S)$ iff it is total, and such that the minimal worlds (with respect to \preceq_S) are those satisfying $B(S)$; i.e., $min(\mathbb{M}, \preceq_S) = [B(S)]$.*

Darwiche and Pearl proved that, for any revision function \circ that satisfies the modified AGM postulates for revision, any belief state S and any $\varphi \in \mathcal{L}$, the following condition, denoted by $(\preceq \circ)$, holds: $B(S \circ \varphi) = th(min([\varphi], \preceq_S))$.

Given the aforementioned background, Darwiche and Pearl introduced four additional postulates, known as the *DP postulates*, to regulate the process of iterated revision.

- (DP1) If $\varphi \models \psi$, then $B((S \circ \psi) \circ \varphi) = B(S \circ \varphi)$.
- (DP2) If $\varphi \models \neg\psi$, then $B((S \circ \psi) \circ \varphi) = B(S \circ \varphi)$.
- (DP3) If $\psi \in B(S \circ \varphi)$, then $\psi \in B((S \circ \psi) \circ \varphi)$.
- (DP4) If $\neg\psi \notin B(S \circ \varphi)$, then $\neg\psi \notin B((S \circ \psi) \circ \varphi)$.

There is a *one-to-one correspondence* between (DP1)–(DP4) and the following constraints on possible worlds:

¹Lehmann showed that one of the DP postulates, as originally stated, was inconsistent with the AGM postulates for revision.

²Due to space limitations, the modified AGM postulates for revision have been omitted; see [Darwiche and Pearl, 1997] for details.

- (R1) If $r, r' \in [\varphi]$, then $r \preceq_S r'$ iff $r \preceq_{S \circ \varphi} r'$.
 (R2) If $r, r' \in [\neg\varphi]$, then $r \preceq_S r'$ iff $r \preceq_{S \circ \varphi} r'$.
 (R3) If $r \in [\varphi]$ and $r' \in [\neg\varphi]$, then $r \prec_S r'$ entails $r \prec_{S \circ \varphi} r'$.
 (R4) If $r \in [\varphi]$ and $r' \in [\neg\varphi]$, then $r \preceq_S r'$ entails $r \preceq_{S \circ \varphi} r'$.

4 Relevance and the DP Postulates

Despite relevance and iteration seem to be unrelated at first glance, dealing with different aspects of the revision process, it turns out that *each one* of the DP postulates is, in a *strong sense*, incompatible with Parikh's axiom (P).

4.1 Parikh's Notion of Relevance

Relevant change is not properly addressed in the classical AGM paradigm, as identified by Parikh [Parikh, 1999]. Accordingly, Parikh proposed a new axiom, named (P), to supplement the AGM postulates for revision. Axiom (P) was further analysed in [Peppas *et al.*, 2015], where two different interpretations of it were identified, called the *weak* and the *strong* version of (P).³ The weak version of (P), denoted by (wP), is sufficient for the promised results. Peppas *et al.* introduced (wP) in terms of belief sets. However, it can quite straightforwardly be recast in terms of *belief states*, simply by replacing K with $B(S)$ and $K * \varphi$ with $B(S \circ \varphi)$:

- (wP) If $B(S) = Cn(x, y)$, $\mathcal{L}_x \cap \mathcal{L}_y = \emptyset$, and $\varphi \in \mathcal{L}_x$, then $B(S \circ \varphi) \cap \overline{\mathcal{L}_x} = B(S) \cap \overline{\mathcal{L}_x}$.

Condition (wP) says that the revision of S by φ will not affect any belief content outside \mathcal{L}_x . In order to present the possible-world characterization of (wP), as introduced in [Peppas *et al.*, 2015], we need some additional terminology.⁴

Definition 2 (Diff). *The difference between two worlds w, r of \mathbb{M} , denoted by $Diff(w, r)$, is: $Diff(w, r) = \{p \in \mathcal{P} : w \models p \text{ and } r \models \neg p\} \cup \{p \in \mathcal{P} : w \models \neg p \text{ and } r \models p\}$.*

The definition of *Diff* can be extended to include the difference between a theory of \mathcal{L} and a world of \mathbb{M} . To this end, the subsequent definition is necessary.

Definition 3 (Theory Splitting, [Parikh, 1999]). *Let K be a theory of \mathcal{L} , and let $Q = \{Q_1, \dots, Q_n\}$ be a partition of \mathcal{P} ; i.e., $\bigcup Q = \mathcal{P}$, $Q_i \neq \emptyset$, and $Q_i \cap Q_j = \emptyset$, for all $1 \leq i \neq j \leq n$. The set Q is a K -splitting iff there exist sentences $\varphi_1 \in \mathcal{L}^{Q_1}, \dots, \varphi_n \in \mathcal{L}^{Q_n}$, such that $K = Cn(\varphi_1, \dots, \varphi_n)$.*

For every theory K of \mathcal{L} , there is a unique *finest* K -splitting, denoted by \mathcal{F} ; i.e., one which refines every other K -splitting [Parikh, 1999].⁵

Definition 4 (Extended Diff, [Peppas *et al.*, 2015]). *Let K be a theory of \mathcal{L} , and let \mathcal{F} be the finest K -splitting. The difference between K and a world r of \mathbb{M} , denoted by $Diff(K, r)$, is: $Diff(K, r) = \bigcup \{F_i \in \mathcal{F} : \text{for some } \varphi \in \mathcal{L}^{F_i}, K \models \varphi \text{ and } r \models \neg\varphi\}$.*

³In the same work, it was shown that (P) is consistent with the full set of AGM postulates for revision.

⁴For the epistemic-entrenchment characterization of (wP), refer to [Aravanis *et al.*, 2017].

⁵A partition Q' refines another partition Q iff, for every $Q'_i \in Q'$, there is a $Q_j \in Q$, such that $Q'_i \subseteq Q_j$.

In the special case of a *complete* theory K (i.e., $[K]$ is a singleton), Definition 4 collapses to Definition 2. Having defined the extended *Diff*, the possible-world characterization of (wP) (in terms of belief states) is the following conditions:

- (SQ1) If $Diff(B(S), r) \subset Diff(B(S), r')$ and $Diff(r, r') \cap Diff(B(S), r) = \emptyset$, then $r \prec_S r'$.
 (SQ2) If $Diff(B(S), r) = Diff(B(S), r')$ and $Diff(r, r') \cap Diff(B(S), r) = \emptyset$, then $r \approx_S r'$.

Finally, and before proceeding with the incompatibility results, we introduce the following useful definition:

Definition 5 (Splittable/Confined Theory). *Let K be a theory of \mathcal{L} . Iff, for some contingent sentences $x, y \in \mathcal{L}$, $K = Cn(x, y)$ and $\mathcal{L}_x \cap \mathcal{L}_y = \emptyset$, we shall say that K is *splittable*. Iff, for some contingent sentence $x \in \mathcal{L}$, $K = Cn(x)$ and $\mathcal{L}_x \subset \mathcal{L}$, we shall say that K is *confined* (to \mathcal{L}_x).*

Note that, if K is confined to a sublanguage \mathcal{L}' of \mathcal{L} , it also “splits” between \mathcal{L}' and $\overline{\mathcal{L}'}$; the $\overline{\mathcal{L}'}$ part is trivial. Moreover, there is no sentence $\xi \in \overline{\mathcal{L}'}$, such that $K \models \xi$.

4.2 Main Incompatibility Results

Theorem 1. *Let K be any splittable theory of \mathcal{L} . There exists a belief state S , associated with K , and a sentence φ of \mathcal{L} , such that no modified AGM revision function \circ satisfies (P) at S and $S \circ \varphi$, in addition to any of the postulates (DP1)–(DP4).*

Proof. Let K be any theory of \mathcal{L} , such that, for some contingent sentences $x, y \in \mathcal{L}$, $K = Cn(x, y)$ and $\mathcal{L}_x \cap \mathcal{L}_y = \emptyset$. We will show that there exists a belief state S associated with K , such that, for some sentence φ of \mathcal{L} , the transition from \preceq_S to $\preceq_{S \circ \varphi}$ cannot obey the dictates of (SQ1)–(SQ2), in addition to the dictates of each one of (R1)–(R4).

To this end, consider any belief state S , such that $B(S) = K$, and assume that (SQ1)–(SQ2) and (R1)–(R4) are satisfied. Moreover, suppose that F_x, F_y are two elements of the finest $B(S)$ -splitting, such that F_x, F_y contain propositional variables of $\mathcal{L}_x, \mathcal{L}_y$, respectively.

• Violation of (R1):

Let w be any $B(S)$ -world. Consider a world r'' , such that it disagrees with *all* $B(S)$ -worlds on some propositional variables in F_x and in F_y (i.e., $r'' \cap \mathcal{L}^{F_x} \neq w' \cap \mathcal{L}^{F_x}$ and $r'' \cap \mathcal{L}^{F_y} \neq w' \cap \mathcal{L}^{F_y}$, for all $w' \in [B(S)]$), and it agrees with w on the remaining variables (if any, namely, in case $\mathcal{P} - (F_x \cup F_y) \neq \emptyset$). Then, $Diff(B(S), r'') = F_x \cup F_y$, and thus $r'' \notin [B(S)]$.⁶ Moreover, consider the two worlds r, r' , such that $r = (r'' \cap \mathcal{L}^{F_x}) \cup (w \cap \mathcal{L}^{\mathcal{P}-F_x})$ and $r' = (r'' \cap \mathcal{L}^{F_y}) \cup (w \cap \mathcal{L}^{\mathcal{P}-F_y})$ (see Figure 1).⁷

⁶To see that $Diff(B(S), r'')$ contains the propositional variables of F_x , observe that, by the choice of r'' , there is a sentence $\xi \in \mathcal{L}^{F_x}$ (i.e., $\xi = r'' \cap \mathcal{L}^{F_x}$), such that $K \models \neg\xi$ and $r'' \models \xi$. The same applies for F_y . To see that $Diff(B(S), r'')$ contains *only* the propositional variables of F_x and F_y , suppose towards contradiction that this is not the case (of course, in case $\mathcal{P} - (F_x \cup F_y) \neq \emptyset$). Then, there is a sentence $\zeta \in \mathcal{L}^{\mathcal{P}-(F_x \cup F_y)}$, such that $K \models \zeta$ and $r'' \models \neg\zeta$. This again entails that $w \models \zeta$, contradicting the fact that w and r'' agree on the variables outside $F_x \cup F_y$.

⁷It is not hard to verify that the worlds r, r', r'' indeed exist.

Then, $\text{Diff}(B(S), r) = F_x$, $\text{Diff}(B(S), r') = F_y$ (see Footnote 6), and thus $r, r' \notin [B(S)]$.

Therefore, we derive that $\text{Diff}(B(S), r) \subset \text{Diff}(B(S), r'')$ and $\text{Diff}(r, r'') \cap \text{Diff}(B(S), r) = \emptyset$, and $\text{Diff}(B(S), r') \subset \text{Diff}(B(S), r'')$ and $\text{Diff}(r', r'') \cap \text{Diff}(B(S), r') = \emptyset$. Then from (SQ1), it follows that $r \prec_S r''$ and $r' \prec_S r''$. Since $\text{Diff}(B(S), r) \cap \text{Diff}(B(S), r') = \emptyset$, (SQ1)–(SQ2) place no constraints on the relative order of r and r' , with respect to \preceq_S (these two worlds are *Diff*-incomparable).

Now, consider the *particular* belief state, and call this S from now on, where it holds that $r \prec_S r' \prec_S r''$. Define φ to be the sentence $\varphi = r \vee r' \vee r''$. According to the definition of \preceq_S , there is *only one* \preceq_S -minimal φ -world, namely r ; therefore from $(\preceq \circ)$, it follows that $[B(S \circ \varphi)] = \{r\}$. Then, again by the choice of r', r'' , we derive that $\text{Diff}(B(S \circ \varphi), r'') = \text{Diff}(r, r'') \subset \text{Diff}(B(S \circ \varphi), r')$. Hence, no matter what the new faithful preorder $\preceq_{S \circ \varphi}$ is, as long as it satisfies (SQ1), it holds that $r'' \prec_{S \circ \varphi} r'$.

However, given that $r', r'' \in [\varphi]$ and $r' \preceq_S r''$, (R1) entails that $r'' \preceq_{S \circ \varphi} r'$. Contradiction.

- **Violation of (R2):**

Consider the three worlds $r, r', r'' \notin [B(S)]$ of the previous case. As stated, the fact that $r \prec_S r''$ and $r' \prec_S r''$ holds for *any* belief state S associated with K . Define φ to be the sentence $\varphi = r$. Therefore from $(\preceq \circ)$, it follows that $[B(S \circ \varphi)] = \{r\}$. Then, we derive that $\text{Diff}(B(S \circ \varphi), r'') = \text{Diff}(r, r'') \subset \text{Diff}(B(S \circ \varphi), r')$. Hence, no matter what the new faithful preorder $\preceq_{S \circ \varphi}$ is, as long as it satisfies (SQ1), it holds that $r'' \prec_{S \circ \varphi} r'$.

However, given that $r', r'' \in [\neg\varphi]$ and $r' \preceq_S r''$, (R2) entails that $r'' \preceq_{S \circ \varphi} r'$. Contradiction.

- **Violation of (R3) and (R4):**

Assume the aforementioned three worlds $r, r', r'' \notin [B(S)]$, and consider again the *particular* belief state S where $r \prec_S r' \prec_S r''$. Defining $\varphi = r \vee r'$, we can derive the desired contradictions, with a totally analogous reasoning. ■

A stronger incompatibility result concerning postulate (DP2) can be established as well. First however, the following definition regarding confined theories is introduced.

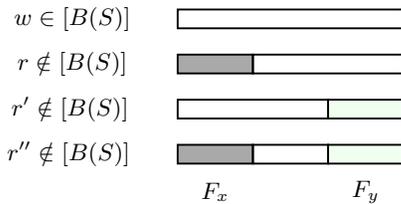


Figure 1: Abstract visualization of the worlds appearing in the proof of Theorem 1. Same color indicates identical propositional variables between worlds.

Definition 6 (Non-Trivially Confined Theory). *We shall say that a theory K of \mathcal{L} is non-trivially confined iff K is confined to a sublanguage \mathcal{L}' of \mathcal{L} , and, moreover, \mathcal{L}' contains at least two propositional variables.*

Clearly, if a theory of \mathcal{L} is non-trivially confined, then the set \mathcal{P} contains *at least three* propositional variables.

Theorem 2. *Let K be any splittable or non-trivially confined theory of \mathcal{L} , and let S be any belief state, associated with K . There exists a sentence φ of \mathcal{L} , such that no modified AGM revision function \circ satisfies (P) at S and $S \circ \varphi$, in addition to postulate (DP2).*

Proof. The proof is divided into two parts. The first part corresponds to the case of any *splittable* theory K , such that, for some contingent sentences $x, y \in \mathcal{L}$, $K = \text{Cn}(x, y)$ and $\mathcal{L}_x \cap \mathcal{L}_y = \emptyset$. The second part corresponds to the case of any *non-trivially confined* theory K , such that, for some contingent sentence $x \in \mathcal{L}$, $K = \text{Cn}(x)$ and $\mathcal{L}_x \subset \mathcal{L}$.

- **Part I: Splittable Theory — $K = \text{Cn}(x, y)$**

This part of the proof is identical to the proof of Theorem 1, concerning the violation of (R2), as the fact that $r \prec_S r''$ and $r' \prec_S r''$ holds for *any* belief state S associated with any splittable theory K .

- **Part II: Non-Trivially Confined Theory — $K = \text{Cn}(x)$**

We will show that, for *any* belief state S associated with K , there exists a sentence φ of \mathcal{L} , such that the transition from \preceq_S to $\preceq_{S \circ \varphi}$ cannot obey the dictates of (SQ1)–(SQ2), *in addition to the dictates of (R2)*.

To this end, consider *any* belief state S , such that $B(S) = K$, and assume that (SQ1)–(SQ2) and (R2) are satisfied. Since $B(S)$ is *non-trivially confined* to \mathcal{L}_x , it follows that $\overline{\mathcal{L}_x}$ contains *at least two* propositional variables, and, moreover, there exist *at least four* worlds outside $[B(S)]$, such that they agree on *all* propositional variables in \mathcal{L}_x .

Let r be a world outside $[B(S)]$. Consider the world $r' \notin [B(S)]$, such that it agrees on all propositional variables with r , except from *one* variable of $\overline{\mathcal{L}_x}$. Moreover, consider the world $r'' \notin [B(S)]$, such that it agrees on all propositional variables with r' , except from *one* variable of $\overline{\mathcal{L}_x}$, and $r'' \neq r$ (see Figure 2).⁸ Then, by the choice of r, r', r'' (that they are outside $[B(S)]$ and agree on all propositional variables in \mathcal{L}_x) and since theory $B(S)$ knows nothing about $\overline{\mathcal{L}_x}$ (i.e., there is no sentence $\xi \in \overline{\mathcal{L}_x}$, such that $B(S) \models \xi$), we have that $\emptyset \neq \text{Diff}(B(S), r) = \text{Diff}(B(S), r') = \text{Diff}(B(S), r'') \subseteq \mathcal{L}_x$, $\text{Diff}(r, r') \cap \text{Diff}(B(S), r) = \emptyset$, and $\text{Diff}(r', r'') \cap \text{Diff}(B(S), r') = \emptyset$. Hence from (SQ2), $r \approx_S r' \approx_S r''$.

Now, define φ to be the sentence $\varphi = r$. Therefore from $(\preceq \circ)$, it follows that $[B(S \circ \varphi)] = \{r\}$. Then, by the choice of r, r', r'' , we derive that $\text{Diff}(B(S \circ \varphi), r') = \text{Diff}(r, r') \subset \text{Diff}(B(S \circ \varphi), r'')$. Hence, no matter what the new faithful preorder $\preceq_{S \circ \varphi}$ is, as long as it satisfies (SQ1), it holds that $r' \prec_{S \circ \varphi} r''$.

However, given that $r', r'' \in [\neg\varphi]$ and $r' \approx_S r''$, (R2) entails that $r' \approx_{S \circ \varphi} r''$. Contradiction. ■

⁸It is not hard to verify that the worlds r, r', r'' indeed exist.

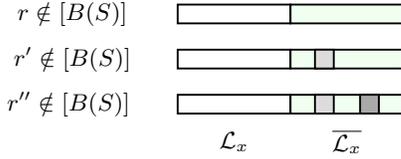


Figure 2: Abstract visualization of the worlds appearing in the proof of Theorem 2 (Part II). Same color indicates identical propositional variables between worlds.

Theorems 1 and 2, essentially, entail that there exists no modified AGM revision function *of any type*, satisfying Parikh’s relevance-sensitive axiom (P) *in addition* to the DP postulates.

5 Corollaries

The incompatibility results of the previous section affect other existing belief-revision proposals.

5.1 Parametrized Difference Revision Operators

Parametrized Difference revision operators (abbrev. *PD operators*) is an important family of operators, recently introduced in [Peppas and Williams, 2016; 2018] as a generalization of Dalal’s concrete revision operator [Dalal, 1988]. It was shown in [Peppas and Williams, 2016] that PD operators satisfy (wP).⁹ Thus, in view of Theorems 1 and 2, we obtain immediately the following two corollaries:¹⁰

Corollary 1. *Parametrized Difference revision operators are inconsistent with the DP postulates.*

Corollary 2. *Dalal’s revision operator is inconsistent with the DP postulates.*

PD operators, as well as Dalal’s operator, are based on a simple and intuitive construction, and they are the *only* among a series of popular operators (including the ones in [Borgida, 1985], [Winslett, 1988] and [Satoh, 1988]) that satisfy the full set of (modified) AGM postulates for revision. Obviously, having such “well-behaved” revision operators violating the DP postulates is a major drawback.

5.2 The Independence Postulate

The *Independence* postulate (Ind) was introduced to remedy problems with the DP ones [Jin and Thielscher, 2007].

(Ind) If $\neg\psi \notin B(S \circ \varphi)$, then $\psi \in B((S \circ \psi) \circ \varphi)$.

The possible-world characterization of (Ind) is as follows:

(IndR) If $r \in [\varphi]$ and $r' \in [\neg\varphi]$, then $r \preceq_S r'$ entails $r \prec_{S \circ \varphi} r'$.

It has been shown in [Jin and Thielscher, 2007] that (Ind) implies both (DP3) and (DP4); hence, the following corollary is an immediate consequence of Theorem 1.

Corollary 3. *Parikh’s axiom (P) is inconsistent with the Independence postulate (Ind).*

⁹In [Aravanis *et al.*, 2019], it was shown that PD operators satisfy strong (P) as well.

¹⁰Although PD operators and Dalal’s operator were introduced in terms of belief sets, they can *straightforwardly* be recast in terms of belief states, as it was done for conditions (wP) and (SQ1)–(SQ2).

5.3 Spohn’s Conditionalization

Spohn uses a *quantitative* structure for representing a belief state, related to a belief set. He calls this structure *ordinal conditional function* (OCF) [Spohn, 1988].

Definition 7 (OCF, [Spohn, 1988]). *An OCF κ is a function from the set \mathbb{M} of possible worlds to the class of ordinals, such that at least one world is assigned the ordinal 0.*

Intuitively, κ assigns a plausibility grading to worlds; the larger $\kappa(r)$ is for a world r , the less plausible r is. This plausibility grading can be extended to sentences; i.e., for any contingent sentence $\varphi \in \mathcal{L}$, $\kappa(\varphi) = \min(\{\kappa(r) : r \in [\varphi]\})$. Lastly, the most plausible worlds define the belief set $B(\kappa)$ that κ is related to; i.e., $B(\kappa) = th(\{r \in \mathbb{M} : \kappa(r) = 0\})$.

Given an OCF κ , the new information φ , and the *degree of firmness* $d \geq 0$ with which φ is incorporated into the new OCF $\kappa \bullet \langle \varphi, d \rangle$, Spohn’s *conditionalization* is as follows:

$$\kappa \bullet \langle \varphi, d \rangle(r) = \begin{cases} \kappa(r) - \kappa(\varphi) & \text{if } r \in [\varphi] \\ \kappa(r) - \kappa(\neg\varphi) + d & \text{if } r \in [\neg\varphi] \end{cases}$$

Darwiche and Pearl showed that (DP1)–(DP4) are satisfied by Spohn’s proposal, restricted to the principal case of *revision* scenarios (where $\kappa(\varphi) > 0$, $\kappa(\neg\varphi) = 0$ and $d > 0$) [Darwiche and Pearl, 1997]. Roughly speaking, this was achieved by constructing an OCF-based revision operator, herein denoted by \diamond , and proving that \diamond satisfies (DP1)–(DP4) (in addition to the modified AGM postulates for revision).¹¹

Therefore, in view of Theorems 1 and 2, we obtain immediately the following result, that indicates the inconsistency between Spohn’s conditionalization (as expressed by means of \diamond) and Parikh’s relevance-sensitive axiom (P).

Corollary 4. *Revision operator \diamond does not satisfy axiom (P).*

6 Belief States and Total Preorders

A belief state represents the current beliefs of an agent. Furthermore, it contains *preference* information which guides the revision process. This information is usually represented by a total preorder over all possible worlds. In this section, we study the relation between belief states and total preorders.

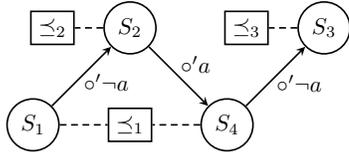
With the assumption that each total preorder of the language is associated with *exactly one* belief state, there is a set Σ of belief states, such that it contains belief states that are in a *one-to-one correspondence* with the preorders. We shall say that such a set Σ is a *simple* set of belief states.

On the other hand, if each preorder is associated with *finitely many* belief states, then more belief states arise (since two *distinct* belief states could be associated with the *same* preorder), which are contained in a “wider” (and obviously not simple) set Σ' of belief states, such that $|\Sigma| < |\Sigma'|$.

Clearly then, there are two sets R and R' of modified AGM revision functions satisfying the DP postulates, and defined over the domains Σ and Σ' , respectively.

Subsequently, we show that there exist revision policies encoded in the revision functions of R' (hence, aligned with the DP postulates), that cannot be “captured” by means of the revision functions of R . This is accomplished in Theorem 3, utilizing the notion of “*simulation*” defined below.

¹¹See Theorem 14 in [Darwiche and Pearl, 1997] for more details.


 Figure 3: Transitions between belief states via o' .

Definition 8 (Simulation). Let Σ be a simple set of belief states, Σ' be a “wider” set of belief states such that $|\Sigma| < |\Sigma'|$, and \circ, o' be two modified AGM revision functions defined over Σ, Σ' , respectively. We shall say that \circ simulates o' iff there exists a function f from Σ' to Σ , such that, for every belief state $S' \in \Sigma'$ and for all sequences of sentences $\varphi_1, \dots, \varphi_n$ of \mathcal{L} , it holds that $B((S' \circ' \varphi_1) \circ' \dots \circ' \varphi_n) = B((f(S') \circ \varphi_1) \circ \dots \circ \varphi_n)$.

Informally, the function \circ simulates the function o' iff, for every belief state in Σ' (the input of o'), there exists a belief state in Σ (the input of \circ), such that, for any sequence of sentences, \circ and o' present the same behaviour (with respect to the belief content of the resulting belief states).

Theorem 3. Let Σ be a simple set of belief states, and let Σ' be a “wider” set of belief states, such that $|\Sigma| < |\Sigma'|$. There exists a modified AGM revision function o' satisfying the DP postulates, defined over Σ' , that cannot be simulated by any modified AGM revision function \circ satisfying the DP postulates, defined over Σ .

Proof. Assume that $\mathcal{P} = \{a, b, c\}$, and that $\preceq_1, \preceq_2, \preceq_3$ are the following three total preorders over the worlds of \mathbb{M} :¹²

$$\begin{array}{ccccccc} abc & \preceq_1 & \bar{a}bc & \preceq_1 & \bar{a}\bar{b}c & \preceq_1 & \bar{a}\bar{b}\bar{c} \\ & & \bar{a}b\bar{c} & \preceq_1 & \bar{a}b\bar{c} & \preceq_1 & \bar{a}\bar{b}\bar{c} \\ & & ab\bar{c} & & a\bar{b}\bar{c} & & \bar{a}\bar{b}\bar{c} \\ \\ \bar{a}bc & \preceq_2 & \bar{a}\bar{b}c & \preceq_2 & \bar{a}\bar{b}\bar{c} & \preceq_2 & \bar{a}\bar{b}\bar{c} \\ & & \bar{a}b\bar{c} & \preceq_2 & \bar{a}b\bar{c} & \preceq_2 & \bar{a}\bar{b}\bar{c} \\ & & abc & & abc & & \bar{a}\bar{b}\bar{c} \\ \\ \bar{a}bc & \preceq_3 & \bar{a}\bar{b}\bar{c} & \preceq_3 & \bar{a}b\bar{c} & \preceq_3 & \bar{a}\bar{b}\bar{c} \\ & & \bar{a}b\bar{c} & \preceq_3 & \bar{a}b\bar{c} & \preceq_3 & \bar{a}\bar{b}\bar{c} \\ & & abc & \preceq_3 & abc & \preceq_3 & \bar{a}\bar{b}\bar{c} \\ & & \bar{a}b\bar{c} & \preceq_3 & \bar{a}b\bar{c} & \preceq_3 & \bar{a}\bar{b}\bar{c} \end{array}$$

Consider four belief states S_1, S_2, S_3 and S_4 of Σ' . We assign at S_1 and S_4 the same preorder \preceq_1 , at S_2 the preorder \preceq_2 , and at S_3 the preorder \preceq_3 . For simplicity, we assume that each one of the remaining preorders over \mathbb{M} is associated with exactly one belief state. Clearly then, Σ' is not a simple set of belief states.

Suppose now that, for a modified AGM revision function o' , defined over Σ' , it holds that $S_1 \circ' \neg a = S_2$, $S_2 \circ' a = S_4$, and $S_4 \circ' \neg a = S_3$, as depicted in Figure 3.

Observe that these three transitions satisfy (DP1)–(DP4). Moreover, observe that $B(\dots((S_1 \circ' \neg a) \circ' a) \circ' \neg a) \circ' (a \vee \neg b \vee \neg c) = B(\dots((S_2 \circ' a) \circ' \neg a) \circ' (a \vee \neg b \vee \neg c)) = B((S_4 \circ'$

$$\neg a) \circ' (a \vee \neg b \vee \neg c) = B(S_3 \circ' (a \vee \neg b \vee \neg c)) = Cn((\neg a \wedge \neg b \wedge c) \vee (\neg a \wedge b \wedge \neg c)).$$

Next, we show that there is no modified AGM revision function satisfying the DP postulates, defined over the simple set of belief states Σ , that simulates o' .

First, let us denote by Q_1, Q_2, Q_3 the belief states of Σ that correspond to the preorders $\preceq_1, \preceq_2, \preceq_3$, respectively.

Now, assume, towards contradiction, that there exists a modified AGM revision function \circ satisfying (DP1)–(DP4), defined over Σ , that simulates o' . It is not hard to verify that, in order for \circ to simulate o' , it must hold that $Q_1 \circ \neg a = Q_2$ and $Q_2 \circ a = Q_1$. Then, $B(\dots((Q_1 \circ \neg a) \circ a) \circ \neg a) \circ (a \vee \neg b \vee \neg c) = B(\dots((Q_2 \circ a) \circ \neg a) \circ (a \vee \neg b \vee \neg c)) = B((Q_1 \circ \neg a) \circ (a \vee \neg b \vee \neg c)) = B(Q_2 \circ (a \vee \neg b \vee \neg c)) = Cn((\neg a \wedge \neg b \wedge c) \vee (\neg a \wedge b \wedge \neg c) \vee (a \wedge b \wedge c)) \neq B(\dots((S_1 \circ' \neg a) \circ' a) \circ' \neg a) \circ' (a \vee \neg b \vee \neg c)$, which is clearly a contradiction. ■

Theorem 3 indicates that a total preorder over worlds is not a sufficient representation for a belief state, in order to achieve “full coverage” of DP-compliant revision policies. This “non-reductionist” thesis was also recently supported in [Booth and Chandler, 2017; 2018], where it is argued that “a preference ordering over the set of possible worlds provides insufficient structure to represent an agent’s commitments to policies of iterated revision” [Booth and Chandler, 2017, p. 413].

7 Conclusion

In this article, important observations on the DP approach were pointed out. At first, we generalized the conflicts between relevance-sensitive and iterated belief revision, proving that there is no modified AGM revision function of any type, satisfying axiom (P) in addition to the DP postulates. These fully-fledged incompatibilities imply significant corollaries; PD operators and Dalal’s operator are incompatible with the DP approach, as well as the Independence postulate and Spohn’s conditionalization (restricted to revision scenarios) are incompatible with axiom (P). Lastly, we shed more light on the nature of belief states, by showing that a preference ordering (over possible worlds) is an insufficient structure for representing a belief state, as the DP postulates allow for more revision policies than the ones that can be captured by identifying belief states with total preorders over worlds.

Arguably, iteration and relevance play a central role in the revision process. The established results show that the dominant models for these two aspects of belief revision are in a deep conflict. Undoubtedly however, any comprehensive formal framework for belief change would need to consistently combine iterated and relevance-sensitive belief revision.

Finding ways to reconcile iteration and relevance — either by weakening the involved postulates or by following an entirely new approach — as well as identifying the exact nature of a belief state constitute imperative future tasks.

Acknowledgements

We are grateful to the anonymous reviewers for their insightful comments on this work.

¹²For the sake of readability, possible worlds are represented as sequences (rather than sets) of literals, and the negation of a propositional variable p is represented as \bar{p} (instead of $\neg p$).

References

- [Alchourrón *et al.*, 1985] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.
- [Aravanis *et al.*, 2017] Theofanis Aravanis, Pavlos Peppas, and Mary-Anne Williams. Epistemic-entrenchment characterization of Parikh’s axiom. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence, IJCAI 2017*, pages 772–778, 2017.
- [Aravanis *et al.*, 2019] Theofanis Aravanis, Pavlos Peppas, and Mary-Anne Williams. An investigation of parametrized difference revision operators. *Annals of Mathematics and Artificial Intelligence*, 2019.
- [Booth and Chandler, 2017] Richard Booth and Jake Chandler. The irreducibility of iterated to single revision. *Journal of Philosophical Logic*, 46:405–418, 2017.
- [Booth and Chandler, 2018] Richard Booth and Jake Chandler. On strengthening the logic of iterated belief revision: Proper ordinal interval operators. In *Proceedings of the 16th International Conference on Principles of Knowledge Representation and Reasoning, KR 2018*, pages 210–219, 2018.
- [Borgida, 1985] Alexander Borgida. Language features for flexible handling of exceptions in information systems. *ACM Transactions on Database Systems*, 10(4):563–603, 1985.
- [Dalal, 1988] Mukesh Dalal. Investigations into theory of knowledge base revision: Preliminary report. In *Proceedings of the 7th National Conference of the American Association for Artificial Intelligence, AAAI 1988*, pages 475–479, 1988.
- [Darwiche and Pearl, 1994] Adnan Darwiche and Judea Pearl. On the logic of iterated belief revision. In *Proceedings of the 5th Conference on Theoretical Aspects of Reasoning About Knowledge, TARK 1994*, pages 5–23, Pacific Grove, California, 1994. Morgan Kaufmann.
- [Darwiche and Pearl, 1997] Adnan Darwiche and Judea Pearl. On the logic of iterated belief revision. *Artificial Intelligence*, 89:1–29, 1997.
- [Gärdenfors, 1988] Peter Gärdenfors. *Knowledge in Flux – Modeling the Dynamics of Epistemic States*. MIT Press, Cambridge, Massachusetts, 1988.
- [Jin and Thielscher, 2007] Yi Jin and Michael Thielscher. Iterated belief revision, revised. *Artificial Intelligence*, 171:1–18, 2007.
- [Katsuno and Mendelzon, 1991] Hirofumi Katsuno and Alberto Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52(3):263–294, 1991.
- [Konieczny and Pérez, 2000] Sébastien Konieczny and Ramón Pino Pérez. A framework for iterated revision. *Journal of Applied Non-Classical Logics*, 10, 2000.
- [Lehmann, 1995] Daniel Lehmann. Belief revision, revised. In *Proceedings of the 14th International Joint Conference on Artificial Intelligence, IJCAI 1995*, pages 1534–1540, Montreal, Quebec, 1995. Morgan Kaufmann Publishers Inc.
- [Parikh, 1999] Rohit Parikh. Beliefs, belief revision, and splitting languages. In Lawrence S. Moss, Jonathan Ginzburg, and Maarten de Rijke, editors, *Logic, Language and Computation*, volume 2, pages 266–278. CSLI Publications, 1999.
- [Peppas and Williams, 2016] Pavlos Peppas and Mary-Anne Williams. Kinetic consistency and relevance in belief revision. In *Proceedings of the 15th European Conference on Logics in Artificial Intelligence, JELIA 2016*, pages 401–414. Springer International Publishing, 2016.
- [Peppas and Williams, 2018] Pavlos Peppas and Mary-Anne Williams. Parametrised difference revision. In *Proceedings of the 16th International Conference on Principles of Knowledge Representation and Reasoning, KR 2018*, pages 277–286, 2018.
- [Peppas *et al.*, 2008] Pavlos Peppas, Anastasios Michael Fotinopoulos, and Stella Seremetaki. Conflicts between relevance-sensitive and iterated belief revision. In *Proceedings of the 18th European Conference on Artificial Intelligence, ECAI 2008*, pages 85–88, Patras, Greece, 2008. IOS Press.
- [Peppas *et al.*, 2015] Pavlos Peppas, Mary-Anne Williams, Samir Chopra, and Norman Foo. Relevance in belief revision. *Artificial Intelligence*, 229:126–138, December 2015.
- [Peppas, 2008] Pavlos Peppas. Belief revision. In Frank van Harmelen, Vladimir Lifschitz, and Bruce Porter, editors, *Handbook of Knowledge Representation*, pages 317–359. Elsevier Science, 2008.
- [Peppas, 2014] Pavlos Peppas. A panorama of iterated revision. In Sven O. Hansson, editor, *David Makinson on Classical Methods for Non-Classical Problems*, pages 71–94. Springer Netherlands, 2014.
- [Satoh, 1988] Ken Satoh. Nonmonotonic reasoning by minimal belief revision. In *Proceedings of the International Conference on Fifth Generation Computer Systems*, pages 455–462. Springer-Verlag, Tokyo, 1988.
- [Spohn, 1988] Wolfgang Spohn. Ordinal conditional functions: A dynamic theory of epistemic states. In William L. Harper and Brian Skyrms, editors, *Causation in Decision, Belief Change, and Statistics*, volume 42 of *The University of Western Ontario Series in Philosophy of Science*, pages 105–134. Springer Netherlands, 1988.
- [Stalnaker, 2009] Robert Stalnaker. Iterated belief revision. *Erkenntnis*, 70(2):189–209, 2009.
- [Winslett, 1988] Marianne Winslett. Reasoning about action using a possible models approach. In *Proceedings of the 7th National Conference of the American Association for Artificial Intelligence, AAAI 1988*, pages 89–93, 1988.