Belief revision, revised *

Daniel Lehmann
Institute of Computer Science
Hebrew University
Jerusalem 91904 Israel
lehmann@cs.huji.ac.il

Abstract

This paper studies properties of iterated revisions. First, a triviality result shows that, in the AGM original framework, the only revision operation that satisfies two reasonable properties is the trivial revision. Then, an alternative to the AGM framework for studying belief revision and rationality postulates is proposed. Iterated revisions are the objects of this formalism and the rationality postulates deal with properties of iterated revisions. A set of rationality postulates is presented, closely related to the AGM postulates. A representation result shows that those postulates imply serious limitations to the way revisions can cope with the Principle of Minimal Change. Those postulates are not suitable for belief update, but their consideration raises doubts about the adequacy of previous treatments of belief update.

1 Introduction

Intelligent agents must gather information about the world, elaborate theories about it and revise those theories in view of new information that, sometimes, contradicts the beliefs previously held. Belief revision is therefore a central topic in Knowledge Representation. It has been studied in different forms: numeric or symbolic, procedural or declarative, logical or probabilistic. One of the most successful framework in which belief revision has been studied has been proposed by Alchourron, Gardenfors and Makinson and is known as the AGM framework. It deals with operations of revision that revise a theory (the set of previous beliefs) by a formula (the new information). It proposes a set of rationality postulates that any reasonable revision should satisfy. A large number of researchers in AI have been attracted by and have developed this approach further, both in the abstract and by devising revision procedures that satisfy the AGM rationality postulates. Even though iterated revisions are a central concern in AI applications, those postulates do not mention iterated revisions explicitly. A closer look shows that the AGM postulates do not say anything on the way a revision may depend on its first argument, the theory being revised. If one considers iterated revisions, shouldn't one put some requirement on this dependence? This paper will show that some such requirements should be considered. It is time to take stock of what has been accomplished and assess the general approach. Noticeable progress has been made in three directions:

- in distinguishing revisions of a theory about an unchanging world from updates of a theory about an evolving situation under an assumption of inertia [Katsuno and Mendelzon, 1992],
- in relating the AGM rationality postulates to properties of nonmonotonic consequence relations [Makinson and Gardenfors, 1989; Gardenfors, 1990; Gardenfors and Makinson, 1994],
- and in the explicit consideration of iterated revisions, their properties and postulates for them.

The third point needs to be expanded upon. Many authors ([Boutilier and Goldszmidt, 1993; Boutilier, 1993; Darwiche and Pearl, 1994; Williams, 1994; Nayak *et al.*, 1995]) have recently expressed interest in iterated revisions. One may definitely consider iterated revisions in the AGM framework, but one may notice that AGM consistently avoided mentioning iterations of their * operator in their rationality postulates. Natural properties of iterated revisions, such as postulate **C1** below, have therefore been largely ignored (until [Darwiche and Pearl, 1994]).

This work proposes a slightly modified, more convenient, framework in which to study iterated belief revisions.

In section 2, the AGM postulates will be recalled. In section 3, two additional postulates concerning iterated revisions will be described, and a triviality result proved. Then, in section 4, a new framework and notation will be described. A set of rationality postulates is proposed and discussed in section 5. Formal properties of the postulates are proved in section 6. A family of models, that of widening ranked models, is described in section 7. The revisions they define are exactly those that satisfy the postulates proposed in section 5. Methodological conse

* This work was partially supported by the Jean and Helene Alfassa fund for research in Artificial Intelligence and by grant 136/94-1 of the Israel Science Foundation on "New Perspectives on Nonmonotonic Reasoning".

quences of this representation result are discussed. Section 8 concludes by stressing open problems.

2 The AGM framework

The AGM framework studies *revision* operations, denoted $*$, that operate on two arguments: a theory (any set of formulas closed under logical deduction) K on the left, and a formula a on the right. Thus $K*a$ is the result of revising theory K by formula a, using revision method $*$. The original AGM rationality postulates are the following. We use the customary notation $Cn(X)$ for the set of all logical consequences of a set X of formulas. The logical connectives always have precedence over the $*$ operation.

K*1 $K*a$ is a theory

K*2 $a \in K*a$

K*3 $K*a \subseteq Cn(K, a)$

K*4 If $\neg a \notin K$, then $Cn(K, a) \subseteq K*a$

K*5 If $K*a$ is inconsistent, then a is a logical contradiction

K*6 If $\models a \leftrightarrow b$, then $K*a = K*b$

K*7 $K*a \wedge b \subseteq Cn(K*a, b)$

K*8 If $\neg b \notin K*a$, then $Cn(K*a, b) \subseteq K*a \wedge b$

3 Additional postulates

The following additional postulate was proposed in [Darwiche and Pearl, 1994].

C1 $K*a*a \wedge b = K*a \wedge b$

The justification for **C1** is that, if an agent learns first some partial information (a) and then the full information ($a \wedge b$), the partial information should not make a difference. Detailed arguments in favor of **C1** may be found in [Darwiche and Pearl, 1994] and below in this paper, after the presentation of **I5** in section 5. That **C1** is consistent with the AGM postulates is seen from the consideration of the trivial revision defined by:

$$K*a = \begin{cases} Cn(a) & \text{if } \neg a \in K \\ Cn(K, a) & \text{otherwise} \end{cases}$$

The trivial revision satisfies **C1** and the AGM postulates. The following shows that **C1** is a powerful postulate.

Theorem 1 *Any revision operation $*$ that satisfies* **K*1**–**K*6** *and* **C1** *satisfies* **K*7**, **K*8** *and the postulates* **C3** *and* **C4** *proposed in [Darwiche and Pearl, 1994].*

In other terms, **C1** is an elegant, slightly more powerful version of the AGM postulates **K*7** and **K*8**. The proof is similar to that of the corresponding lemmas of section 6.

The postulate **C2** proposed in [Darwiche and Pearl, 1994] has been shown inconsistent with the AGM postulates in [Freund and Lehmann, 1994]. Even the weaker **C2'** proposed in [Nayak *et al.*, 1995] is inconsistent. An even weaker postulate is:

D1 $K*\neg a*a \subseteq K*a$

The justification for **D1** is that, if an agent successively learns two contradictory pieces of information, it should not, at the end, hold beliefs it would not hold if it had not learned the false, intermediate information. Such beliefs could only be the result of information known to be incorrect. The next result is a triviality result.

Theorem 2 *The only revision operation that satisfies* **K*1**–**K*6**, **C1** *and* **D1** *is the trivial revision defined above.*

A lemma is needed.

Lemma 1 *Let $*$ be a revision satisfying* **K*1**–**K*6**, **C1** *and* **D1**. *If $b \models \neg a$, then $K*a*b \subseteq Cn(K, b)$.*

Proof: By **C1**, $K*a \wedge \neg b*b = K*a \wedge \neg b*\neg a \vee b*b$. But the left hand side is $K*a*b$, and $K*a*b \subseteq Cn(K*a \wedge \neg b*\neg a \vee b, b)$. One concludes by **D1**: $K*a*b \subseteq Cn(K, b)$. ∎

The proof of Theorem 2 will be presented now.

Proof: Suppose $\neg a \in K$, the only interesting case. Let k be the conjunction of all formulas of K (a finite language is assumed) and $K_\emptyset$ be the set of all tautologies. Then $K = K_\emptyset*k$ and $a \models \neg k$. By Lemma 1, then $K*a = K_\emptyset*k*a \subseteq Cn(K_\emptyset, a) = Cn(a)$. ∎

4 The revised framework

Central to the present analysis is the concept of a belief state resulting from a finite sequence of revisions. In this work, the underlying language is supposed to be a propositional calculus on a *finite* set of atomic propositions. The individual revisions considered are represented by *consistent* formulas: the formula describes the new information received that triggers the revision. The letters a, b and c will be used to stand for arbitrary consistent formulas. Arbitrary finite sequences of individual revisions, i.e., finite sequences of consistent formulas, will be represented by: ρ, σ and τ. The empty sequence is denoted Λ. Concatenation of sequences is denoted $\cdot$ and formulas will be identified with sequences of length one. For example, the sequence $\sigma \cdot a \cdot \tau$ consists of first the formulas of the sequence σ, then the formula a and finally the formulas of the sequence τ. The belief set resulting from a sequence σ of individual revisions, will be denoted by $[\sigma]$. The square brackets ([ ]) represent a revision procedure. For example, $[\Lambda]$ is the belief set resulting from revising the initial belief set by an empty sequence of revisions: it is the initial belief set. A theory is a set of formulas that is closed under logical consequence. Material implication is denoted $\rightarrow$.

In order to understand and evaluate the set of new rationality postulates presented in section 5, it helps to consider the correspondence between the new notation introduced above and the traditional AGM notation. As explained above $[\sigma \cdot a]$ represents the belief set that results from revising the initial belief set by, first, the formulas of σ and then the formula a. Similarly $[\sigma]$ represents the belief set that is the result of revising the initial belief set with the elements of σ. The belief set $[\sigma \cdot a]$ therefore represents the result of revising the belief set $[\sigma]$ by the formula a. It may seem that we may identify

$[\sigma, a]$ with what AGM would have denoted $[\sigma] * a$. Nevertheless, an important difference exists. The notation used by AGM implicitly implies that, if $[\sigma] = [\tau]$, then $[\sigma] * a = [\tau] * a$, whereas the set of postulates of section 5 does not imply

Non – postulate If $[\sigma] = [\tau]$, then $[\sigma, a] = [\tau, a]$

The framework of this paper allows the agent to base its revision, i.e., $[\sigma, a]$, not only on the belief set $[\sigma]$ but also on the way the agent got to this belief set, i.e., on the sequence σ itself. It must be pointed at that, though the framework presented here does not include the non-postulate, this postulate is expressible and therefore may be added to the set of basic postulates to be presented below, if one wishes to do so. But theorem 5 will show that the only revision that satisfies the rationality postulates of section 5 and the non-postulate is the trivial revision.

5 Rationality Postulates

5.1 The postulates

The following postulates are proposed to characterize reasonable revision methods. They have to be understood for any sequences σ, τ and for any formulas a, b.

I1 $[\sigma]$ is a consistent theory

I2 $a \in [\sigma, a]$

I3 If $b \in [\sigma, a]$, then $a \rightarrow b \in [\sigma]$

I4 If $a \in [\sigma]$, then $[\sigma, \tau] = [\sigma, a, \tau]$

I5 If $b \models a$, then $[\sigma, a, b, \tau] = [\sigma, b, \tau]$

I6 If $\neg b \notin [\sigma, a]$, then $[\sigma, a, b, \tau] = [\sigma, a, a \wedge b, \tau]$

I7 $[\sigma, \neg b, b] \subseteq Cn([\sigma], b)$

Note that the formula $a \wedge b$ appearing in **I6** is indeed a consistent formula if $\neg b \notin [\sigma, a]$ by **I1** and **I2**.

5.2 Discussion of the postulates

Postulate **I1** requires belief sets to be consistent theories. It corresponds exactly to the AGM postulates **K*1** and **K*5**. The only reason **I1** looks stronger than **K*5** is that the only revisions considered here are revisions by consistent formulas. **I1** therefore does not represent any strengthening of the AGM postulates.

Postulate **I2** requires the last element of a sequence of revisions to be a member of the belief set obtained. Since a is a consistent formula, this requirement is compatible with **I1**. It corresponds exactly to **K*2**.

Postulate **I3** expresses the fact that the agent should be aware of the results of possible revisions: if it would accept b after a revision by a, it should already now accept that if a is true then b is true. It is similar to the property (S) of [Kraus *et al.* 1990]. It corresponds exactly to AGM's **K*3**.

The next three postulates have the same structure: *they assert that two different* sequences of revisions are equivalent, i.e., they produce the same belief set. Postulate **I4** deals with the case a revision is superfluous and can be eliminated by virtue of previous revisions: a revision by a formula that is a member of the current belief set is useless. As an example, consider an agent in a shuttered room that believes the sun is shining. Suppose it opens a shutter and sees that, indeed, the weather outside conforms to its beliefs. Postulate **I4** says that this agent should not modify its belief set and, moreover, also says that this confirmation of the sunny weather outside *should in no way influence further revisions*. This seems reasonable as far as theory revision is concerned; it seems unreasonable for theory update (in a changing world). Suppose indeed that an hour after the scenario described above, our agent opens a shutter and realizes it is raining. In the case it did not open the shutter earlier, it may well want to assume it has been raining for some time and believe a flood warning has ben put out already. In the case it did open the shutter, on the contrary, it will believe the flood warning has not been put out yet. Note that, in the study of theory update found in [Katsuno and Mendelzon 1992], the equivalence $\varphi \diamond \mu \equiv \varphi$ when $\varphi \models \mu$ is their postulate **(U2)**, and is therefore part of the system. It follows that, for them, if $\varphi \models \mu$, for any ν, $\varphi \diamond \mu \diamond \nu \equiv \varphi \diamond \nu$. The fact that the similar **I4** is unacceptable for theory update raises serious doubts as to whether the Katsuno-Mendelzon approach is capable of dealing properly with iterated updates. Notice that **I4** is unreasonable if the theory being revised is inconsistent, but this case is excluded in the present framework: revising the inconsistent theory by a consistent formula must change the theory, by **I2**. Postulate **I4** expresses a property akin to the Cumulativity of [Kraus *et al.* 1990]: learning that a plausible consequence is indeed true does not change the set of plausible consequences. Note also that the property: If $\varphi \in K$, then $K * \varphi = K$ follows from the AGM postulates when the theory K is consistent. Postulate **I4** therefore does not imply any *strengthening of the AGM postulates, or of the Katsuno-Mendelzon ones for theory updates*, as was pointed out just above.

Postulate **I5** is perhaps the most remarkable of the list. It concerns the case a revision is made superfluous by virtue of the following revision: if one revises first by a and then by $a \wedge b$, the second revision confirms all the information contained in the first one and, therefore, the first revision is superfluous. Suppose, for example, that an agent listens to some reliable source of information saying: "c is true ... and d is also true". It should not make a difference for the agent to either wait until the end of the sentence and revise by $c \wedge d$, or revise first by c on hearing the first part of the sentence and then revising by $c \wedge d$ on realizing it did not hear all of the sentence. Note that revising first by c and then by d is not equivalent to the above, since the agent learned that c and d are true together, whereas learning d alone, even after learning c, could convince the agent to abandon c. Postulate **I5** is an adaptation to the new framework of **C1**. In this new framework, it is not enough to say that, in the situations considered, the first revision (by a) cannot influence the belief set that results from the second one (by b): the postulate must also say explicitly that the first revision cannot influence the result of any further revision (by τ). Postulate **I5** does not correspond to any of the AGM postulates and indeed **C1** does not follow from the AGM postulates. Postulate **I5** therefore represents a definite strengthening of the AGM postulates but it seems very difficult to find reasons to reject it. The question whether **I5** should be accepted for updates is more delicate. My present feeling is that it should: learning first a and then $a \wedge b$ should not make me think that there have been two changes in the world but probably that, when I learned a, $a \wedge b$ was already true.

Postulate **I6** deals with the case in which a formula a, used to revise a belief set, may be repeated in the next revision, i.e., when revising first by a and then by b is equivalent to revising by a and then by $a \wedge b$. Note immediately that, in the presence of **I5**, **I6** is equivalent to

I6′ If $\neg b \notin [\sigma, a]$, then $[\sigma, a, b, \tau] = [\sigma, a \wedge b, \tau]$

I6′ says that, if a certain condition is met, two successive revisions may be collapsed into one revision by the conjunction of the two formulas involved. The condition to be met is that the second revision is a *mild* one, i.e., $\neg b \notin [\sigma, a]$. It stands to reason that there may be a difference between revising by $a \wedge b$ one one hand and revising σ first by a and then by b only if the second revision is a *severe* revision, i.e., $\neg b \in [\sigma, a]$. If it is not a severe revision, we expect the result to contain the formula a, and, therefore, we expect **I6** to be satisfied. Postulate **I6′** corresponds to AGM's postulate **K∗8** and a special case of **K∗7**. Postulate **I6** therefore does not represent any strengthening of the AGM postulates.

Postulate **I7** is an instance of the Principle of Minimal Change. If, after a sequence σ of revisions, an agent learns $\neg b$ and then b, it certainly believes b but any additional belief it may have can only come from the beliefs it held after the sequence σ: none of the beliefs triggered by learning $\neg b$ should survive learning b. Postulate **I7** is not a consequence of **K∗1**–**K∗8**, and even not a consequence of the additional **K∗9** of [Freund and Lehmann 1994]. Translated back in the traditional AGM framework, one gets **D1** above, which is a weakening of **C2**. It is the strongest possible weakening of **C2** that is consistent, and should therefore be considered as the correct expression of the intuitions behind it.

If one considers the postulates above, not from the point of view of theory revision, i.e., changing beliefs on a static situation, but from the point of view of theory update [Katsuno and Mendelzon 1992], i.e., changing beliefs on a changing situation, one would probably accept all the postulates above, under the usual assumptions (priority to new information), except for **I4** and **I7**. The confirmation of the truth of a proposition believed to be true should probably influence further updates by giving a posterior time-stamp to the information. Learning that a proposition b has changed from false to true may provide the agent with new information expressing the consequences of a change.

The set of rationality postulates presented above is very similar to the set proposed by AGM. On one point it is slightly weaker: the non-postulate, and two other points it is slightly stronger: **I5** and **I7**. Since postulates **I5** and **I7** seem secure, i.e., difficult to reject, the postulates **I1**–**I7** may probably be considered as a reasonable formalization of the intuitions of AGM. **I7** is probably less secure than **I5**, but the system **I1**-**I6** is amenable to a study that is very close to the one to presented now.

6 Properties of the postulates

Some consequences of postulates **I1**–**I6** (the last postulate is not used in this section) will be described now. Many of the proofs have been omitted.

In the rest of this section, $\sigma \mapsto [\sigma]$ is supposed to be a revision system satisfying **I1**–**I6**. One interesting aspect of the postulates above is that no postulate similar to **K∗6**, expressing invariance under replacement of a formula by a logically equivalent formula, is needed.

Lemma 2 *If a is logically equivalent to a', then* $[\sigma, a, \tau] = [\sigma, a', \tau]$.

Proof By **I2** and **I1**, since $a \models a'$, $a' \in [\sigma, a]$. By **I4**, then, we have $[\sigma, a, \tau] = [\sigma, a, a', \tau]$. But since $a' \models a$, by **I5**, $[\sigma, a, a', \tau] = [\sigma, a', \tau]$. ∎

The next results, lemmas 3 and 4, show that our system enjoys all the properties of the AGM system.

Lemma 3 *If $\neg a \notin [\sigma]$, then $[\sigma, a] = Cn([\sigma], a)$.*

Proof By **I3**, $[\sigma, a] \subseteq Cn([\sigma], a)$. By **I1** and **I2**, $[\sigma, a]$ is a theory that contains a. To show $Cn([\sigma], a) \subseteq [\sigma, a]$, it is therefore enough to show that $[\sigma] \subseteq [\sigma, a]$. Suppose $b \in [\sigma]$. By **I4**, we have both $[\sigma, b] = [\sigma]$ and $[\sigma, b, a] = [\sigma, a]$. We conclude that $\neg a \notin [\sigma, b]$ and that, by **I6′**, we have $[\sigma, b, a] = [\sigma, b \wedge a]$. By **I1** and **I2**, then, $b \in [\sigma, b, a] = [\sigma, a]$. ∎

The next result expresses the AGM postulates **K∗7** and **K∗8**.

Lemma 4 $[\sigma, a \wedge b] \subseteq Cn([\sigma, a], b)$ *and, if $\neg b \notin [\sigma, a]$, then one also has* $Cn([\sigma, a], b) \subseteq [\sigma, a \wedge b]$.

Proof If $\neg b \in [\sigma, a]$, $Cn([\sigma, a], b)$ is the full language and the the first claim is proved. It is left to us to prove that if $\neg b \notin [\sigma, a]$ then $Cn([\sigma, a], b) = [\sigma, a \wedge b]$. But, by lemma 3, $Cn([\sigma, a], b) = [\sigma, a, b]$ and by **I6′**, $[\sigma, a, b] = [\sigma, a \wedge b]$. ∎

We have shown that, for a fixed σ, the theories $[\sigma, a]$ obey all of the AGM postulates. The next batch of results, lemmas 5, 6 and 7, concern the influence of a mild revision before a severe revision.

Lemma 5 *If $\neg a \notin [\sigma]$, but $\neg b \in [\sigma]$, then* $[\sigma, a, b, \rho] = [\sigma, b, \rho]$.

Proof By **I5**, $[\sigma, a, b, \rho] = [\sigma, a \vee b, a, b, \rho]$. Since $\neg a \notin [\sigma]$, we know that $\neg(a \vee b) \notin [\sigma]$, and, by lemma 3, $[\sigma, a \vee b] = Cn([\sigma], a \vee b)$. Since $\neg b \in [\sigma]$, we conclude that $a \in [\sigma, a \vee b]$. Postulate **I5** then implies that we have $[\sigma, a \vee b, a, b, \rho] = [\sigma, a \vee b, b, \rho]$. By **I5** again, $[\sigma, a \vee b, b, \rho] = [\sigma, b, \rho]$ and we are done. ∎

The next result concerns disjunction; it has an interest of its own. As one would expect, if the same conclusion is arrived at by two different but parallel sequences of revisions, σ, b, τ and σ, c, τ, where c has replaced b, then it could have been arrived at by the sequence $\sigma, b \vee c, \tau$.

Lemma 6 *If $a \in [\sigma\ b\ \tau]$ and $a \in [\sigma\ c\ \tau]$, then we have $a \in [\sigma\ b \vee c\ \tau]$.*

We may now state our main result concerning the effect of an intermediate mild revision. This result is very important because it is, in a sense, paradoxical. One of the main aims of Belief Revision is to provide revision systems that satisfy the Principle of Minimal Change: as much as possible of the *old* theory should be kept in the revised theory. In case of a mild revision this principle is fully enforced by lemma 3. But in the case of a severe revision, the postulates (ours or AGM's) do not seem to enforce the principle in any way. The next lemma shows that, essentially, the Principle of Minimal Change for severe revisions is *inconsistent* with the very natural postulates **I1–I6**. The effect of a mild revision essentially fades away at the first severe revision: the identity of the formula by which one performed a mild revision is forgotten after the first severe revision. Therefore no part of the *old* theory that was brought about by the mild revision can be kept in the *new* theory. The best one could hope for is to keep those parts of the *old* theory that were brought about by the previous severe revision. We shall not try, here, to discuss the implications of this result, if there are, for the epistemology of science.

Lemma 7 *If one has $\neg a \notin [\sigma]$, $\neg a' \notin [\sigma]$, $\neg b \in [\sigma\ a]$ and $\neg b \in [\sigma\ a']$, then $[\sigma\ a\ b\ \tau] = [\sigma\ a'\ b\ \tau]$.*

The previous lemmas present basic consequences of the postulates **I1–I6**. They are also needed for proving the representation theorem of section 7. The next lemmas show that the postulates **I1–I6** imply properties that are similar to the **C3** and **C4** of [Darwiche and Pearl, 1994].

Lemma 8 describes a strengthening of both **I5** and property **C3**. It resembles **I4**, but the order between a and b is different. The proof of lemma 8 provides a proof that **C3** is a consequence of **C1** and the AGM postulates.

Lemma 8 *If $b \in [\sigma\ a]$, then $[\sigma\ a\ \rho] = [\sigma\ b\ a\ \rho]$.*

Proof By **I5**, $[\sigma\ a\ \rho] = [\sigma\ a \vee b\ a\ \rho]$. Since $b \in [\sigma\ a]$, $b \in [\sigma\ a \vee b]$ and by **I4** $[\sigma\ a \vee b\ a\ \rho] = [\sigma\ a \vee b\ b\ a\ \rho]$ and we conclude by **I5**. ∎

Next lemma is similar to Darwiche and Pearl's **C4**. Its proof provides a proof that **C4** follows from **C1** and the AGM postulates.

Lemma 9 *If $\neg b \notin [\sigma\ a]$, then $\neg b \notin [\sigma\ b\ a]$.*

7 Iterated revisions in widening ranked models

A class of revision procedures, defined by a family of ranked models, will be presented now. Any procedure of this class satisfies the postulates presented in section 5. A word of caution and explanation is needed here to clarify the role those models. The models to be presented are an extremely useful technical tool to study rationality postulates; they are not intended to describe the *ontology* of theory revision. The reader may justifiably feel that the models presented do not provide for a proper ontology of theory revision because the revisions they define do not satisfy, or do not comply enough with, the Principle of Minimal Change. This point will be discussed at the close of this section and taken again in more depth in section 8.

The general framework is that of a ranked model, defined in [Lehmann and Magidor, 1992]. The agent uses a plausibility ranking over possible state of affairs. Each state of affairs is labelled with a propositional model. After any sequence of revisions the agent sees itself at a certain implausibility rank n and considers plausible a *non-empty set of states of affairs of rank n*. When revising then by a formula a, the agent examines whether there is a state of affairs among the ones it considers plausible that satisfies a. If there is one, it continues to see itself at the same implausibility rank n, but downsizes its set of plausible states to those that satisfy a. If, on the contrary, there is no state that satisfies a among the plausible states, it increases its implausibility rank to the smallest rank m above n at which there is a state of affairs that satisfies a. It considers plausible any state of rank m that satisfies a.

Two additional restrictions are put on the ranked model. First, the set of possible states of affairs (i.e. propositional models) at an implausible rank is a superset of the set at a more plausible rank: the set of propositional models widens with higher implausibility ranks. This enforces the condition that less assumptions, less *default information* is available at less plausible ranks: a sequence of contradicting revisions (by b and $\neg b$ successively) essentially cannot give the agent new information. Secondly, any propositional model must appear at some implausibility rank, and therefore at any rank of sufficient implausibility. The purpose of this last requirement is to ensure that the agent always, i.e. after any sequence of revisions, ends up with a non empty set of plausible states of affairs.

Even though the underlying language considered is finite, we cannot restrict our attention to models in which the ranks are natural numbers: we need to allow ordinals as ranks. Consider, for example, a revision system in which $b \in [a\ \neg a\ a\ \ldots\ a]$ and $b \in [a\ \neg a\ a\ \ldots\ \neg a]$ for any sequence of revisions by a and its negation alternatively. The rank corresponding to $[\neg b]$ must be infinite. The formal definition of a widening ranked model will be given now. As explained above, the ranks are ordinals. Note that, since the ranks are ordinals (a well ordered family), the smoothness condition of [Lehmann and Magidor, 1992] becomes superfluous. Some initial segment of the class of ordinals will be denoted by $\mathcal{N}$ and the set of all propositional models by $\mathcal{W}$. It is easy to see that the initial segment $\mathcal{N}$ does not need to be very long: $\omega + \omega$ is always sufficient, if the language is finite.

Definition 1 *An widening ranked model consists of is a function $M : \mathcal{N} \mapsto 2^{\mathcal{W}} - \emptyset$ such that*

1. *for any $n, m \in \mathcal{N}$, if $n \leq m$, then $M(n) \subseteq M(m)$, and*
2. *for any $w \in \mathcal{W}$, there is some $n \in \mathcal{N}$ such that $w \in M(n)$.*

The value of $M(n)$ is the set of all models that represent some state of affairs of implausibility n. Notice that, on a finite language, in any widening ranked model, there is a rank n such that for every $m \geq n$, $M(m) = W$. Any widening ranked model M defines a system of iterated revisions in the following way. Any sequence of revisions σ defines a rank $r(\sigma)$ and some set of propositional models $p(\sigma) \subseteq M(r(\sigma))$.

Definition 2 *The definition is by induction on the length of* σ.

- $r(\Lambda) = 0$ *and* $p(\Lambda) = M(0)$.
- *If there is in* $p(\sigma)$ *some element that satisfies* α, *one has* $r(\sigma \cdot \alpha) = r(\sigma)$ *and* $p(\sigma \cdot \alpha)$ *is defined to be the (non-empty) subset of* $p(\sigma)$ *containing all models of* $p(\sigma)$ *that satisfy* α. *If no element of* $p(\sigma)$ *satisfies* α, $r(\sigma \cdot \alpha)$ *is defined to be the smallest* $n > r(\sigma)$ *such that some element of* $M(n)$ *satisfies* α. *Note that definition 1 implies there is such an* n. *The set* $p(\sigma \cdot \alpha)$ *is defined to be the (non-empty) subset of* $M(n)$ *containing all models that satisfy* α.

Now, given any widening ranked model M, a system of iterated revisions may be defined.

Definition 3 *For any sequence* σ, $[\sigma]$ *is the set of formulas that are satisfied in all models of* $p(\sigma)$.

The following soundness result is straightforward.

Theorem 3 *The revision procedure defined (as in definition 3) by any widening ranked model satisfies postulates* **I1**–**I7**.

The soundness of **I7** follows from part 1 of definition 1. The fact that $[\sigma]$ is consistent is ensured by the fact that $p(\sigma)$ is never empty.

The widening ranked models are technically much easier to manipulate and grasp than the rationality postulates **I1**–**I6**. It is therefore important to prove the converse of theorem 3.

Theorem 4 *Any revision procedure that satisfies* **I1**–**I7** *is defined by some widening ranked model.*

The proof of this result has been obtained, so far, only under the assumption of a finite number of atomic propositions, and is too long to be presented here.

Since the only widening ranked model that defines a revision that satisfies the non-postulate is the trivial widening ranked model in which all ranks contain all propositional models, we have the following.

Theorem 5 *The only revision that satisfies* **I1**–**I7** *and the non-postulate is the trivial revision.*

The trivial revision is defined by:

$$[\sigma \cdot \alpha] = \begin{cases} Cn(\alpha) & \text{if } \neg\alpha \in [\sigma] \\ Cn([\sigma], \alpha) & \text{otherwise} \end{cases}$$

As noted above, on a finite language, any widening ranked model is trivial from a certain rank onwards: all ranks are full, they contain all possible propositional models. Therefore, any revision satisfying **I1**–**I7** becomes trivial for long enough sequences of revisions.

It seems, at the present stage of this work, that **I7** corresponds exactly to the widening property of ranked models. It is clear that any ranked model in which every propositional model appears at unbounded ranks defines a revision that satisfies **I1**–**I6**. The converse seems to hold, but has not been fully proved yet.

I expressed, at the beginning of this section, the feeling that the revisions defined by widening ranked models are probably not the revisions one would like to use because they seem to provide severe revisions that methodically neglect the Principle of Minimal Change. This point will be discussed in section 8, but two preliminary remarks may be made here. First, it may be the case that suitable restrictions on the ranked models ensure compliance with the Principle of Minimal Change. The trivial ranked model in which every propositional model appears at every rank certainly does. Secondly, since any revision satisfying **I1**–**I7** may be defined by such a model, the conclusion could be that no reasonable revision procedure can satisfy the Principle of Minimal Change in a meaningful way. Since postulates **I1**–**I7** are very closely related to the AGM original postulates, this would probably signal that no reasonable revision satisfying the AGM postulates complies with the Principle of Minimal Change.

8 Conclusion

A new framework, syntax, semantics and postulates, for theory revision has been presented. Much of what has been done in the new framework may be translated back into the original AGM framework, but this new framework seems more conducive to experimentation with new postulates: for example, the strength of postulate **C1** or the triviality result of section 5 were only discovered in the new framework. A set of natural rationality postulates has been proposed. It is surprisingly powerful. A number of questions have received only partial answers. Representation theorems for weaker systems should be sought for. The quest for a suitable system for belief update is probably the most intriguing question of this type. Such a system must reject **I4** and must therefore be essentially different from [Katsuno and Mendelzon, 1992]. The postulate **I5** provides an intriguing new axiom to be introduced into Conditional Logic, or, equivalently, in the treatment of embedded conditional assertions: $\alpha > (\beta > \gamma) \rightarrow \beta > \gamma$ if $\beta \models \alpha$.

It is probably the opinion of most readers that revisions defined by widening ranked models (or just ranked models for that effect) are not satisfactory because they do not properly implement the Principle of Minimal Change. This work shows that this defect is a necessary consequence of two postulates: **I5** and **I6**, that embody some other principle of informational economy. The proper conclusion is probably that not all reasonable principles of revision are compatible.

The underlying philosophical assumptions of AGM seem to imply that contraction is the more basic operation, revision being only a derived operation consisting of a retraction followed by an expansion (the Levi identity). AI researchers will probably find this thesis questionable. It seems the operation of revising a theory in view of new information is more basic to AI than that of retracting previously received information upon learning that this previous information was obtained from some unreliable source. If such a retraction has to be performed, it seems that it will be because new information implying the unreliability of the source has been gained and therefore this retraction is really a revision by this new information. The question of the meaning of the new rationality postulates for retraction needs further study.

To conclude, I would like, using a remark of P. Schobbens, to give an example of the kind of philosophical questions that are raised by this work, but do not seem to have been remarked before. Suppose an agent learns first, a long conjunction $a \wedge b \wedge \ldots \wedge z$ and then its negation $\neg a \vee \neg b \vee \ldots \vee \neg z$. Postulate I7 implies it will forget about the first information since it has been contradicted by the second one. But one could argue that this is not the right thing to do. Granted, the first information is incorrect, but it could be almost correct. If one makes this assumption, upon receiving the second information one will conclude that few components, perhaps only one, of the conjunction are false, most of them still believed to hold true. This analysis distinguishes, I think, between two kinds of retractions. In the first one, one retracts a proposition because the source from which it has been obtained is now known to be unreliable. In this case, there is no reason to suppose the proposition is approximately correct. In the second kind of retraction, one retracts a proposition because some new information came to contradict it. In this case, one may have reason to believe the proposition is still approximately correct. This distinction should lead to two different sets of postulates for retractions.

9 Acknowledgments

This work takes on where a previous joint work with Michael Freund [Freund and Lehmann 1994] leaves off. The work of Darwiche and Pearl [Darwiche and Pearl 1994] was instrumental in convincing me to look for rationality postulates for iterated revisions. Thanks are due to David Makinson, Mark Ryan, Pierre-Yves Schobbens and Emil Weydert for useful comments.

References

[Boutilier and Goldszmidt 1993] Craig Boutilier and Moises Goldszmidt. Revision by conditional beliefs. In *Proceedings of the 11th National Conference on Artificial Intelligence (AAAI)*, pages 649–651, Morgan Kaufmann, Washington, D.C., July 1993.

[Boutilier 1993] Craig Boutilier. Revision sequences and nested conditionals. In Ruzena Bajcsy, editor, *Proceedings of the 13th I.J.C.A.I.*, pages 519–525, Morgan Kaufmann, Chambery, Savoie, France, August 1993.

[Darwiche and Pearl 1994] Adnan Darwiche and Judea Pearl. On the logic of iterated belief revision. In Ronald Fagin, editor, *Proceedings of the fifth Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 5–23, Morgan Kaufmann, Pacific Grove, CA, March 1994.

[Freund and Lehmann 1994] Michael Freund and Daniel Lehmann. Belief revision and rational inference. Technical Report TR 94-16, The Leibniz Center for Research in Computer Science, Institute of Computer Science, Hebrew University, July 1994.

[Gardenfors and Makinson 1994] Peter Gardenfors and David Makinson. Nonmonotonic inference based on expectations. *Artificial Intelligence*, 65(1):197–245, January 1994.

[Gardenfors 1990] Peter Gardenfors. Belief revision and nonmonotonic logic: Two sides of the same coin? In L. Carlucci Aiello, editor, *Proceedings of the ninth European Conference on Artificial Intelligence*, pages 768–773, London, 1990. Pitman Publishing.

[Katsuno and Mendelzon 1992] Hirofumi Katsuno and Alberto O. Mendelzon. On the difference between updating a knowledge base and revising it. In Peter Gardenfors, editor, *Belief Revision*, number 29 in Cambridge Tracts in Theoretical Computer Science, pages 183–203. Cambridge University Press, 1992.

[Kraus *et al.* 1990] Sarit Kraus, Daniel Lehmann, and Menachem Magidor. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44(1–2):167–207, July 1990.

[Lehmann and Magidor 1992] Daniel Lehmann and Menachem Magidor. What does a conditional knowledge base entail? *Artificial Intelligence*, 55(1):1–60, May 1992.

[Makinson and Gardenfors 1989]
D. Makinson and P. Gardenfors. Relations between the logic of theory change and nonmonotonic logic. In Fuhrmann A. and M. Morreau, editors, *The Logic of Theory Change, Workshop, Lecture Notes in Artificial Intelligence, Volume 465*, Konstanz, FRG, October 1989. Springer Verlag.

[Nayak *et al.* 1995] Abhaya C. Nayak, Norman Y. Foo, Maurice Pagnucco, and Abdul Sattar. Changing conditional beliefs unconditionally. private communication, 1995.

[Williams 1994] Mary-Anne Williams. Transmutations of knowledge systems. In Jon Doyle, Erik Sandewall, and Pietro Torasso, editors, *Proceedings of the fourth International Conference on Principles of Knowledge Representation and Reasoning*, pages 619–629, Morgan Kaufmann, Bonn, Germany, May 1994.