Iterated Belief Revision, Revised

Yi Jin and **Michael Thielscher**
Department of Computer Science
Dresden University of Technology
{yijin, mit}@inf.tu-dresden.de

Abstract

The AGM postulates for belief revision, augmented by the DP postulates for iterated belief revision, provide generally accepted criteria for the design of operators by which intelligent agents adapt their beliefs incrementally to new information. These postulates alone, however, are too permissive: They support operators by which all newly acquired information is canceled as soon as an agent learns a fact that contradicts some of its current beliefs. In this paper, we present a formal analysis of the deficiency of the DP postulates, and we show how to solve the problem by an additional postulate of *independence*. We give a representation theorem for this postulate and prove that it is compatible with AGM and DP.

1 Introduction

Belief revision is the process of changing the beliefs of an agent to accommodate new, more precise, or more reliable evidence that is possibly inconsistent with the existing beliefs. In situations where the new evidence is consistent with the existing beliefs, the two can just be merged; we call this *mild revision*. More interesting and complicated are situations where the evidence conflicts with the prior beliefs, in which case the agent needs to remove some of its currently held beliefs in order to accommodate the new evidence. This kind of revision is referred to as *severe revision*.

In the literature, the classical approach of belief revision is the AGM framework [Alchourrón *et al.*, 1985; Gärdenfors and Makinson, 1988; Gärdenfors, 1988]. Given an underlying logic language $\mathcal{L}$, the beliefs of an agent are represented by a set of sentences in $\mathcal{L}$ (known as *belief set*) which is closed under logical consequence. New evidence is also a sentence in $\mathcal{L}$, and a belief revision operator maps the current belief set and the new evidence to a revised belief set. To provide general design criteria for belief revision operators, Alchourrón, Gärdenfors, and Makinson (AGM) have developed a set of postulates [Alchourrón *et al.*, 1985]. The guiding principle is that of *economy of information* or *minimal change* of belief sets, which means not to give up currently held beliefs and not to generate new beliefs unless necessary.

For the incremental adaptation of beliefs, these postulates proved to be too weak [Darwiche and Pearl, 1994; 1997]. This has led to the development of additional postulates for iterated belief revision by Darwiche and Pearl (DP), among others.

Still, however, the two sets of postulates are too permissive in that they support belief revision operators which assume arbitrary dependencies among the pieces of information which an agent acquires along its way. A universal dependency among all new evidences has a drastic effect when the agent makes an observation which contradicts its currently held beliefs: The agent has to cancel everything it has learned up to this point [Nayak *et al.*, 1996; 2003]. In this paper, we first give a formal analysis of this problem of implicit dependence, and then we present, as a solution, an *Independence* postulate for iterated belief revision. We give a representation theorem for our postulate and prove its consistency by defining a concrete belief revision operator. We also contrast the postulate of independence to the so-called Recalcitrance postulate of [Nayak *et al.*, 1996; 2003] and argue that the latter is too strict in that it rejects reasonable belief revision operators.

The rest of the paper is organized as follows. In the next section, we recall the classical AGM approach in a propositional setting as formulated by [Katsuno and Mendelzon, 1991], followed by the approach of [Darwiche and Pearl, 1994] for iterated belief revision. In Section 3, we formally analyze the problem of the DP postulates to be overly permissive. In Section 4, we present an additional postulate to overcome this deficiency. We give a representation theorem for the postulate along with a concrete revision operator. We conclude in Section 5 with a detailed comparison to related work.

2 Preliminaries

It has been observed by many researchers that belief sets alone are not sufficient to determine a unique strategy for belief revision [Gärdenfors and Makinson, 1988; Spohn, 1988]. Any concrete belief revision operator requires additional information concerning the firmness of different beliefs to determine the revision strategy. In particular, this (extra-logical) information should uniquely determine a set of *conditional beliefs*: An agent is said to hold a conditional belief $\alpha \gg \beta$ (with α, β sentences in $\mathcal{L}$) precisely when it will believe β after a revision with α [Gärdenfors, 1988; Boutilier, 1993]. The *Triviality Theorem* of [Gärdenfors and Makinson, 1988] shows that using the AGM postulates it is improper to include conditional beliefs into the belief sets. As a consequence, we need to distinguish a belief set (referred to as *propositional beliefs*) from a *belief state* (also called *epistemic state*). The latter contains, in addition to its belief set, the conditional beliefs which determine the revision strategy.

2.1 (Modified) KM Postulates

Katsuno and Mendelzon (KM) rephrased the AGM postulates for the propositional setting [Katsuno and Mendelzon, 1991]. The beliefs of an agent are represented by a sentence ψ in a finitary propositional language $\mathcal{L}$.[1] Any new evidence is a sentence μ in $\mathcal{L}$, and the result of revising ψ with μ is also a sentence (denoted as $\psi * \mu$) belonging to $\mathcal{L}$.

To avoid an inconsistency with their postulates for iterated revision, [Darwiche and Pearl, 1994; 1997] suggested to weaken the original KM postulates by regarding belief revision operators as functions on belief states (rather than on belief sets). This resulted in the Postulate (R*1)–(R*6) shown below.

For the sake of the simplicity, we will abuse notation by using interchangeably a belief state Ψ and its belief set $\mathit{Bel}(\Psi)$. For example, Ψ and $\Psi * \mu$ in Postulate (R*1) refer, respectively, to the current belief state and to the posterior belief state; and $\Psi * \mu \models \mu$ is just shorthand for $\mathit{Bel}(\Psi * \mu) \models \mu$. The following are the (modified) KM postulates, which in turn are a reformulation of the AGM postulates:

(R*1) $\Psi * \mu \models \mu$.
(R*2) If $\Psi \wedge \mu$ is satisfiable, then $\Psi * \mu \equiv \Psi \wedge \mu$.
(R*3) If μ is satisfiable, then $\Psi * \mu$ is also satisfiable.
(R*4) If $\Psi_1 = \Psi_2$ and $\mu_1 \equiv \mu_2$, then $\Psi_1 * \mu_1 \equiv \Psi_2 * \mu_2$.
(R*5) $(\Psi * \mu) \wedge \phi \models \Psi * (\mu \wedge \phi)$.
(R*6) If $(\Psi * \mu) \wedge \phi$ is satisfiable, then $\Psi * (\mu \wedge \phi) \models (\Psi * \mu) \wedge \phi$.

Darwiche and Pearl have given a representation theorem for Postulates (R*1)–(R*6) wrt. a revision mechanism based on total pre-orders over possible worlds:

Definition 2.1 Let $\mathcal{W}$ be the set of all worlds (interpretations) of a propositional language $\mathcal{L}$. A function that maps each belief state Ψ to a total pre-order $\leq_\Psi$ on $\mathcal{W}$ is called a *faithful assignment* iff

- If $w_1, w_2 \models \Psi$, then $w_1 =_\Psi w_2$.
- If $w_1 \models \Psi$ and $w_2 \not\models \Psi$, then $w_1 <_\Psi w_2$.
- If $\Psi = \Phi$, then $\leq_\Psi = \leq_\Phi$.

where $w_1 =_\Psi w_2$ is defined as $w_1 \leq_\Psi w_2$ and $w_2 \leq_\Psi w_1$; and $w_1 <_\Psi w_2$ means $w_1 \leq_\Psi w_2$ and $w_2 \not\leq_\Psi w_1$.

The intuitive meaning of $w_1 \leq_\Psi w_2$ is that w_1 is at least as plausible as w_2 in Ψ.

[1] We assume that the corresponding belief set $\mathit{Bel}(\psi)$ contains all logical consequences of ψ.

Theorem 2.2 *[Darwiche and Pearl, 1997] A revision operator $*$ satisfies Postulates (R*1)–(R*6) iff there exists a faithful assignment that maps a belief state Ψ to a total pre-order $\leq_\Psi$ such that*

$$\mathit{Mods}(\Psi * \mu) = \min(\mathit{Mods}(\mu), \leq_\Psi)$$

Here, $\mathit{Mods}(\mu)$ denotes the set of all models of μ, and $\min(\mathit{Mods}(\mu), \leq_\Psi)$ is the set of minimal elements of $\mathit{Mods}(\mu)$ wrt. the pre-order $\leq_\Psi$.

2.2 DP Postulates

In order to allow for successive revisions, in each revision step it must be fully specified how the conditional beliefs are to be modified. Following the principle of economy of information, minimal change should also be applied to the conditional beliefs. Unfortunately, although the KM postulates put much emphasis on the preservation of propositional beliefs, they do not constrain the modification of conditional beliefs. Darwiche and Pearl have shown that the KM postulates alone are too weak to adequately characterize iterated belief revision, because they support unreasonable revision behaviors [Darwiche and Pearl, 1994]. To overcome this deficiency, they proposed these four additional postulates [Darwiche and Pearl, 1997]:

(C1) If $\alpha \models \mu$, then $(\Psi * \mu) * \alpha \equiv \Psi * \alpha$.
(C2) If $\alpha \models \neg\mu$, then $(\Psi * \mu) * \alpha \equiv \Psi * \alpha$.
(C3) If $\Psi * \alpha \models \mu$, then $(\Psi * \mu) * \alpha \models \mu$.
(C4) If $\Psi * \alpha \not\models \neg\mu$, then $(\Psi * \mu) * \alpha \not\models \neg\mu$.

Motivation and interpretation of these postulates can be found in [Darwiche and Pearl, 1994; 1997].

To provide formal justifications, Darwiche and Pearl have presented an extension of the above representation theorem for Postulates (C1)–(C4):

Theorem 2.3 *[Darwiche and Pearl, 1997] Suppose that a revision operator satisfies Postulates (R*1)–(R*6). The operator satisfies Postulates (C1)–(C4) iff the operator and its corresponding faithful assignment satisfy:*

(CR1) If $w_1, w_2 \models \mu$, then $w_1 \leq_\Psi w_2$ iff $w_1 \leq_{\Psi*\mu} w_2$.
(CR2) If $w_1, w_2 \not\models \mu$, then $w_1 \leq_\Psi w_2$ iff $w_1 \leq_{\Psi*\mu} w_2$.
(CR3) If $w_1 \models \mu$ and $w_2 \not\models \mu$, then $w_1 <_\Psi w_2$ implies $w_1 <_{\Psi*\mu} w_2$.
(CR4) If $w_1 \models \mu$ and $w_2 \not\models \mu$, then $w_1 \leq_\Psi w_2$ implies $w_1 \leq_{\Psi*\mu} w_2$.

Furthermore, [Darwiche and Pearl, 1997] showed that their additional postulates are consistent with the (modified) KM postulates. They did so by defining a concrete revision operator which satisfies both (R*1)–(R*6) and (C1)–(C4).

2.3 Why Not Absolute Minimization?

A different approach to studying iterated belief revision is by defining concrete revision operators. For instance, [Boutilier, 1993] proposed a specific revision operator (known as *natural revision*) which satisfies the modified KM postulates and the following one:

(CB) If $\Psi * \mu \models \neg\beta$, then $(\Psi * \mu) * \beta \equiv \Psi * \beta$.

As shown by [Darwiche and Pearl, 1997; Boutilier, 1996], postulate (CB) imposes absolute minimization on the change of conditional beliefs:

Theorem 2.4 *Suppose that a revision operator satisfies Postulates (R*1)–(R6). The operator satisfies Postulate (CB) iff the operator and its corresponding faithful assignment satisfy*

(CBR) If $w_1 \models \neg(\Psi * \mu)$ *and* $w_2 \models \neg(\Psi * \mu)$ *then* $w_1 \leq_\Psi w_2$ *iff* $w_1 \leq_{\Psi * \mu} w_2$

Condition (CBR) says that the agent should keep as much of the ordering information in Ψ as possible while satisfying the KM postulates. At first glance, therefore, it seems that Condition (CBR) complies with the principle of economy of information. However, Postulate (CB) is too radical since any severe revision necessarily cancels all previous evidences [Darwiche and Pearl, 1994; Zhang, 2004]. This suggests that absolute minimization of the change of conditional beliefs is overly strict and not desirable in general. It is easy to see that the DP postulates are a weakening of Postulate (CB), in the sense Postulate (CB) implies all of the DP postulates but not vice versa.

3 The Problem of Implicit Dependence

Although most counter-examples in [Darwiche and Pearl, 1997] against the KM postulates are solved by adding the DP postulates, several open problems remain. For instance, the DP postulates are consistent with (CB), hence they do not block counter-examples against natural revision, like the following [Darwiche and Pearl, 1997]:

Example 1 We encounter a strange new animal and it appears to be a bird, so we believe the animal is a bird. As it comes closer to our hiding place, we see clearly that the animal is red, so we believe that it is a red bird. To remove further doubts about the animal birdhood, we call in a bird expert who takes it for examination and concludes that it is not really a bird but some sort of mammal. The question now is whether we should still believe that the animal is red.

As argued in [Darwiche and Pearl, 1997], we have every reason to keep our belief that the animal is red, since birdhood and color are not correlated. However, natural revision enforces us to give up the belief of the animal's color: According to Postulate (CB), since $\textit{bird}{*}\textit{red} \models \neg(\neg \textit{bird})$ it follows that $(\textit{bird}{*}\textit{red}){*}\neg\textit{bird} \equiv \textit{bird}{*}\neg\textit{bird}$.

In being compatible with (CB), the DP postulates are not strong enough to guarantee that the belief of the animal's color is retained. This can be intuitively explained as follows: After observing the animal's color, we are actually acquiring a new conditional belief, namely, that the animal is red even if it were not a bird i.e., $\neg\textit{bird} \gg \textit{red}$. However, the DP postulates do not enforce the acquisition of conditional beliefs. In the sequel, we first give a formal analysis of this weakness of the DP postulates, and then we present an additional postulate by which this problem is overcome.

It is well known (see, e.g., [Gärdenfors and Makinson, 1988]) that if a belief state Ψ suffices to uniquely determine a revision strategy that satisfies the AGM (or the KM) postulates, then the belief state determines a unique, total pre-order $\leq_\Psi$ (known as *epistemic entrenchment*) over $\mathcal{L}$ which satisfies certain conditions. Given an epistemic entrenchment, the corresponding belief revision operator is defined by the following condition: For any β,

(C*) $\Psi * \mu \models \beta$ if either $\models \neg\mu$ or $\neg\mu <_\Psi \neg\mu \vee \beta$

Other forms of total pre-orderings on $\mathcal{L}$ have been proposed, e.g., [Rott, 1991; Williams, 1992]. All of these orderings require extra-logical information, that is, they cannot be determined by pure logical relations among the sentences. Although we focus on epistemic entrenchment, the analysis that follows does not depend on this specific ordering; it can be easily adapted to the approaches just mentioned.

To begin with, we define the notion of dependence between sentences wrt. a belief state as follows [Fariñas del Cerro and Herzig, 1996]:

Definition 3.1 A sentence β *depends* on another sentence μ in belief state Ψ precisely when $\Psi \models \beta$ and $\Psi{*}\neg\mu \not\models \beta$. Two sentences μ, β are called *dependent* in Ψ if either μ depends on β or β depends on μ in Ψ.

Consider, now, a (non-tautological) new evidence μ. Whenever $\Psi \models \beta$, condition (C*) implies that if $\mu \geq_\Psi \mu \vee \beta$, then β is (implicitly) dependent on μ in Ψ. This kind of dependency could be problematic. In particular, it is possible that two initially independent sentences become, undesirably, dependent after a revision step. In Example 1, say, *red* becomes dependent on *bird* after revising by *red* if natural revision is used.

Epistemic entrenchment reveals that the problem of natural revision is that it assigns the lowest degree of belief to a new evidence without asserting conditional beliefs for independence. Thus the new evidence depends on all other beliefs which survive the revision process. This explains why severe revision always cancels all previous evidences.

4 A Postulate of Independence

The analysis in the previous section shows that in order to overcome the problem of implicit dependence, the revision operator must explicitly assert some conditional beliefs. It is easy to see that the DP postulates only require the *preservation* of conditional beliefs when a belief state Ψ is revised with μ: Postulates (C1) and (C2) neither require to add nor to remove certain conditional beliefs (namely, those conditioned on β) in case $\beta \models \mu$ or $\beta \models \neg\mu$; Postulate (C3) requires to retain the conditional belief $\beta \gg \mu$; finally, Postulate (C4) requires not to obtain the new conditional belief $\beta \gg \neg\mu$. Since none of the DP postulates requires to make independence assumptions, a new postulates is necessary to avoid undesired dependencies.

The revision process may introduce undesirable dependencies in both directions. That is, it could be that the new evidence becomes dependent on existing beliefs, or the other way around. Prior to stating the new postulate, we show that the DP postulates impose some constraints on the retention of the independence information in one direction. In the presence of the KM postulates, Postulate (C2) implies the following:

(C2′) If $\Psi * \neg\mu \models \beta$, then $(\Psi * \mu){*}\neg\mu \models \beta$

This essentially means that if β is not dependent on the new evidence μ in Ψ, then it also does not depend on μ in $\Psi * \mu$.

In order to ensure the explicit assertion of independence information in the other direction, we propose the following postulate of *Independence* to complement the DP postulates:

(Ind) If $\Psi*\beta \not\models \neg\mu$ then $(\Psi*\mu)*\beta \models \mu$

Basically, Postulate (Ind) says that if the belief in $\beta \supset \neg\mu$ is not strongly held in Ψ, then μ does not depend on $\neg\beta$ in $\Psi * \mu$.

Postulate (Ind) is sufficient to overcome the problem of implicit dependence, as can be shown by reconsidering Example 1. According to (Ind), $(\mathit{bird}*\mathit{red})*\neg\mathit{bird} \models \mathit{red}$, given that $\mathit{bird}*\neg\mathit{bird} \not\models \neg\mathit{red}$. This shows that the new postulate blocks unreasonable behaviors which are admitted by the DP postulates. In Section 5, we will also argue that Postulate (Ind) is not overly strict.

In order to formally justify our new postulate, we will first provide a representation theorem along the line of Theorem 2.3. Thereafter, we will give a concrete belief revision operator which satisfies (Ind).

4.1 A Representation Theorem

For the proof of the representation theorem, we need the following observation, which is a consequence of Theorem 2.2.

Observation 4.1 *Suppose that a revision operator satisfies Postulates (R*1)–(R*6). If* $\not\models \neg\beta$, *then* $\Psi*\beta \models \mu$ *precisely when there exists a world* w *such that* $w \models \mu \wedge \beta$ *and* $w <_\Psi w'$ *for any* $w' \models \neg\mu \wedge \beta$, *where* $\leq_\Psi$ *is the corresponding faithful assignment.*

Theorem 4.2 *Suppose that a revision operator satisfies Postulates (R*1)–(R*6). The operator satisfies Postulate (Ind) iff the operator and its corresponding faithful assignment satisfy:*

(IndR) If $w_1 \models \mu$ *and* $w_2 \models \neg\mu$, *then* $w_1 \leq_\Psi w_2$ *implies* $w_1 <_{\Psi*\mu} w_2$.

Proof: *"$\Leftarrow$": Assume* $\Psi*\beta \not\models \neg\mu$. *From Observation 4.1, it follows that for any world* $w \models \beta \wedge \neg\mu$, *there exists another world* $w' \models \beta \wedge \mu$ *such that* $w' \leq_\Psi w$. *Hence, since* $\leq_\Psi$ *is total, there must be a world* w_1 *such that* $w_1 \models \mu \wedge \beta$ *and* $w_1 \leq_\Psi w_2$ *for any* $w_2 \models \neg\mu \wedge \beta$. *Condition (IndR) then implies that* $w_1 <_{\Psi*\mu} w_2$ *for any* $w_2 \models \neg\mu \wedge \beta$. *Due to Observation 4.1, we have* $(\Psi*\mu)*\beta \models \mu$.

"$\Rightarrow$": Assume $w_1 \models \mu$, $w_2 \models \neg\mu$, *and* $w_1 \leq_\Psi w_2$. *Let* β *be such that* $\mathit{Mods}(\beta) = \{w_1, w_2\}$. *From Theorem 2.2, it follows that* $w_1 \in \mathit{Mods}(\Psi*\beta)$. *Hence* $\Psi*\beta \not\models \neg\mu$. *Postulate (Ind) implies* $(\Psi * \mu) * \beta \models \mu$. *Due to Postulates (R*1) and (R*3),* $\mathit{Mods}((\Psi * \mu) * \beta) = \{w_1\}$. *From Theorem 2.2, it follows that* $w_1 <_{\Psi*\mu} w_2$. □

An immediate consequence of Theorem 2.3 and 4.2 is that Postulate (Ind) implies both (C3) and (C4). Theorem 4.2 also shows that Postulate (Ind) is not overly constrained: Condition (IndR) only requires to decrease the plausibility of the worlds violating the new evidence μ and not to decrease the plausibility of the worlds confirming μ.

4.2 Properties of the Independence Postulate

We suggest to use the KM postulates along with Postulates (C1), (C2), and (Ind) to govern iterated belief revision. To show that these postulates together are consistent, we present a concrete revision operator which satisfies all of them. The operator is based on Spohn's proposal of revising *ordinal conditional functions* [Spohn, 1988], which can be viewed as a qualitative version of Jeffrey's Rule of probabilistic conditioning [Goldszmidt, 1992].

An ordinal conditional function is a function k from a given set of worlds into the class of ordinals. Intuitively, the ordinals represent degrees of plausibility. The lower its ordinal, the more plausible is a world. An ordinal conditional function encodes both a belief set and the conditional beliefs. The belief set $\mathit{Bel}(k)$ is the set of sentences which hold in all worlds of rank 0 (the smallest ordinal):

$$\mathit{Mods}(\mathit{Bel}(k)) = \{w | k(w) = 0\}$$

From now on, we use an ordinal conditional function and its belief set interchangeably; e.g., $\mu \in k$ means $\mu \in \mathit{Bel}(k)$, and $k \wedge \mu$ denotes $\bigwedge \mathit{Bel}(k) \wedge \mu$. A ranking of worlds can be extended to a ranking of sentences as follows:

$$k(\mu) = \begin{cases} (\max_w \ k(w)) + 1 & \text{If } \models \mu \\ \min_{w \models \neg\mu} k(w) & \text{Otherwise} \end{cases} \qquad (1)$$

Put in words, the rank of a formula is the lowest rank of a world in which the formula does not hold. Hence, the higher the rank of a sentence, the firmer the belief in it. In fact, it is not hard to see that an ordinal conditional function k determines an epistemic entrenchment as follows:

$$\alpha \leq_k \beta \text{ iff } k(\alpha) \leq k(\beta)$$

Our belief revision operator allows to assign different plausibility degrees to new evidences; standard KM/DP revision is easily obtained as a special case by using a fixed value in all iterations [Darwiche and Pearl, 1997]. An ordinal conditional function k is revised according to new evidence μ with plausibility degree $m > 0$ as follows:

$$(k^*_{\mu.m})(w) = \begin{cases} k(w) - k(\neg\mu) & \text{If } w \models \mu \\ k(w) + m & \text{If } w \models \neg\mu \end{cases} \qquad (2)$$

Assuming the same degree of plausibility for any new evidence, satisfiability of the KM postulates along with Postulates (C1), (C2), and (Ind) by the revision defined in (2) is an almost direct consequence of Theorem 2.2, 2.3, and 4.2. But we can even prove a stronger result with varying plausibility values m.

Theorem 4.3 *For any* $m > 0$, *the belief revision operator defined by (2) satisfies all KM postulates (R*1)–(R*6), where* $\mathit{Bel}(\Psi)$ *and* $\Psi * \mu$ *are, respectively, identified with* $\bigwedge \mathit{Bel}(k)$ *and* $k^*_{\mu,m}$.

Lemma 4.4 *Let* k *be an arbitrary ordinal conditional function and* μ *a new evidence with plausibility degree* m, *then for any non-tautological sentence* β,

$$k^*_{\mu,m}(\beta) = \begin{cases} k(\beta) + m & \text{If } \models \mu \supset \beta \\ t - k(\neg\mu) & \text{Else if } t = k(\beta) \\ min(t - k(\neg\mu), k(\beta) + m) & \text{Otherwise} \end{cases}$$

where $t = k(\mu \supset \beta)$.

Theorem 4.5 *For arbitrary $m_1, m_2 > 0$, the belief revision operator defined by (2) satisfies the following conditions:*[2]

(EC1) If $\alpha \models \mu$, then $(k^*_{\mu,m_1})^*_{\alpha,m_2} \equiv k^*_{\alpha,m_2}$.

(EC2) If $\alpha \models \neg\mu$, then $(k^*_{\mu,m_1})^*_{\alpha,m_2} \equiv k^*_{\alpha,m_2}$.

(EInd) If there exists m such that $k^*_{\neg\beta,m} \not\models \neg\mu$, then $(k^*_{\mu,m_1})^*_{\neg\beta,m_2} \models \mu$

***Proof**: If $\models \neg\alpha$, Condition (EC1) holds trivially. Assume that $\alpha \models \mu$ and $\not\models \neg\alpha$. By (2),*

$$k^*_{\alpha,m_2}(w) = 0 \textit{ iff } w \models \alpha \textit{ and } k(w) = k(\neg\alpha) \tag{3}$$

Likewise,

$$(k^*_{\mu,m_1})^*_{\alpha,m_2}(w) = 0 \textit{ iff } w \models \alpha \textit{ and } k^*_{\mu,m_1}(w) = k^*_{\mu,m_1}(\neg\alpha) \tag{4}$$

Since $\alpha \models \mu$, for any $w \models \alpha$ we have $k^*_{\mu,m_1}(w) = k(w) - k(\neg\mu)$ by (2). Since $\mu \supset \neg\alpha \equiv \neg\alpha$ and $\not\models \neg\alpha$, it follows from Lemma 4.4 that $k^*_{\mu,m_1}(\neg\alpha) = k(\neg\alpha) - k(\neg\mu)$. Hence, (4) is equivalent to

$$(k^*_{\mu,m_1})^*_{\alpha,m_2}(w) = 0 \textit{ iff } w \models \alpha \textit{ and } k(w) = k(\neg\alpha)$$

This and (3) implies $(k^*_{\mu,m_1})^*_{\alpha,m_2} \equiv k^*_{\alpha,m_2}$. Condition (EC2) can be proved analogously.

We prove Condition (EInd) by contradiction. To begin with, from the assumption that $k^*_{\neg\beta,m} \not\models \neg\mu$ it follows that $\not\models \beta$ and $\not\models \mu \supset \beta$. Furthermore, there exists w such that $k^*_{\neg\beta,m}(w) = 0$, $w \models \neg\beta \wedge \mu$, and $k(w) = k(\beta)$. With the help of (1), this implies $k(\beta) = k(\mu \supset \beta)$.

Now assume that $(k^*_{\mu,m_1})^*_{\neg\beta,m_2} \not\models \mu$. It follows that there exists w' such that $(k^*_{\mu,m_1})^*_{\neg\beta,m_2}(w') = 0$, $w' \models \neg\beta \wedge \neg\mu$, and $k^*_{\mu,m_1}(w') = k^*_{\mu,m_1}(\beta)$. Since $k(w) = k(\beta)$ and $w' \models \neg\beta$, we have $k(w') \geq k(w)$. Hence by (2), $k^*_{\mu,m_1}(w') = k(w') + m_1 > k(\beta)$. But from Lemma 4.4 it follows that $k^*_{\mu,m_1}(\beta) \leq k(\beta)$, since $\not\models \beta$ and $\not\models \mu \supset \beta$. This contradicts $k^*_{\mu,m_1}(w') = k^*_{\mu,m_1}(\beta)$. □

Theorem 4.3 and 4.5 show that Postulate (Ind) is consistent with the KM and DP postulates. On the other hand, (Ind) does not follow from these postulates, which can be seen by the fact that (Ind) is incompatible with (CB), the postulate that characterizes natural revision.

It is worth mentioning that revisions based on ordinal conditional functions are particularly suitable for implementations of belief revision. For instance, [Jin and Thielscher, 2004] presented a method (and its implementation) for revision of belief bases which is equivalent to the belief revision defined by (2).

5 Related Work and Conclusions

The first DP approach [Darwiche and Pearl, 1994] seemed excessively strong because one of the postulates, (C2), is inconsistent with the classical AGM framework, as pointed out by [Freund and Lehmann, 1994]. Apart from the solution by [Darwiche and Pearl, 1997], namely, to use a modified version of KM (which we followed in this paper), several other proposals were made to overcome the inconsistency. For instance, Lehmann proposed a different way of weakening the AGM postulates accompanied by quite different additional postulates for iterated revision, in which a belief state consists of a sequence of revision sentences [Lehmann, 1995]. Nayak *et al.* suggested to retain the original AGM framework and to consider a belief revision operator as a unary function (associated with a belief state). The idea is to view belief revision as dynamic, in the sense that the operator evolves after each step [Nayak *et al.*, 1996; 2003]. On the other hand, one point has been generally accepted, namely, that iterated belief revision should not be considered as a purely set-theoretical change of belief sets but as an evolution of belief states, which encapsulate beliefs with some kind of information of the firmness of beliefs [Zhang, 2004].

The problem of the DP postulates to be overly permissive was already studied by Nayak *et al.*, [1996; 2003]. They suggested to strengthen the DP postulates by the following so-called postulate of Conjunction:

(Conj) If $\mu \not\models \neg\beta$, then $(\Psi * \mu) * \beta \equiv \Psi * (\beta \wedge \mu)$.

It is easy to see that (Conj) implies another postulate, which they called postulate of Recalcitrance:

(Rec) If $\not\models \beta \supset \neg\mu$, then $(\Psi * \mu) * \beta \models \mu$.

In the following, we argue that Postulate (Conj), while strengthening the DP postulates, is overly strict. To this end, we show that even Postulate (Rec) is too strong. The latter says that, as long as $\beta \supset \neg\mu$ is not a tautology, it should be canceled after a successive revision by μ followed by β, no matter how strong the initial belief in $\beta \supset \neg\mu$. A simple example shows that this behavior is not reasonable:

Example 2 All her childhood, Alice was taught by her parents that a person who has told a lie is not a good person. So Alice believed, initially, that if Bob has told a lie then he is not a good person. After her first date with Bob, she began to believe that he is a good guy. Then a reliable friend of Alice warns her that Bob is in fact a liar, and Alice chooses to believe her. Now, should Alice still believe that Bob is a good guy?

According to Postulate (Rec), Alice should not challenge Bob's morality and still believe he is good, and hence to disbelieve what her parents taught her. But it is at least as reasonable to give up the belief that Bob is good. This shows that Postulate (Rec) is too strict a criterion for belief revision operators.

With regard to the postulate we have proposed in this paper, it is easy to see that (Ind) is a weakening of Postulate (Rec). This raises the question whether Postulate (Ind) weakens too much. Let us consider an example, taken from [Nayak *et al.*, 1996], which, at first glance, seems to show that this is indeed the case.

Example 3 Our agent believes that Tweety is a singing bird. However, since there is no strong correlation between singing and birdhood, the agent is prepared to retain the belief that

[2] Note that, as before, we abuse notation by simply writing $k^*_{\mu,m}$ instead of $\bigwedge Bel(k^*_{\mu,m})$ etc.

Tweety sings even after accepting the information that Tweety is not a bird, and conversely, if the agent were to be informed that Tweety does not sing, she would still retain the belief that Tweety is a bird. Imagine that the agent first receives the information that Tweety is in fact not a bird, and later learns that Tweety does not sing.

Nayak *et al.* claimed that it is only reasonable to assume that the agent should, in the end, believe that Tweety is a non-singing non-bird. Indeed, with $\Psi \equiv \mathit{singing} \wedge \mathit{bird}$ it follows from Postulate (Rec) that $(\Psi * \neg \mathit{bird}) * \neg \mathit{singing} \models \neg \mathit{bird}$, since $\not\models \neg \mathit{singing} \supset \mathit{bird}$. Postulate (Ind), on the other hand, does not apply in this case. But the behavior which is claimed to be the only reasonable one is not generally justified. Suppose, for example, the agent initially believes firmly that $\neg \mathit{singing} \supset \mathit{bird}$. It is then possible, after revising by $\neg \mathit{bird}$, that the belief in $\neg \mathit{singing} \supset \mathit{bird}$ is stronger than the belief in $\neg \mathit{bird}$. In this case, after further revising by $\neg \mathit{singing}$, the agent believes that Tweety is a bird after all.

In [Zhang, 2004], it has also been argued that Postulate (Rec) is too radical because only those revision operators are admissible which assign the highest belief degree to the new evidence. This can be considered a further justification of our Postulate (Ind). Moreover, it is easy to verify that Postulate (Ind) is satisfied by all revision operators with memory [Konieczny and Pérez, 2000].

We have formally justified our new postulate by a representation theorem and a concrete revision operator which satisfies both (Ind) and KM/DP. Other concrete revision operators have been proposed, e.g., in [Williams, 1995], which also use ordinal conditional functions. The main difference is that these operators do not satisfy the Independence postulate.

References

[Alchourrón *et al.*, 1985] C. E. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *The Journal of Symbolic Logic 50(2),* pages 510–530, 1985.

[Boutilier, 1993] C. Boutilier. Revision sequences and nested conditionals. In *Proceedings of IJCAI*, pages 519–525, 1993.

[Boutilier, 1996] C. Boutilier. Iterated revision and minimal change of conditional beliefs. *Journal of Philosophical Logic*, 25(3), 1996.

[Darwiche and Pearl, 1994] A. Darwiche and J. Pearl. On the logic of iterated belief revision. In R. Fagin, editor, *Proceedings of TARK*, pages 5–23. Kaufmann, San Francisco, CA, 1994.

[Darwiche and Pearl, 1997] A. Darwiche and J. Pearl. On the logic of iterated belief revision. *Artificial Intelligence 89*, 1–29, 1997.

[Fariñas del Cerro and Herzig, 1996] L. Fariñas del Cerro and A. Herzig. Belief change and dependence. In Yoav Shoham, editor, *Proceedings of TARK*, pages 147–162. Morgan Kaufmann, 1996.

[Freund and Lehmann, 1994] M. Freund and D. Lehmann. Belief revision and rational inference. Technical Report TR-94-16, Institute of Computer Science, The Hebrew University of Jerusalem, Jerusalem, 91904, Israel, 1994.

[Gärdenfors and Makinson, 1988] P. Gärdenfors and D. Makinson. Revisions of knowledge systems using epistemic entrenchment. In *Proceedings of TARK*, pages 83–95, Asilomar, CA, 1988.

[Gärdenfors, 1988] P. Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press, 1988.

[Goldszmidt, 1992] M. Goldszmidt. *Qualitative probabilities: a normative framework for commonsense reasoning*. PhD thesis, University of California at Los Angeles, 1992.

[Jin and Thielscher, 2004] Y. Jin and M. Thielscher. Representing beliefs in the fluent calculus. In R. L. de Mántras and L. Saitta, editors, *Proceedings of ECAI*, pages 823–827, Valencia, Spain, August 2004. IOS Press.

[Katsuno and Mendelzon, 1991] H. Katsuno and A. O. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52(3):263–294, 1991.

[Konieczny and Pérez, 2000] S. Konieczny and R/ Pérez. A framework for iterated revision. *Journal of Applied Non-Classical Logics*, 10(3-4), 2000.

[Lehmann, 1995] D. J. Lehmann. Belief revision, revised. In C. S. Mellish, editor, *Proceedings of IJCAI*, pages 1534–1540, 1995.

[Nayak *et al.*, 1996] A. C. Nayak, N. Y. Foo, M. Pagnucco, and A. Sattar. Changing conditional beliefs unconditionally. In Yoav Shoham, editor, *Proceedings of TARK*, pages 119–135. Morgan Kaufmann, San Francisco, 1996.

[Nayak *et al.*, 2003] A. C. Nayak, M. Pagnucco, and P. Peppas. Dynamic belief revision operators. *Artificial Intelligence*, 146(2):193–228, 2003.

[Rott, 1991] H. Rott. A nonmonotonic conditional logic for belief revision. In A. Fuhrmann and M. Morreau, editors, *The logic of theory change*, pages 135–183. Springer, Berlin, 1991.

[Spohn, 1988] W. Spohn. Ordinal conditional functions: A dynamic theory of epistemic state. In William L. Harper and Brian Skyrms, editors, *Causation in Decision: Belief, Change and Statistics: Proceedings of the Irvine Conference on Probability and Causation*, volume II, pages 105–134, Dordrecht, 1988. Kluwer Academic Publisher.

[Williams, 1992] M-A. Williams. Two operators for theory bases. In *Proceedings of the Australian Joint Artificial Intelligence Conference*, pages 259–265. World Scientific, 1992.

[Williams, 1995] M.-A. Williams. Iterated theory base change: A computational model. In C. S. Mellish, editor, *Proceedings of IJCAI*, pages 1541–1547, 1995.

[Zhang, 2004] D. Zhang. Properties of iterated multiple belief revision. In *7th International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR)*, pages 314–325. Springer, 2004.