Brian Hill

Towards a "Sophisticated" Model of Belief Dynamics. Part II: Belief Revision.

Abstract. In the companion paper (*Towards a "sophisticated" model of belief dynamics. Part I*), a general framework for realistic modelling of instantaneous states of belief and of the operations involving them was presented and motivated. In this paper, the framework is applied to the case of belief revision. A model of belief revision shall be obtained which, firstly, recovers the Gärdenfors postulates in a well-specified, natural yet simple class of particular circumstances; secondly, can accommodate iterated revisions, recovering several proposed revision operators for iterated revision as special cases; and finally, offers an analysis of Rott's recent counterexample to several Gärdenfors postulates [32], elucidating in what sense it fails to be one of the special cases to which these postulates apply.

Keywords: Representations of belief, bounded rationality, logical omniscience, awareness, logical locality, belief dynamics, iterated revision, Gärdenfors postulates, rational choice theory, framing effect.

In the companion paper (*Towards a "sophisticated" model of belief dynamics. Part I*, henceforth just referred to as 'Part I'), a framework for modelling beliefs and the situations they are involved in was proposed. The aim of the current paper is to use this framework to provide a model that deals, in a realistic way, with the principal aspects of belief revision: that is, not only with the traditional postulates which are meant to apply to belief revision, but equally with the apparent infractions of these postulates. This project is construed only as an example demonstrating the power and usefulness of general framework developed in Part I.

To apply the general framework to the case of belief revision, sufficient machinery to represent beliefs and their changes will need to be added. This machinery mainly comes from existing work on belief revision, but is by no means intended to be the only possibility for developing a model of belief revision in the proposed framework. In that sense, the model presented here is just one example of a realistic model of belief revision that can be obtained using this framework. It has several natural and attractive properties: firstly, it allows recovery of the Gärdenfors postulates as applying in particular circumstances; secondly, it can accommodate iterated revisions,

Presented by **Hannes Leitgeb**; *Received* October 26, 2006

Studia Logica (2008) 89: 291–323
DOI: 10.1007/s11225-008-9130-9

recovering several proposed revision operators for iterated revisions as special cases; and finally, it offers an analysis of Rott's aforementioned counterexample to several of the Gärdenfors postulates [32], elucidating in what sense it fails to be among the circumstances to which these postulates apply. The general framework proposed in Part I thus provides a model which answers the difficult challenges posed by belief revision; this should be taken as motivation for using the framework to develop realistic models of other phenomena involving belief.

1. Introduction and state of play

A model of belief revision typically consists of (1) a model of the belief state, (2) a representation of the new information with which the state is to be revised, and (3) an operation representing the revision of the former by the latter, enjoying appropriate properties. Firstly, the operation is generally taken to satisfy a certain number of belief revision postulates, of which the most popular are the so-called Gärdenfors postulates [1, 11]; furthermore, one often expects the framework to be general enough to accommodate any revision operation which satisfies these postulates (it supports a representation theorem). Typically, in the AGM paradigm, the state of belief is taken to be a set of sentences (of a given language L) closed under a (given) logical consequence relation, and the new information consists of a sentence of this language [1, 11].[1] A popular model of belief revision in this paradigm, proposed by Grove [13], will serve as a useful example. It uses a (reflexive) order $\leq$ on the set S of maximal consistent sets of sentences of the language L — or if you prefer, possible worlds with respect to L — which has the following properties:

(S $\leq$ 1) $\leq$ is connected ($\forall x, y \in S$, $x \leq y$ or $y \leq x$);

(S $\leq$ 2) $\leq$ is transitive;

(S $\leq$ 3) $\leq$ is finitarily stoppered: for all $\phi \in L$, $|\phi| \neq \emptyset$ implies that $\{x \in |\phi| \mid x \leq y \ \forall y \in |\phi|\} \neq \emptyset$, where $|\phi| = \{x \in S \mid x \vDash \phi\}$ (that is, the set of worlds where ϕ is true).

Such an order shall henceforth be called a Grove order.

In this model, the set of beliefs is the set of sentences true in the $\leq$-minimal worlds. A sentence ψ is true after revision by ϕ if it is true in

[1] In AGM theory, the operation of contraction — removal of a belief — is taken as primitive [11]. Here only the question of belief revision shall be dealt with; contraction will not be discussed.

all the $\preceq$-minimal worlds satisfying ϕ. This model satisfies the Gärdenfors postulates (and supports a representation theorem with respect to them).[2]

Since the original models of belief revision were proposed, two other properties deemed desirable for models of belief revision have come to fore. On the one hand, there is the question of *iterated belief*: it is desirable to have a model such that, whatever results from the revision of beliefs, it can itself be revised in the face of subsequent information. The traditional AGM models, and indeed the Grove model described above, do not satisfy this condition. The Grove model, for example, yields a set of sentences after revision, but no order $\preceq'$ on S that would be appropriate for use in further revision. Since then, several models supporting iterated revision, and indeed various postulates on iterated revisions that these models satisfy, have been proposed [7, 36, 28, 25, 31, 34]. It is a generally accepted desideratum for models of belief revision that they permit iterated revisions, and therefore satisfy at least some of the proposed postulates.

A second sort of development has already been discussed: the question of the *realism* of the proposed theories of belief revision. Doubts over this issue have taken several forms: some of the debates have been localised to the validity of particular postulates [29, 30], whereas other authors have seen a general need for more "sophisticated" or "realistic" models which can account for apparent failings of certain postulates in particular situations, or overcome the apparent idealisations underpinning traditional models. Hansson [15] counts as the first problem of belief revision that of finding models which are more faithful to the finiteness of agents; Rott [32] proposes his counterexample to two of the Gärdenfors postulates in order to motivate a search for more "sophisticated" models of belief and belief change. It is incumbent upon any model seeking to capture more accurately such phenomena, such as the model of belief revision which shall be proposed in this part of the paper, to show how it deals with such problematic examples. Rott's counterexample shall be taken as a test case: a model should make it clear to what extent the Gärdenfors postulates hold, and why they do not seem to hold in this counterexample.

In the following section, a model of belief revision shall be proposed. In Section 3, it shall be shown how it satisfies the desiderata mentioned above: it satisfies the Gärdenfors postulates in appropriate cases, it models iterated revision, and recovers iterated revision operations proposed in the literature as special cases, and it supports an enlightening, "sophisticated" analysis of Rott's counterexample.

[2] For the uninitiated, it may be useful to compare this model with Lewis' semantics for counterfactuals [27]; for a detailed comparison, see [13].

Before presenting this model, it is worth noting one final development in the literature on belief revision: namely, the proposal of models or theories of belief revision where the language contains operators expressing attitudes such as belief ("the agent believes that ...") as well as operators allowing propositions to be formulated about the results of revisions ("after revision by information α ..."). Many such proposals easily extend to the multi-agent case by adding a belief operator for each agent. [40, 42, 4, 5] are some important examples of such theories.

The model proposed below will work with interpreted algebras as defined in Section I.1,[3] so that specific belief operators for different agents and operators expressing change will not be present in the model. As noted in Section I.1 and Literature Remark I.1, the model may be extended to include these operators, in a manner similar to those proposed in the literature. The choice not to include such operators is defendable for the aims of this paper, on two grounds. Firstly, given that the purpose of this section is *illustrative* — to show how the general framework proposed above can be applied to the particular problem of belief revision to yield a "realistic" model — a simpler application may prove heuristically more advantageous than a complicated one. Secondly, it is not clear to what extent one can add such operators to the language *whilst retaining the realistic credentials of the model.* Such an extension of the notion of interpreted algebra defined and motivated in Section I.1 would represent the local logical structure the agent is using at a given moment as containing sentences describing his beliefs about any sentence which is in play at that moment, his beliefs about those beliefs, and so on, as well as sentences describing the effect of learning any sentence which is in play on those beliefs, the effect of such learning on the effect of such learning, and so on, as well, in the multi-agent case, as sentences describing other agents' beliefs, and agents' beliefs about his own beliefs, and so on. It may be objected that this is not an accurate model of the local logical structure; in particular, it is questionable whether all these sentences are "in play", and thus whether he has a definite doxastic attitude — believes to be true, believes to be false, believes both the sentence and its negation to be open possibilities — to each one (as ordinary models of beliefs in terms of sets of states imply; see Section I.1). This objection is itself debatable, and in particular, may be parried if one manages to extend some of the points made in the discussion of Definition I.1 to the case of such richer languages. The second reason for shying away from the use of modal operators in this paper is to avoid such intricate debates.

[3]Henceforth, the prefix I indicates that the reference is to be found in Part I.

2. A model of belief revision

2.1. Belief states

In Section I.1, a model of the local logical structure effective at a particular moment was proposed, in the form of an interpreted algebra. The locality of this language (not all sentences of some overarching language are necessarily contained in an interpreted algebra), and of its logic (the logical consequence involved in the interpreted algebra does not necessary coincide with some global or absolute notion of logical consequence) respond well to certain limits in real agents' belief states. The basic proposal for modelling the belief state of an individual is to employ traditional models of beliefs, but, instead of using some fixed language and logical structure, to consider the beliefs of the agent at a particular moment as couched in a local logical structure which is effective at that moment. This is, so to speak, a model of the beliefs regarding that of which the agent is aware, a model adopting the agent's own point of view at a particular moment. In this section, the main details of the model shall be introduced and motivated.

The simplest model of beliefs would be, following the tradition, as a set of sentences closed under logical consequence — that is, the logical consequence of the local logical structure which is operative at the appropriate moment.[4] However, it has been suggested that correct representations of the belief *states* (sometimes called "epistemic states") of an individual should include information not only about his current beliefs, but also about how he would revise them, or, alternatively, about how "entrenched" they are [7]. Such a model of belief states shall be employed. It consists in adding a Grove order — representing not only the agent's beliefs but potential revisions of these beliefs — to the interpreted algebra representing the local logical structure in play at the moment in question. The resulting structure is called an ordered algebra.

Definition 1 (Ordered algebra). An *ordered algebra* is a pair $(\mathbf{B}, \preceq)$ where $\mathbf{B} = (B_I, B, q)$ is an interpreted algebra and $\preceq$ is a reflexive order on the atoms of B which is connected, transitive and finitarily stoppered (ie. satisfies conditions (S $\preceq$ 1-3)).[5]

[4] This model implies that the agent is *locally* logically omniscient. However, this assumption does not seem excessive for normal agents: if, at a given moment, an agent is aware of all three sentences 'A', 'if A, then B', and 'B' and he explicitly believes the first two, then it would seem that he believes the last one.

[5] Given that, in general, the interpreted algebra involved here are finite, the condition (S $\preceq$ 3) is redundant. It is retained for coherence.

Here are three examples of ordered algebra (see Example I.1 for terminology and notation).

EXAMPLE 1. For any trivial interpreted algebra $\mathbf{B}$, $(\mathbf{B}, \preceq_0)$ is a *trivial ordered algebra*, where $\preceq_0$ is the empty order.

The *point ordered algebra* for sentence ϕ, $(\mathbf{B}_{\phi_\mathbf{p}}, \preceq_{\phi_p})$ has, as interpreted algebra, the point algebra for ϕ, and, as order, the only possible one.

The *simple ordered algebra* for sentence ϕ, $(\mathbf{B}_\phi, \preceq_\phi)$, has, as interpreted algebra, the simple algebra for ϕ, and, as order, the order favouring ϕ: $q(\phi) \preceq\ q(\neg\phi)$.

Explication. Ordered algebras provide subtle models of the agent's doxastic states, in that they permit, for any given sentence, a range of "doxastic statuses". To clarify the discussion, a little preliminary terminology will prove useful.

DEFINITION 2. For an element X of the base algebra of an interpreted algebra $\mathbf{B} = (B_I, B, q)$, let $|X| = \{\phi \in \mathbf{B} | q(\phi) \geqslant X\}$.[6]

The *centre* of an ordered algebra $(\mathbf{B}, \preceq)$, denoted $|(\mathbf{B}, \preceq)|$, is $|\{x|\ x\ \preceq$-minimal$\}|$; that is, the set of elements of $\mathbf{B}$ true in all the small worlds minimal with respect to $\preceq$. The centre of a trivial ordered algebra $(\mathbf{B_0}, \preceq_0)$ shall be taken to be the set of its elements.

An element ϕ of $\mathbf{B}$ is a *generator* if $q(\phi) = \{x|\ x \preceq$-minimal$\}$.[7] In a trivial ordered algebra, every element is a generator.

An element ϕ of $\mathbf{B}$ is a *local tautology* if $q(\phi) = \top$.[8] All the elements of a trivial ordered algebra are local tautologies.

For a sentence ϕ in an ordered algebra $(\mathbf{B}, \preceq)$, there are two general senses in which it may be "believed". On the one hand, it may be a local tautology of the algebra; on the other hand, it may only belong to the centre (all local tautologies belong to the centre, but the converse does not always hold). The first case corresponds what have been called "doxastic commitments" or "irrevocable beliefs" [36]: no revision of such beliefs is admissible, since there is no world (falling under the relation $\preceq$) where they do not hold. The second case is what are called "beliefs" in Grove's model [11]: those "beliefs" which

[6]Recall that $\geqslant$ is the order of the Boolean algebra. In traditional notation (read for example in terms of small worlds), this definition would be expressed as $|X| = \{\phi | \{x|\ x \vDash \phi\} \supseteq \{x|\ x \in X\}\}$ (that is, $|X| = \{\phi | \forall x \in X,\ x \vDash \phi\}$).

[7] Since q is surjective, there is always a generator. This is natural in the finite case appropriate here.

[8]This notion can be defined on interpreted algebras, but shall only be used in the context of ordered algebras.

are not "irrevocable" may be revised if new information forces one to move to worlds where they do not hold.

However, because the ordering only applies to the local logical structure, these possibilities take on different meanings than in traditional frameworks. Invariably, a fixed language and notion of logical consequence is presupposed in the literature, so that the "irrevocable beliefs" are just the tautologies of this language. However, in the framework proposed here, where the language and the logic are local, it is *not* necessary that the *local* tautologies are tautologies of some fixed language (see Section I.1); in this sense, the believer is not modelled as (globally) omniscient. Moreover, as opposed to most traditional models, not only the centre of the ordered algebra modelling the agent's belief state, but also the local tautologies, may change over time.

The local tautologies, rather than being some fixed set of logical truths in a strong sense, are more adequately understood as those opinions which the agent cannot envisage giving up *at that particular moment.* Perhaps the term 'commitment' is appropriate for them, perhaps '*presupposition*' is more fitting; both capture the idea of acceptance without question (though the second is more amenable to change). The sentences true in the centre are the most preferred sentences *amongst those which are in play*; assuming the traditional terminology, they shall be called (explicit, instantaneous) '*beliefs*'. For an ordered algebra representing the agent's belief state, the centre will be the set of beliefs, often denoted by K (or $K_{(\mathbf{B},\leq)}$ where necessary). A generator is a belief (such as the conjunction of elements of K) which sums up exactly the belief set.

Among those sentences which are neither beliefs nor presuppositions (nor presupposed false), there are two sorts. Firstly, there are those which are in play for the agent — that is, which are contained in the interpreted algebra $\mathbf{B}$. These are true in some small worlds, but not all of the $\leq$-minimal ones. The ordinary method for revising beliefs in Grove models apply to such sentences: the set of beliefs after revision by ϕ are those sentences true in all $\leq$-minimal small worlds where ϕ. The ordered algebra thus represents the agent's opinion on how he *would* revise his beliefs by sentences which are in play for him at that moment. As shall be discussed shortly, this cannot be the *actual* revision operation, since such an operation would have to take account of revision by sentences not belonging to the original local language. So the ordered algebra only provides *envisaged revision* (or, equivalently, *envisaged* entrenchment of beliefs), but not a full measure of *actual revisions.* Note that, since revision using a Grove order satisfies the Gärdenfors postulates [13], these envisaged revisions satisfy Gärdenfors postulates relative to the local logical structure effective at that moment (ie. $\mathbf{B}$).

Observation 1. *An ordered algebra* $(\mathbf{B}, \leq)$ *satisfies the Gärdenfors postulates relative to the interpreted algebra* $\mathbf{B}$ *(that is, for the sentences of this algebra, and its notion of logical consequence).*

Relation to the Literature, Remark 1. The use of orders such as Grove orders to model beliefs (in a given logical structure) is widespread; in particular, it is popular in the dynamic logic tradition [40, 42, 4]. In this latter tradition, the order yields a notion of conditional belief, which is comparable to the notion of envisaged revision, modulo differences in the strength of language and interpretation (see Literature Remark I.1). Similarly, the distinction between envisaged revision and actual revision is similar to the distinction between conditional belief and belief revision made in [40, 4, 5].

Finally, a sentence may simply not figure in the underlying interpreted algebra; it may simply be out of play. The agent has no opinions at this moment about such a sentence. However, for the purposes of revision, it is as necessary to model the agent's reaction to such a sentence as it is to capture his response to a sentence of which he is already aware. The next question is thus that of the representation of the new information.

However, before turning to this question, let us illustrate some of these distinctions on the toy example, introduced in Part I of the paper.

Toy Example, Analysis. Part 1. As noted in part I.1 of the analysis, an appropriate interpreted algebra for modelling the local logical structure involved in Leon's belief state at the beginning of the story is $\mathbf{B_1^E}$ (see Figure I.1). $\mathbf{B_1^E}$ already captures the fact that χ (the possibility of a strengthening of the right) is out of play at that moment. A model of his beliefs at that moment will be couched in this local logical structure, but will have to contain extra structure — notably, an order on the small possible worlds. An appropriate model of the beliefs described in the story is the ordered algebra $(\mathbf{B_1^E}, \leq_1^E)$, shown in Figure 1 (p. 305). $\leq_1^E$ represents several aspects of Leon's belief state at that moment. Firstly, the fact that s_2^1 is the least preferred state represents Leon's conviction that that the far-left support would increase ($\neg\psi$) if his support were to suffer ($\neg\phi$). Secondly, the fact that s_4^1 is more preferred than s_1^1 and s_3^1 represents his belief that his party's vote is low.[9]

[9] $(\mathbf{B_1^E}, \leq_1^E)$ is but one possible representation of his belief state (given the information furnished by the story) because a choice has been made about the relationship between s_1^1 and s_3^1, namely, that they are equivalent according to the order $\leq$. As remarked at the beginning of part I.1, other choices are possible, resulting in other orders, but they can generally be dealt with along similar lines.

On the other hand, the logical structure Leon would have been using had he taken the threat of the right into account would have been different: it is the algebra $\mathbf{B}^{\mathbf{E}}_{\mathbf{1'}}$ discussed in part I.1 of the analysis. Accordingly, the beliefs he would have been represented by an order on this algebra: the ordered algebra obtained, $(\mathbf{B}^{\mathbf{E}}_{\mathbf{1'}}, \preceq^{E}_{1'})$, is shown in Figure 1. The centre of $(\mathbf{B}^{\mathbf{E}}_{\mathbf{1'}}, \preceq^{E}_{1'})$ is the same as that of $(\mathbf{B}^{\mathbf{E}}_{\mathbf{1}}, \preceq^{E}_{1})$: this reflects the fact that he would initially have had the same beliefs, even if he had not overlooked the right's threat (and in particular, he does not have any specific beliefs about the right's support, that is, about χ). However, had he not overlooked the possibility of the right gaining support, his conviction that his support is tied to the support for the far-left would have been qualified. He would still accept that, if the right had not gained support, then any fall in the far-left's support would mean a gain for him: this is represented by the fact that there is no state where all parties lose support (no state where $\neg\phi \wedge \psi \wedge \neg\chi$). On the other hand, if the right had made gains, it is plausible that a fall in the far-left's support would be accompanied by a fall in Leon's support: thus the state $s^{1'}_{2}$ may be taken as equally plausible (according to $\preceq^{E}_{1'}$) as the other states which are not quite believed.

2.2. New information

Often new information with respect to which beliefs are to be revised is treated as a simple sentence of some fixed language. However, in the framework proposed here, where no use is made of such a fixed language, incoming information will generally require a local logical structure of its own, because it may involve sentences which do not figure in the local logical structure relevant for the current belief state. Following the considerations of Section I.2, this structure shall be modelled as an interpreted algebra. However, the interpreted algebra only models the logical structure in which the incoming information is couched; it does not specify which sentences in this structure are learnt, or the extent to which the sentences of the local language are to be accepted.

For example, if the local logical structure pertinent for a case where ϕ is learnt is modelled as a simple algebra (Example I.1), some supplementary structure on this algebra would be needed to represent the fact that it is ϕ and not $\neg\phi$ which is to be accepted. It would seem natural to represent this fact with an order on the states or small worlds of this algebra which favours (the small world where) ϕ to (that where) $\neg\phi$. So doing, one obtains an ordered algebra — in fact, one obtains a *simple ordered algebra* (Example 1). On the other hand, if one uses a point algebra as model for the local logical structure, ϕ is automatically specified as input information, because it is a local tautology of this algebra. But in this case too, the structure modelling the incoming information is (trivially) an ordered algebra (the point ordered algebra, Example 1). Hence, in the most basic cases, new information can be represented as ordered algebra; the suggestion is that this sort of representation is appropriate in general.

Indeed, the representation of new information by ordered algebras inherits the advantages of ordered algebras which have been emphasised for the case of representation of beliefs discussed above (Section 2.1). As for that case, different statuses of the various elements of incoming information may be captured by ordered algebras. A sentence learnt *irrevocably* — accepted without any envisaged possibility that the new information may not hold — corresponds to a local tautology of the ordered algebra representing the incoming information (cf. the point algebra example above). On the other hand, often the information acquired comes in a context which admits that it is reliable only under certain conditions: the information gleaned from a scientific experiment, for example, does not consist of the bare result, but a collection of conditions and assumptions relating to the details of the experiment, which, if found to be false, would undermine the result. Such a situation is best represented by an ordered algebra which contains not only the sentence expressing the result of the experiment, say ϕ, but sentences expressing the appropriate conditions. In such an algebra, ϕ is not a local tautology, since conditions are envisaged in which the result of the experiment would be revoked. It is, however, the most preferred or natural option: it is in the *centre* of the ordered algebra (cf. the case of beliefs in Section 2.1). To be more pedantic, what is learnt from the experiment is characterised *precisely* by any sentence which is true only in the minimal (or most preferred) worlds of the ordered algebra — that is, by any *generator* of the ordered algebra (Definition 2). To capture revision properly, such a level of precision shall prove necessary: incoming information shall be modelled by an ordered algebra, where the sentence learnt can be thought of as a generator of the algebra.

Finally, this representation permits — thanks to the locality of the framework — that not all sentences of some fixed global language figure in the ordered algebra representing the new information; such sentences are *out of play* in the context in which this information is acquired. As has been noted in Section I.2, this aspect endows the model with the flexibility to capture accurately both the "standard" cases of inputs figuring single sentences (cf. the examples of point and simple algebras discussed above) and the more complicated cases featuring details on the conditions under which the information was acquired, possible revisions of the new information after subsequent discoveries, and so on. Citing the latter sorts of cases, the idea of representing new information by the same type of structure as that used to model the belief state has been proposed previously in the literature [28, 25].[10] However, it has always been done in the framework of a fixed global language, where the idea seems less palatable: for example, such structures determine the consequences of a storm in Tahiti for the validity of the results of a scientific experiment carried out in Tübingen. The locality of the current framework, and notably the fact that certain sentences may be out of play, allows the current proposal to avoid such counterintuitive consequences.[11]

Let us illustrate some of these points turning once again to the toy example.

Toy Example, Analysis. Part 2. Consider the second revision in the example, where Leon reads the paper. Recall from part I.2 that the logical structure involved can be represented by the interpreted algebra $\mathbf{B_3^E}$ with interpreting algebra generated by ψ ("support for the far-left has fallen") and χ ("support for the right has increased") and base algebra with the two small worlds where ψ holds. Adding the appropriate order, one obtains the ordered algebra $(\mathbf{B_3^E}, \preceq_3^E)$ representing the new information, shown in Figure 1. This representation reflects the different strengths of the information received. ψ (the fall in the far-left's support) is learnt with a large degree of certainty: it is thus a prime candidate for a sentence learnt *irrevocably*. This is represented in the model by the fact that ψ is true in all the small worlds (in other words, it is a local tautology). On the other hand, the sentence χ is by no means certain: there are small worlds where it is does not hold. However, it is argued to be true in the newspaper article: this weight in favour of χ can be represented by a preference for the state where χ over the state where $\neg\chi$. It is precisely this preference which is represented by $\preceq_3^E$.

---

[10]In a certain sense, [4, 5] also model the incoming information with the same sort of structure as that used to model the belief state. To this extent, the objection stated in the text may apply to that framework. However, as noted in Literature Remark I.1, the structure modelling the incoming information receives a significantly different interpretation from the structure modelling the belief state, and this complicates the comparison. See Literature Remark 4 for further discussion.

[11]The best that can be done in a framework that insists on ordering worlds of a fixed global language is to represent the weather in Tahiti as being *independent* from the experimental result (for example, at each rung of the order, there are worlds where the weather is good, and others where the weather is bad). However, as noted in the introduction, *being out of play is not the same thing as being independent*: if it were, then bringing something previously out of play into play would never make a difference. This delicate point, tangential to the present discussion, is considered in more detail in [19, §4.2.4].

Technically, the representation of new information by a structure of the same sort as that which represents the belief states (an ordered algebra) implies that the representation of belief revision should consist of an operation which takes two ordered algebras to a third. Such a *fusion* operation on ordered algebras can be defined from simple operations on interpreted algebras and on orders.

2.3. Revision operation

Under the current proposal, both the belief state and the new information are represented by interpreted algebras with appropriate orders on them (ordered algebras); the revision operation will somehow combine these algebras. The operation which combines the interpreted algebras (local logical structures) has already been defined and motivated: it is the fusion operation $*$ of Section I.2. The following definition and observation state that the orders on the initial interpreted algebras can be mapped canonically into the fusion interpreted algebra; furthermore, the images of the orders are still Grove orders.

DEFINITION 3. For $\preceq_i$ an order on the atoms of $\mathbf{B_i}$, for $i \in \{1,2\}$, let the *image* of $\preceq_i$ in $\mathbf{B_1} * \mathbf{B_2}$, also called $\preceq_i$, be defined as follows:

$$x \otimes y \preceq_i x' \otimes y' \text{ iff } \begin{cases} x \preceq_1 x' & \text{if } i = 1 \\ y \preceq_2 y' & \text{if } i = 2 \end{cases}$$

OBSERVATION 2. *For $i \in \{1,2\}$, if $\preceq_i$ satisfies ($S \preceq 1-3$), as an order on (the atoms of) $\mathbf{B_i}$, then its image in $\mathbf{B_1} * \mathbf{B_2}$ satisfies ($S \preceq 1-3$).*

It remains to specify how these orders, coming from the respective algebras, are to be combined. There is a selection of operations which may be employed here, several of which have been discussed in some form or another in the literature. Recall that the focus of this paper is the general framework, and the model of belief revision is just an extended example: thus, for the purposes of this paper, it would not be appropriate to enter into detailed considerations of the operations which may be used. It will suffice to pick a natural candidate and develop a revision operation built on this operation on orders. Although this candidate, and the revision operation constructed from it, has several interesting, attractive and useful properties, let it be emphasised that other operations on orders may prove equally useful, and may result, using a similar procedure to that carried out below, in equally interesting revision operations. The operation on orders used here is the lexicographic product:

DEFINITION 4 (Lexicographic product). Given two orders $\preceq_1$ and $\preceq_2$ on a set S, the *lexicographic product* $\preceq_1 \times_L \preceq_2$ is an order on S with, for all $a, b, c, d \in S$,

$$(a,b) \preceq_1 \times_L \preceq_2 (c,d) \text{ iff } \begin{cases} \textit{either } b \prec_2 d \\ \textit{or } b \cong_2 d \ \& \ a \preceq_1 c \end{cases}$$

Relation to the Literature, Remark 2. The lexicographic product features in several models of belief revision, including [28, 25, 4] (in the latter it is called the "anti-lexicographic" product). The special case where $\preceq_2$ merely distinguishes the states where ϕ from the states where $\neg\phi$ (for some ϕ) is often called "lexicographic" revision, and has received significant attention (see for example [34, 40]). For further discussion of the relation between these uses of the lexicographic product, and a comparison with the use made of the notion here, see Section 3.2 and Literature Remark 4 (as well as the remarks at the end of Section 2.2).

This product has the following two useful properties. Firstly, it is non commutative: indeed, it gives priority to one of the orders over the other. This fits well with the idea that new information should have priority over previous beliefs. Secondly, as an operation on Grove orders, it yields a Grove order, and thus, when used to combine ordered algebras, it guarantees that the resulting structure will be an ordered algebra.

OBSERVATION 3. *If $\preceq_1$ and $\preceq_2$ satisfy ($S \preceq 1-3$), then $\preceq_1 \times_L \preceq_2$ satisfies ($S \preceq 1-3$).*

Therefore the *fusion* operation, consisting of the fusion of interpreted algebras (Definition I.5) and the lexicographic product of the images of the initial orders in this fusion, is a well-defined operation taking ordered algebras to ordered algebras.

DEFINITION 5 (Fusion $*$ of ordered algebras). Let $(\mathbf{B_1}, \preceq_1)$ and $(\mathbf{B_2}, \preceq_2)$ be ordered algebras, and $\simeq$ an identification relation between them. The *fusion* is defined as:

$$(\mathbf{B_1}, \preceq_1) *_{\simeq} (\mathbf{B_2}, \preceq_2) = (\mathbf{B_1} *_{\simeq} \mathbf{B_2}, \preceq_1 \times_L \preceq_2)$$

The subscript $\simeq$ shall be omitted when clear from the context.

$(\mathbf{B_1}, \preceq_1)$ represents the initial belief state: its centre is the set of beliefs (Section 2.1). $(\mathbf{B_2}, \preceq_2)$ represents the new information: the sentence learnt is a generator (Section 2.2). $(\mathbf{B_1}, \preceq_1) * (\mathbf{B_2}, \preceq_2)$ represents the resulting belief state: its centre is the new set of beliefs. Note that, by the definition of lexicographic order given above (Definition 4), priority is accorded to new information over previous beliefs.

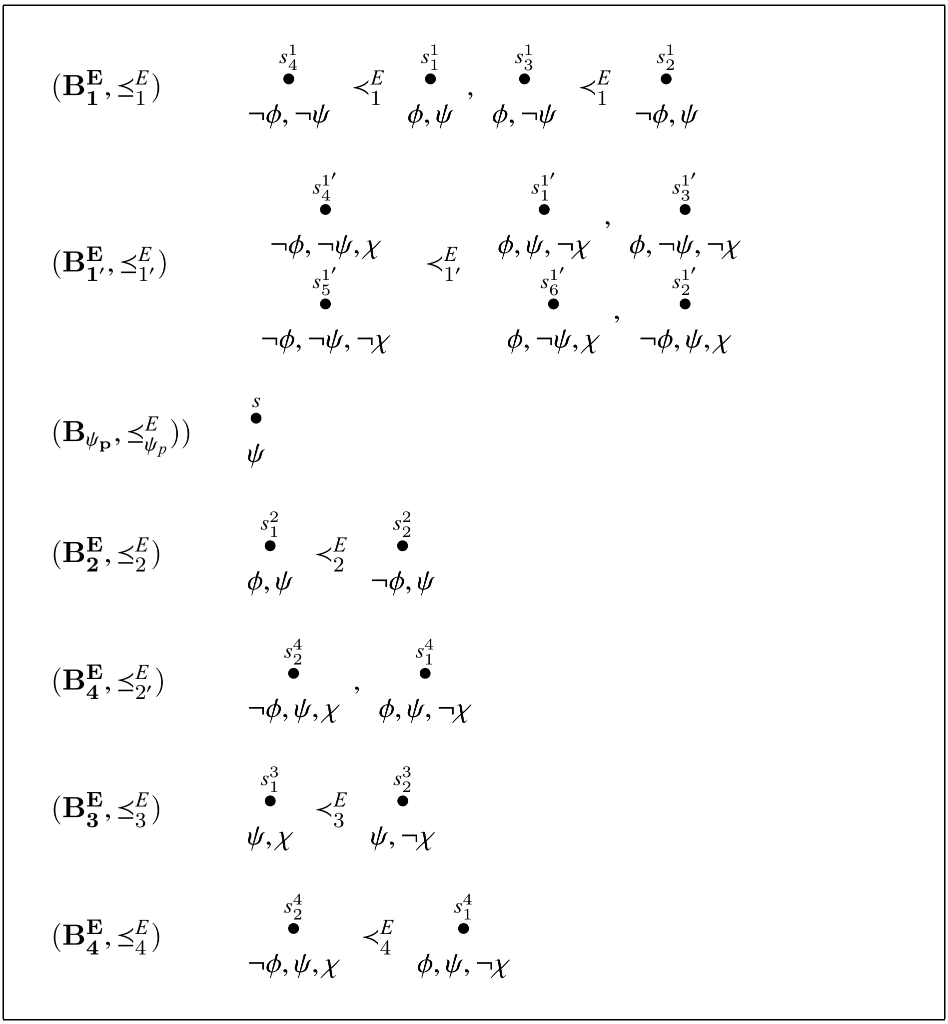
Having introduced the fusion operation on ordered algebras, the analysis of the toy example may now be completed.

TOY EXAMPLE, ANALYSIS. PART 3. Recall (part I.2 of the analysis) that the logical structure of the new information involved in the first revision (before the rally) can be modelled by the point algebra $\mathbf{B}_{\psi_\mathbf{p}}$ (where ψ is "support for the far-left has fallen"); accordingly the new information itself — obtained by adding the appropriate order to the local logical structure — will be modelled by the point ordered algebra $(\mathbf{B}_{\psi_\mathbf{p}}, \preceq_{\psi_p})$ (Example 1). The revision is thus modelled by the fusion of the initial belief state, $(\mathbf{B}^\mathbf{E}_\mathbf{1}, \preceq^E_1)$ (see part 1) with this point algebra: this yields the ordered algebra $(\mathbf{B}^\mathbf{E}_\mathbf{2}, \preceq^E_2)$, represented in Figure 1. The centre of this ordered algebra contains ϕ ("support for Leon's party has risen"): this represents the fact that he has come to believe that his party's support has increased.

On the other hand, the revision he would have undertaken, had he had in mind the possibility of the right gaining support, is represented by the fusion of $(\mathbf{B}^\mathbf{E}_{\mathbf{1'}}, \preceq^E_{1'})$ with $(\mathbf{B}_{\psi_\mathbf{p}}, \preceq_{\psi_p})$; this yields the ordered algebra $(\mathbf{B}^\mathbf{E}_\mathbf{4}, \preceq^E_{2'})$ shown in figure 1 (see also part 1). The centre of this algebra contains ψ but not ϕ: this reflects the fact that, if the possibility that the right has made gains had been in play, Leon would have been unsure as to whether to attribute the far-left's drop in support to a strengthening of the right (and thus a fall in his party's support) or a strengthening of his party's support (without a strengthening of the right). Learning ψ would not lead him to believe ϕ: the analysis thus captures the differences between the case where the possibility of the right's strengthening is overlooked and the case where it is recognised.

Turning to the second revision, this is represented by the fusion of $(\mathbf{B}^\mathbf{E}_\mathbf{2}, \preceq^E_2)$ (the belief state resulting from the first revision) with $(\mathbf{B}^\mathbf{E}_\mathbf{3}, \preceq^E_3)$ (the new information; see part 2). This produces the ordered algebra $(\mathbf{B}^\mathbf{E}_\mathbf{4}, \preceq^E_4)$, represented in Figure 1 (see parts I.3 and I.4 on the fusion of the interpreted algebras). It nicely represents three of the main factors of the change in Leon's belief state upon reading the news report: firstly the fact that the question of the right's support comes into play (χ is in $\mathbf{B}^\mathbf{E}_\mathbf{4}$); secondly, the fact that he has reneged on his former belief that his party is doing well (ϕ is no longer in the centre); and finally the fact that, given the argument of the newspaper article, has come to believe that he is worse off ($\neg\phi$ is in the centre). This revision is thus well-modelled by the concepts developed here, even if it may seem to contradict some of the traditional Gardenförs postulates, to the extent that the information learnt for sure — ψ — was already known. As shall be discussed in Section 3.1, and more at length in Section 3.3, the fusion operation is completely coherent with the Gardenförs postulates construed as applying only in a special sort of case; this example, just as Rott's purported counterexample to the postulates, is not one of these special cases.

Figure 1. Toy example: Representations of the ordered algebras



3. Properties of the model

The toy example suggests that the operator $*$, taken with the interpretation of ordered algebras as representations of belief states and incoming information, provides a flexible modelling framework capable of capturing some of the subtleties of actual belief revision. It also possesses the technical properties of belief revision operators mentioned in Section 1. Firstly, it satisfies an appropriate version of the well-known Gärdenfors postulates for belief revision. Secondly, since it satisfies the property of *categorial matching* [12] — the representation of the belief state after revision (ordered algebra) is of the same format as the representation before revision (and thus appropriate for further revision) — it is automatically an *iterated* revision operator. Furthermore, two important iterated revision operators proposed in the literature [36, 28, 25] can be recovered in the proposed framework as *special cases* corresponding to particular constraints placed on the incoming information. Finally, the proposed framework supports an analysis of Rott's aforementioned counterexample to two of the Gärdenfors postulates [32], identified in Section 1 as an important litmus test of "realistic" models of belief revision. These points shall be dealt with successively.

3.1. Gärdenfors postulates

The theorem detailing the postulates satisfied by $*$ is complicated by the fact that traditional statements of the Gärdenfors postulates suppose a global language, whereas such an assumption is explicitly avoided here. To formulate the theorem, the following preliminary definition is thus required.

Definition 6. Let $(\mathbf{B}, \preceq)$ be an ordered algebra with generator ϕ. For ψ in $\mathbf{B}$, the *refinement* of $(\mathbf{B}, \preceq)$ with generator $\phi \wedge \psi$ is the ordered algebra $(\mathbf{B}, \preceq')$, with

$$x \prec' y \text{ iff } \begin{cases} x \leqslant q(\psi), y \nleqslant q(\psi) & \text{if } x, y \leqslant q(\phi) \\ x \prec y & \text{otherwise} \end{cases}$$

This ordered algebra has generator $\phi \wedge \psi$.

Using the model of belief states and new information proposed in Section 2, with the interpretation of the set of beliefs and the sentence learnt as the centre and the generator of the respective ordered algebras, the operator $*$ satisfies the Gärdenfors postulates in the following sense.

Theorem 1. *Let* $(\mathbf{B_1}, \preceq_1)$ *be a non trivial ordered algebra with centre* K*, let* $(\mathbf{B_2}, \preceq_2)$ *contain sentences* ϕ *and* ψ *and have generator* ϕ*, let* $(\mathbf{B_2}, \preceq_3)$ *be its* *refinement with generator* $\phi \wedge \psi$*, and let* $\simeq$ *be the identification relation between* $\mathbf{B_1}$ *and* $\mathbf{B_2}$*. Let* $K*\phi$ *(resp.* $K*(\phi\wedge\psi)$*) be the centre of* $(\mathbf{B_1}, \preceq_1) *_{\simeq} (\mathbf{B_2}, \preceq_2)$*, (resp.* $(\mathbf{B_1}, \preceq_1) *_{\simeq} (\mathbf{B_2}, \preceq_3)$*). Then*

(K ∗ 1) $K * \phi = Cn_{12}(K * \phi)$*;*

(K ∗ 2) *If* $\bot_1 \not\simeq_B q_2(\phi)$ *or* $\top_2 \Rightarrow_2 \phi$*, then* $\phi \in K * \phi$*;*

(K ∗ 3) $K * \phi \subseteq Cn_{12}(K \cup \{\phi\})$*;*

(K ∗ 4) *If, for each* χ *such that* $q_1(\chi) \simeq_B q_2(\phi)$*,* $\neg\chi \notin K$*, then* $Cn_{12}(K \cup \{\phi\}) \subseteq K * \phi$*;*

(K ∗ 5) *If* $\bot_1 \not\simeq_B q_2(\phi)$*, then* $K * \phi$ *is consistent under* Cn_{12}*;*

(K ∗ 6) *If* $\phi \Leftrightarrow_2 \chi$*, then* $K * \phi = K * \chi$*;*

(K ∗ 7) $K * (\phi \wedge \psi) \subseteq Cn_{12}(K * \phi \cup \{\psi\})$*;*

(K ∗ 8) *If* $\neg\psi \notin K * \phi$*, then* $Cn_{12}(K * \phi \cup \{\psi\}) \subseteq K * (\phi \wedge \psi)$*.*

where $\Rightarrow_1$*,* $\Rightarrow_2$*,* $\Rightarrow_{12}$*,* Cn_{12} *and so on are the consequence relations and sets of consequences in* $\mathbf{B_1}$*,* $\mathbf{B_2}$*, and* $\mathbf{B_1} * \mathbf{B_2}$ *respectively.*

Discussion. The Gärdenfors postulates are normally expressed in terms of sentences and sets of sentences (see, for example [11, §3.3]), whereas the basic notion here is that of ordered algebra. As discussed above, ordered algebras offer a more flexible and general representation of beliefs and new information; accordingly, the subtlety in Theorem 1 is required to curb this flexibility and recover the simpler case dealt with traditionally.

The most important aspect is the translation of the traditional notions of sets of beliefs and sentences learnt as centres and generators of the ordered algebras representing the belief state and the new information ($(\mathbf{B_1}, \preceq_1)$ and $(\mathbf{B_2}, \preceq_2)$). As discussed in Section 2, the interpretation of 'belief' and 'sentence learnt' as "most preferred sentence" is *only one* of the interpretations which could be afforded to these terms by a model consisting of ordered algebras, albeit a particularly natural and intuitive one. It is thus important to emphasise that, to the extent that Theorem 1 states that the Gärdenfors postulates are satisfied, it *only* affirms that they are satisfied when *applied to the centres and generators of ordered algebras.* In particular, the postulates *do not* necessarily apply to the *local tautologies* of the respective algebras, that is, to sentences which have been called "commitments", "presuppositions", or "irrevocable sentences" in Section 2. This is one concrete sense in which the model of belief revision proposed here recovers the traditional theory as an *idealisation*: the Gärdenfors postulates hold, but only in the *special cases* where centres and generators (and the corresponding notions of "belief" or "information") are used. Hence, this theorem shows that one of the desiderata of a more realistic model of belief dynamics — namely, to exhibit the sense in which previous theories are idealisations — is fulfilled.

With respect to the details, there turn out to be two principal types of difference between the postulates in Theorem 1 and the standard formulation of the Gärdenfors postulates [11, §3.3], both of which arise from the increased generality of the proposed framework.

(i) The traditional formulation supposes a fixed consequence relation which figures in all the postulates. Since all the consequence relations involved here are local — relative to particular interpreted algebras — each of the Gärdenfors postulates have to be modified to specify *which* consequence relation is involved. This factor requires particular care when specifying the ordered algebra used to represent the new information $\phi \wedge \psi$ in axioms (K $*$ 7 – 8) — thus the appeal to the notion of refinement (Definition 6).

(ii) Several of the traditional Gärdenfors postulates involve considerations relating to the case where the new information ϕ is contradictory; for example (K $*$ 4) and (K $*$ 5) traditionally contain clauses requiring that ϕ not be contradictory or not be contradictory with the belief set. In the current framework, where ϕ may belong to a different interpreted algebra than that in which the beliefs are couched, what counts is not whether it is contradictory (respectively, contradictory with the belief set) in the interpreted algebra involved in the model of the new information, but whether its image in the interpreted algebra obtained by fusion is contradictory. Thus, for the case of (K $*$ 4) and (K $*$ 5) for example, the traditional formulations are replaced by clauses stating that ϕ is not $\simeq_B$-equivalent to the false element or to an element contradictory with the belief set.
Similarly, (K $*$ 2) — originally of the form "$\phi \in K * \phi$" — relies on the assumption that if ϕ is contradictory in the logical structure of the belief set, then it is contradictory in the logical structure of the incoming information. However, this is not necessarily the case in this setup, and there are cases where the original form of the postulate is violated. To illustrate the point, suppose that $\neg\phi$ is a local tautology of the ordered algebra representing the belief state $(\mathbf{B}, \preceq)$ and that the new information is represented by the simple algebra $(\mathbf{B}_\phi, \preceq_\phi)$: since there are no ϕ-small worlds to place before the $\neg\phi$-small worlds in the lexicographic product, fusion yields the initial ordered algebra — and ϕ

is still not believed. This is exactly the case ruled out by the condition added to (K $*$ 2).[12]

These complications result from a liberalisation of the traditional framework, in the sense that richer representations of the belief state, of the new information, and of the logical structures involved, are permitted. This can be seen most clearly by remarking that, when some of these liberalisations are removed, the complications dissolve. In particular, if one uses the most basic algebra to model the incoming information — that is, a point algebra (Example 1) — the conditions discussed in point (ii) are automatically satisfied, and the postulates stated in Theorem 1 reduce to the standard Gärdenfors postulates.

Remark 1. Recall (Observation 1) that the order on an ordered algebra defines a revision operation with respect to sentences belonging to the algebra and this operation satisfies the Gärdenfors postulates. As one would expect, in the case of revision by a sentence already belonging to the ordered algebra representing the belief state, this revision operation coincides with the more general revision operation $*$; this result is stated as Proposition 1 in the Appendix. It exhibits the extent to which the envisaged revisions agree with actual revisions.

Relation to the Literature, Remark 3. Most AGM-styled approaches to belief revision show that the Gärdenfors postulates — or appropriate Gärdenfors-like postulates — hold. On the other hand, Dynamic Logic-styled approaches generally do not emphasise this aspect as much, largely because the postulates do not hold, in their original form, when beliefs can be expressed in the language (that is, when using a language with a belief operator; see [40, 42]). As noted previously, the model presented here is situated more in the former tradition, at least in so far as there are no modal operators in the language. In the models developed in the latter tradition, the Gärdenfors postulates purportedly appear as a special case:[13] they apply only on the fragment of the language not containing modal operators, the fragment of so-called "factual statements" [40]. However, given that all the statements are factual here, the sense in which the postulates are said to apply in a special case in those theories must be different from the sense in which they apply in a special case in the model proposed here. As noted, the important point here is

[12] This particular infringement of Gärdenfors (K $*$ 2) is in fact quite widespread, affecting the "lexicographic upgrade" and the "elite change" in [40], as well as the "successful minimal belief revision" in [42] ([42, §6] does not seem to notice this fact, whereas [40, Remark 14] expresses it technically without noting it explicitly).

[13] But see footnote 12.

not so much the "content" of the information (factual or doxastic), but its position in the context in which it is couched: roughly, certain elements of the ordered algebra modelling the belief set and the incoming information satisfy Gärdenfors-like postulates (the centre and the generator), others do not (the local tautologies). This is a difference which cannot be accounted for on any approach which models the new information simply as a sentence (such as [40, 42] and the AGM tradition). For discussion of other approaches which have complex representations of the information, see Section 2.2 and Literature Remarks I.1 and 4.

3.2. Iterated Revision

As noted above, $*$ is an iterated revision operator, in that it yields a structure (ordered algebra) fit for subsequent revision (using $*$). As in the discussion of the Gärdenfors postulates above, the realistic credentials of this model can be further bolstered by showing in what sense several notions of iterated revision proposed in the literature can be recovered as special cases of revision with $*$, indeed, as special cases where the new information comes in a particular format. The two iterated revision operators considered, called "radical" and "moderate" revision, are characterised by the following postulates.

Definition 7. **(Rad)** $(K * \phi) * \psi = K * (\phi \wedge \psi)$

(Mod)

$$(K * \phi) * \psi = \begin{cases} K * (\phi \wedge \psi) & \text{if } \psi \text{ is consistent with } \phi \\ K * \psi & \text{otherwise} \end{cases}$$

By requiring that the sentences with respect to which the revisions are made be represented as coming in a particular form — respectively as point or simple algebras (Example 1) — one or other of the postulates are automatically satisfied.

Theorem 2. *Let K be the centre of the ordered algebra $(\mathbf{B}, \preceq)$, and let consistency in Definition 7 be understood as it not being the case that $q(\phi) \simeq_B q(\neg\psi)$, where $\simeq$ is the identification relation between ordered algebras representing the information ϕ and ψ respectively. Then*

(Rad) *If ϕ and ψ are modelled by $(\mathbf{B}_{\phi_\mathbf{p}}, \preceq_{\phi_p})$ and $(\mathbf{B}_{\psi_\mathbf{p}}, \preceq_{\psi_p})$ respectively, then **(Rad)** is satisfied.*

(Mod) *If ϕ and ψ are modelled by $(\mathbf{B}_{\phi}, \preceq_{\phi})$ and $(\mathbf{B}_{\psi}, \preceq_{\psi})$ respectively, then **(Mod)** is satisfied.*

Remark 2. As is usual in the more general localist framework adopted here, there are several possible interpretations which can be given to the notion of "consistency" in the postulate for moderate revision: one might ask for consistency in so far as the sentences appear as elements of the algebra representing the initial belief state K (a condition which assumes they both appear in this algebra), or as a relation between the sentences as they appear in the algebras representing the new information. The second option is adopted in the statement of the theorem; however, as can be noted from the proof, a similar result holds if the first notion of consistency is used (see Remark 5 in the Appendix).

Relation to the Literature, Remark 4. The terms "radical" revision and "moderate" revision are coined by Rott in [31, 34]. The former has been proposed and defended by [36], where it was called "irrevocable" revision; it has also been called "maximal" belief revision in [42] and "belief change under hard information" in [40, §3.3] and [4, Ex 3(b)]. The latter was apparently first suggested in [37], first defended by [28], and discussed in [25] (where it is called the "basic memory operator"), [40, §4.1] and [4, Ex 3(c)] (where it is called "lexicographic update").

Note that, in most of these cases, different models of revision are proposed for the different operations, and the relationship between them is seldom discussed. The notable exception is [4]. There, new information is modelled by "action plausibility models", which are comparable to ordered algebra, except for the issues regarding the explicitness and the richness of the language (see Literature Remarks I.1 and 1); furthermore, update is modelled by the "anti-lexicographic" product, which is essentially equivalent to what is here called the lexicographic product (Definition 4; see also Literature Remark 2). As one would expect given this technical similarity, action models corresponding to point and simple ordered algebra yield the sorts of revision operations described above [4, Exs 3(b), 3(c)]. There is, however, a subtle difference in interpretation. The model in [4] is considered to be a general model of types of actions — that is, in this case, types of revision operations (Literature Remark I.1). The moderate and radical revision operations are thus considered to be different operations, modelled by different action models, albeit action models which can be defined in a common framework. To this extent, that approach is in keeping with the tradition which treats the operations separately. By contrast, the model proposed here involves one operation — fusion (Definition 5) — but a range of formats or structures which the new information can take. What has traditionally been thought of as different revision operations now come to look like special cases of a single revision operation; namely, special cases where the incoming information has a particular form (respectively, that of a point or a simple algebra). This is the sense in which this framework recovers the traditional, independently proposed revision operations and singles out the cases where they occur. Finally, this interpretation opens the question, muted below, as to whether the range of different belief revision operations could be replaced by a single operation, and a sophisticated model of the incoming information.[14]

This theorem counts as a further illustration of the fruitfulness of the proposed model of belief revision, and more generally of the framework of which it is an instance (the framework developed in Part I). Iterated revision operations proposed in the literature are apparently recovered as special cases of the form of the input information. In the sense in which they suppose that the input information takes a particular form, they are *idealisations*; in the sense in which the model proposed here does not make this supposition, and indeed can accommodate a multiplicity of possible formats for the incoming information, it is more *realistic*. Although further discussion will take us too far from the motivational purpose to which this example has been put here, it is worth noting that the idea that the plurality of belief revision operators might boil down to a single general operation with a plurality of formats for incoming information is not only philosophically (and technically) interesting, but may find support, if not a precursor, from unexpected quarters. Friedman and Halpern [9] argue for a plurality of belief revision operators, each appropriate to a different "ontology". However, given that their "ontologies" are differentiated largely by properties of the incoming information (whether it is absolutely certain or only has a certain plausibility relative to other beliefs), one is justified in hoping that the plethora of revision operators may reduce to a single, general, revision operator, and a range of possible formats for the new information.

3.3. Rott's example

The final sort of problem for an alleged realistic model of belief and the operations involving it is to account for the apparent counterexamples to properties or postulates proposed by existing theories. These come in the form of cases where it seems that actual agents do or would behave in a way that does not conform to existing theories. They pose a challenge to any

[14]Further research is required to ascertain whether this is possible. Certainly the lexicographic-based fusion operation proposed here will not be able to serve this purpose (at least in the case of classical propositional languages considered here), and is used merely for purposes of illustration.

modelling project, such as the one undertaken here: such a project cannot just dismiss this sort of behaviour as "irrational" or "illogical", but must show how the proposed model can account for it.

In the case of belief revision dealt with in this part of the paper, the cases in question are those where an agent seems to revise his belief in a way which does not adhere to the Gärdenfors postulates or other axioms posed by models of revision. The toy example contained such a case, and the analysis offered by the model developed here already indicates that it can naturally account for cases like these. In this final section of the paper, the point will be demonstrated more systematically, by using the model to analyse a significant test case, namely the counterexample that Rott has recently proposed to several of the Gärdenfors postulates [32]. To the knowledge of the author, no analysis of this counterexample has been offered, although this was precisely what Rott called for. The challenge is to propose a model of belief revision which makes clear why and to what extent the Gärdenfors postulates do not hold in the example. The model proposed here shall rise to the challenge, explaining the counterexample by relying on the fact, underlined above, that the Gärdenfors postulates only apply in special cases. It shall show that the situation involved in the counterexample does not fall among those special cases to which the postulates apply. Such counterexamples are thus not considered to be rejections of the postulates, but rather symptoms of their limited range of validity.

The counterexample. The story proposed by Rott [32] is a counterexample to the Gärdenfors postulates (K $*$ 7) and (K $*$ 8),[15] which consists of a story concerning an agent who considers the candidates for a metaphysics post in a philosophy department. Four candidates appear in the story: d, who is by far the best, a, who excels in metaphysics (but has no pedigree in logic), b, who is good at metaphysics and logic, and c, who is outstanding in logic but mediocre in metaphysics. The agent is thus originally of the opinion that d will get the post. However, d cannot take the post, and so the agent's beliefs will need to be revised. Rott considers two alternative revision scenarios.

I. The agent is told (by a reliable source) that a or b will get the post; he according revises his belief, and takes the opinion that a will get it.

II. The agent is told (by a reliable source) that a or b or c will get the post. This "triggers off a rather subtle line of reasoning" [32, p. 230].

[15]It is in fact a counterexample to a slightly weaker postulate than (K $*$ 8), but that shall not matter for present purposes.

> Given c's pedigree in logic, and lack of pedigree in metaphysics, and given that the committee is still considering c's application, the agent figures that not only competence in metaphysics, but also competence in logic, are criteria for the post. Therefore, he reasons, b's competence in both domains, combined with a's restriction to metaphysics alone, give b the advantage: he comes to believe that b will get the post.

This contradicts Gärdenfors' (K $*$ 7) in the following way.[16] Without risk of confusion, let a, b, c, d be the propositions expressing that a, b, c and d respectively get the post. Let K be the agent's prior set of beliefs. The described revision patterns yield:

$$K * ((a \vee b \vee c) \wedge (a \vee b)) = K * (a \vee b) = Cn(a) \qquad \text{(Sitn I.)}$$
$$Cn((K * (a \vee b \vee c)) \cup \{a \vee b\}) = Cn(b) \qquad \text{(Sitn II.)}$$

Whence $K * ((a \vee b \vee c) \wedge (a \vee b)) \not\subseteq Cn((K * (a \vee b \vee c)) \cup \{a \vee b\})$, *contra* (K $*$ 7), which states that this inequality holds. Any satisfactory realistic model of belief revision should explain *why* the agent revises his belief in such a way, and in exactly what sense he does and does not respect the Gärdenfors postulates.[17]

Analysis. As the setup is described, there are basically four sentences in play — a, b, c, d — which cover all the possibilities and are mutually exclusive. The (minimal) interpreted algebra representing the local logical structure in which the agent's initial belief state is couched, call it $\mathbf{B}$, will contain (essentially) these four sentences, and four small worlds, such that each sentence shall be true in one and only one world. Without risk of confusion, the four worlds shall also be called a, b, c and d, according to which of the sentences is true in that world. As the initial belief state of the agent is described (he thinks only metaphysics is important), it should be represented by the order $\preceq$ on $\mathbf{B}$, where $d \preceq a \preceq b \preceq c$.

Note that, if the information learnt in situation I., *a or b*, is represented as a point algebra, the revision operation $*$ yields the desired result: a becomes the new belief ($\preceq$-minimal state). If, on the other hand, a point algebra representation of the information is used in situation II., revision with $*$ yields a and not b as in the story. Indeed, according to the proposed model,

[16] The case of (K $*$ 8) is similar and shall not be reproduced here.

[17] As noted in the introduction, this requirement on models of belief revision does not imply that they treat such cases as "rational", nor that they cannot offer a notion of rationality.

the counterexample arises because the point algebra *does not* properly represent the information the agent sees himself as receiving, or to put it another way, the information with respect to which the agent revises is not *just* the bare sentence *a or b or c*. In fact, a different representation is more appropriate, and under revision with this representation, the Gärdenfors postulates do not necessarily hold.

To see that the point algebra is not an accurate representation of the incoming information, compare situation II. with situation III.[18]

III. The agent is told (by a reliable source) quite simply that *d* will not take the job. No "subtle line of reasoning" is triggered off by this information: the agent alters his opinion and plumps for the next best candidate. He comes to believe that *a* will get the job.

Situations II. and III. yield different revisions of belief, *despite the fact that the information acquired in the two cases are equivalent* in the context of the example.[19] This state of affairs is reminiscent of a well-known phenomenon in choice theory, the *framing effect* [39], where equivalent choice problems[20] extract different choices from agents. Here, equivalent formulations of the same input information yield different revisions of beliefs. And, just as Rott's example apparently contradicts postulates of belief revision, the framing effect apparently invalidates axioms of rational choice.

Given that the problem in the framing effect is that the *presentation* of the choice problem plays a decisive role, a natural approach is to explicitly model the effect of the presentation *on the choice problem which the agent sees himself as faced with.* Due to differences in presentation, two choice problems which are *logically* equivalent will result, after processing of the presentation, in *logically non-equivalent* choice problems which the agent sees himself as solving. This is the sort of tactic adopted in Kahneman & Tversky's prospect theory [24, 39], where the choice process involves a "phase of framing and editing" which precedes evaluation and yields the choice problem which the agent effectively evaluates. According to this sort of theory, the basic evaluation of a choice problem is always the same — and indeed it satisfies traditional rationality postulates — but differences in the presen-

[18]In the conclusion of his paper, Rott alludes to a comparison with a case similar to the one described below. A detailed discussion of the differences, and similarities, between Rott's remarks and the analysis proposed here lies beyond the limits of this paper.

[19]More rigorously, *a or b or c* and *not d* are equivalent sentences of the interpreted algebra involved in the representation of the agent's belief state.

[20]Or, if you prefer: different formulations of the same choice problem.

tation of a choice problem may yield logically non-equivalent results of the "phase of framing and editing", and thus different choices by the agent.

Given the considerations above, which extend the similarities between choice theory and belief revision which have already been identified by some theorists, Rott's example seems ripe for this sort of treatment.[21] Comparison between the simple story in situation III. and the intricate story in situation II. seems to suggest that, while in situation III. the incoming information (*not d*) is directly employed to revise beliefs, in situation II., the information seems subject to a preliminary treatment before being used for revision. How else should one understand the "subtle line of reasoning" involved in situation II., if not as a phase of *pre-processing* of the incoming information?

Understood in this way, the information by which the agent *effectively* revises in situation II. is not the simple sentence *a or b or c*; in other words, the new information *post-processing* is thus not properly represented by the point algebra $\mathbf{B}_{\{\mathbf{a\ or\ b\ or\ c}\}_\mathbf{p}}$. Rather, by his "subtle line of reasoning", the agent establishes a preference among the options a, b and c. A better representation of the information with respect to which he revises is an ordered algebra whose interpreted algebra, $\mathbf{B}'$, contains the sentences a, b, c, and three worlds (also called a, b and c), each validating one and only one of the sentences and whose order, $\preceq'$, is such that $b \preceq' a \preceq' c$. The agent accepts as certain that *a or b or c* (it is a local tautology), but, on reflection, accords a preference to b over a and c. On the other hand, in situation III., which involves less pre-processing, the sentence *not d* is directly and simply used to revise: the incoming information can thus be represented by the point ordered algebra $(\mathbf{B}_{\{\mathbf{not\ d}\}_\mathbf{p}}, \preceq_{\{not\ d\}_p})$.[22]

In the context of the story, $(\mathbf{B}_{\{\mathbf{not\ d}\}_\mathbf{p}}, \preceq_{\{not\ d\}_p})$ and $(\mathbf{B}', \preceq')$ have equivalent sets of *local tautologies*. In particular, both have *not d*, or equivalently *a or b or c*, as local tautologies: this information is *irrevocably* learnt in both cases. This is the precise sense in which the two situations involve revision

[21] The tight relationship between belief revision and choice theory has already been explored in the literature, notably in [30]; indeed Rott's counterexample is in a certain sense the fruit of such work. However, to the knowledge of this author, little work has been done on the relationships between the "realism" of the postulates of choice theory and those of belief revision. This section can be seen as a first step in this direction. Note finally that such relationships are perfectly coherent with the project of this paper: to provide a general framework which uses similar tools to treat a range of different issues.

[22] As for the toy example (see part I.1 of the analysis), different choices in the analysis — of sitauation II. (using an algebra containing d, or a different order between a and c) and of situation III. (using a simple algebra instead of a point algebra) — will yield similar results, though they may require a wider range of operations on, and relations between, ordered algebra; see Section I.2 and [19, Ch 5].

by equivalent information: the *local tautologies* are equivalent. However, as was underlined in Section 3.1, the Gärdenfors postulates *do not apply* to the local tautologies but only to the centres and generators of the ordered algebras (Theorem 1). So it is not expected that revision by these two algebras yield the same result. Indeed, it is straightforward to show that revision by these two algebras yield the results described in the story: a for $(\mathbf{B}_{\{\mathbf{not\ d}\}_\mathbf{p}}, \preceq_{\{not\ d\}_p})$ (Situation III.); b for $(\mathbf{B}', \preceq')$ (Situation II.). Moreover, no Gärdenfors postulates are violated in either revision.

This analysis calls for further discussion. There are things to be said about the relationship between the notions of local tautology and generator, and the certainty and irrevocability of the sentence learnt. Similarly, work is required on the pre-processing phase, in which the appropriate representation of the incoming information (ordered algebra) is formed. Some of this may require the use of operators on ordered algebras other than the fusion operation defined here (see [19, Ch 5] and [21] for some relevant operators). Without such a discussion, a full realistic theory of belief revision has not been completed. But this was not the aim of this second part of the paper: rather, the discussion of belief revision was intended to illustrate the fruitfulness of the general framework introduced in Part I. This goal has been achieved: the model of belief revision generated using only a limited range of tools is already capable of recovering the traditional postulates and those for iterated belief revision, revealing in what sense they are idealisations (they apply in special cases). Furthermore, it provides an *analysis* of apparent counterexamples to these postulates, in so far as it exhibits in what sense the cases do not satisfy the conditions required for the postulates to apply. In other words, where traditional theories and models of belief revision cannot cope with these examples, the proposed framework provides an *understanding* of them. This is an indication of the strength of framework in so far as it is used to *represent* belief revision: it captures phenomena that simpler frameworks miss. Moreover, it is an indication of the framework's strength as a *conceptual tool* for studying belief revision: the analysis of the counterexample suggests an approach towards a full realistic theory of belief revision, according to which the question is factorised into two parts — the pre-processing part and the revision proper — of which the latter is captured by the operation $*$ and the former may benefit from expression in terms of the framework proposed here.[23]

[23] In the concluding discussion of his counterexample, Rott suggests that a formalisation is needed which is sufficiently rich to take account of examples such as the one he proposed, whilst remaining operational (or "processable") and eschewing gratuitous addition

One can conclude that the model of belief revision proposed in the second part of this paper is "realistic" and "sophisticated", and indeed seems to open up fruitful possibilities for development into a full theory of belief revision. The general framework on which this model rests, which was presented in the first part of this paper, has proved promising in the case of belief revision; it would not be exaggerated to expect similar success when applied to other questions where belief is involved and "realism" is an issue.

Appendix

PROOF OF THEOREM 1. Let $\preceq_{12}$ stand for $\preceq_1 \times_L \preceq_2$ and similarly for $\preceq_{13}$. Without risk of confusion, the image of the element ϕ of $\mathbf{B_i}$ in $\mathbf{B_1} * \mathbf{B_2}$ shall also be called ϕ.

(K ∗ 1) For any element X in the base algebra of an interpreted algebra $\mathbf{B}$, $|X|$ is closed under the consequence relation on $\mathbf{B}$, by Definitions I.1 and 2. $K * \phi$ is such a $|X|$ in $\mathbf{B_1} * \mathbf{B_2}$, so it is closed under the consequence relation for $\mathbf{B_1} * \mathbf{B_2}$.

(K ∗ 2) There are two cases:

$\perp_1 \not\simeq_B q_2(\phi)$. It follows from the hypothesis that $\phi \not\Leftrightarrow_{12} \perp_{12}$. By definition of the lexicographic order, $\{x | \; x \;\; \preceq_{12}\text{-min}\} \leqslant q(\phi)$, so ϕ is in $K * \phi$.

$\perp_1 \simeq_B q_2(\phi)$ and $\top_2 \Rightarrow_2 \phi$. $\mathbf{B_1} * \mathbf{B_2}$ is trivial, and so, by Definition 2, the centre $K * \phi$ contains ϕ.

(K ∗ 3) Let θ be a generator of $(\mathbf{B_1}, \preceq_1)$, and consider the images of θ and ϕ in $\mathbf{B_1} * \mathbf{B_2}$. $Cn_{12}(K \cup \{\phi\}) = \{\psi \in \mathbf{B_1} * \mathbf{B_2} | \; \theta \wedge \phi \Rightarrow \psi\} = |q_{12}(\theta) \cap q_{12}(\phi)|$. There are two cases:

$\theta \wedge \phi \not\Leftrightarrow_{12} \perp_{12}$. Since $q_{12}(\theta)$ (resp. $q_{12}(\phi)$) is the (non-empty) set of $\preceq_1$-minimal (resp. $\preceq_2$-minimal) worlds in $\mathbf{B_1} * \mathbf{B_2}$, $q_{12}(\theta) \cap q_{12}(\phi)$ is the set of $\preceq_{12}$-minimal worlds in $\mathbf{B_1} * \mathbf{B_2}$. So $K * \phi = |q_{12}(\theta) \cap q_{12}(\phi)| = Cn_{12}(K \cup \{\phi\})$;

$\theta \wedge \phi \Leftrightarrow_{12} \perp_{12}$. Therefore $Cn_{12}(K \cup \{\phi\}) = |\perp| \supseteq K * \phi$.

(K ∗ 4) Immediate consequence of the reasoning in the first case of axiom (K ∗ 3).

---

of information. The analysis proffered of Rott's counterexample, the simple definition of the revision operation ∗, and the solid motivations and interpretations given in Part I and Section 2 seem to suggest that the model proposed here satisfies all of these requirements.

(K $*$ 5) Since $\mathbf{B_1}$ is non trivial, it follows from the hypothesis that $\mathbf{B_1} * \mathbf{B_2}$ is non trivial. So there is a non empty set of $\preceq_{12}$-minimal worlds; since $q_{12}(\bot) = \bot$, the set of sentences true in these worlds, $K * \phi$, is consistent under Cn_{12}.

(K $*$ 6) By Definition 2, if ϕ is a generator of $(\mathbf{B_2}, \preceq_2)$, and $\phi \Leftrightarrow_2 \chi$, then χ is a generator of $(\mathbf{B_2}, \preceq_2)$. So both $K * \phi$ and $K * \chi$ are the centre of $(\mathbf{B_1} * \mathbf{B_2}, \preceq_{12})$, and hence they are equal (by Definition 2, there is a unique centre).

(K $*$ 7) There are two cases:

$\bot_1 \simeq_B q_2(\phi)$**.** If $\mathbf{B_1} * \mathbf{B_2}$ is trivial, then the condition trivially holds. If not, let θ be a generator of $(\mathbf{B_1} * \mathbf{B_2}, \preceq_{12})$. Since ϕ is not a consequence of θ, and since $\preceq_2$ and $\preceq_3$ coincide on $\neg\phi$ (Definition 6), θ is also a generator of $(\mathbf{B_1} * \mathbf{B_2}, \preceq_{13})$. Thus $K * (\phi \wedge \psi) = |q_{12}(\theta)| \subseteq |q_{12}(\theta) \cap q_{12}(\psi)| = Cn_{12}(K * \phi \cup \{\psi\})$.

$\bot_1 \not\simeq_B q_2(\phi)$**.** Let θ be a generator of $(\mathbf{B_1} * \mathbf{B_2}, \preceq_{12})$. There are two cases:

$\theta \wedge \psi \not\Leftrightarrow_{12} \bot_{12}$**.** $q_{12}(\theta) \cap q_{12}(\psi)$ is thus the (non empty) set of $\preccurlyeq_{12}$-minimal worlds in $\mathbf{B_1} * \mathbf{B_2}$ where ψ holds. However, by Definition 6, this is exactly the set of $\preceq_{13}$-minimal worlds. Therefore, $K * (\phi \wedge \psi) = |q_{12}(\theta) \cap q_{12}(\psi)| = Cn_{12}(K * \phi \cup \{\psi\})$;

$\theta \wedge \psi \Leftrightarrow_{12} \bot_{12}$**.** Therefore $Cn_{12}(K * \phi \cup \{\psi\}) = |\bot| \supseteq K * (\phi \wedge \psi)$.

(K $*$ 8) Immediate consequence of the reasoning in the first case of the second case in axiom (K $*$ 7). ■

Remark 3. The operation of extension — classically taken to be the addition of information to a belief set (possibly resulting in a contradictory belief set) — is involved in the Gardenförs postulates. As can be seen from the proof of (K $*$ 3) and (K $*$ 7), it can be defined by taking the fusion of the interpreted algebras (Definition I.5) and then the conjunction of images of the new information and the belief set.

PROPOSITION 1. *Let* $(\mathbf{B_1}, \preceq_1)$ *be an ordered algebra with centre* K*,* $(\mathbf{B_2}, \preceq_2)$ *an ordered algebra with generator* ϕ*, such that the identification relation* $\simeq$ *between* $\mathbf{B_1}$ *and* $\mathbf{B_2}$ *is* total *on* $\mathbf{B_2}$ *(to every element of* $\mathbf{B_2}$*, it associates an element of* $\mathbf{B_1}$*), and the consequence relation is preserved under this relation.*[24] *Let* $K * \phi$ *be the centre of* $(\mathbf{B_1}, \preceq_1) * (\mathbf{B_2}, \preceq_2)$ *(the result of revising according to* $*$*), and* $K *^e \phi = |\{x \in B_1 \mid x \preceq_1\text{-minimal st. } x \preccurlyeq q_1(\phi)\}|$

[24] For $q_1(\phi) \simeq_B q_2(\phi')$, $q_1(\psi) \simeq_B q_2(\psi')$, $\phi \Rightarrow_1 \psi$ iff $\phi' \Rightarrow_2 \psi'$.

(the result of revising by the revision operation induced by $\preceq_1$ *on* $\mathbf{B_1}$*). Then,* $K * \phi = K *^e \phi$*.*

PROOF OF PROPOSITION 1. By the condition on $\simeq$, $\mathbf{B_1} * \mathbf{B_2} = \mathbf{B_1}$. By the definition lexicographic product, the $\preceq_{12}$-minimal worlds are exactly the $\preceq_1$-minimal worlds where ϕ. These are just those worlds relevant for $K *^e \phi$. ■

PROOF OF THEOREM 2. Since all the operations used to define $*$, the fusion operator on ordered algebras, are associative, $*$ is associative. Let $(\mathbf{B}, \preceq)$, $(\mathbf{B_2}, \preceq_2)$, $(\mathbf{B_3}, \preceq_3)$ be ordered algebras with centre K and generators ϕ and ψ respectively. Let $\phi * \psi$ be a generator of $(\mathbf{B_2}, \preceq_2) * (\mathbf{B_3}, \preceq_3)$, and $K * -$ denote the centre of the appropriate fusion of algebras. The associativity of $*$ implies that[25]

$$\text{(Ass)} \quad (K * \phi) * \psi = K * (\phi * \psi)$$

Remark 4. It follows that all iterated revision operators which are supported by this $*$ (as defined with the lexicographic product) satisfy a formula of the form $(K * A) * B = K * \mathcal{F}(A, B)$, for some function taking pairs of sentences to sentences. This property is called "right-associativity" in [31, §8.3].

Both the radical and moderate iterated revision operators are characterised by conditions of this form. It remains to show that, when $(\mathbf{B_2}, \preceq_2)$ (resp. $(\mathbf{B_3}, \preceq_3)$) are point (resp. simple) algebras for ϕ and ψ, **(Rad)** (resp. **(Mod)**) is satisfied.

(Rad) By the definition of point algebras (Example 1):

$$(\mathbf{B}_{\phi_\mathbf{p}}, \preceq_{\phi_p}) * (\mathbf{B}_{\psi_\mathbf{p}}, \preceq_{\psi_p}) = \begin{cases} (\mathbf{B}_{\psi_\mathbf{p}}, \preceq_{\psi_\mathbf{p}}) & \text{if } q(\phi) \simeq_B q(\psi) \\ ((B_{\{\psi\}}, \mathbf{0}, q_0), \preceq_0) & \text{if } q(\phi) \simeq_B q(\neg\psi) \\ ((B_{\{\phi,\psi\}}, \mathbf{1}, q), \preceq) & \text{otherwise} \end{cases}$$

where $q : \phi \wedge \psi \mapsto \top$ and $\preceq$ is the only possible order. In all cases, $*$ yields an ordered algebra with generator $\phi \wedge \psi$; using this ordered algebra to revise by $\phi \wedge \psi$, (Ass) ensures that **(Rad)** is satisfied.

(Mod) By the definition of simple algebras (Example 1):

$$(\mathbf{B}_{\phi}, \preceq_{\phi}) * (\mathbf{B}_{\psi}, \preceq_{\psi}) = \begin{cases} (\mathbf{B}_{\psi}, \preceq_{\psi}) & \text{if } q(\phi) \simeq_B q(\psi) \\ (\mathbf{B}_{\psi}, \preceq_{\psi}) & \text{if } q(\phi) \simeq_B q(\neg\psi) \\ ((B_{\{\phi,\psi\}}, \mathbf{4}, q), \preceq) & \text{otherwise} \end{cases}$$

[25] As noted in Section I.2, the definition of the identification relation (Definition I.2) extends naturally to the case of more than two algebras. The canonical identification relations between the algebras are assumed in this result.

where **4** has four atoms, the images, under q, of $\phi \wedge \psi$, $\neg\phi \wedge \psi$, $\phi \wedge \neg\psi$ and $\neg\phi \wedge \neg\psi$, and the order imposed by $\preceq$ is that in which they are listed. $*$ yields an ordered algebra with generator $\phi \wedge \psi$ when ϕ and ψ are not contradictories, and ψ when they are. Using this ordered algebra to revise by $\phi \wedge \psi$, (Ass) ensures that **(Mod)** is satisfied. ■

Remark 5. Examination of the proof shows that, if ψ is not a contradictory element of **B**, then **(Rad)** and **(Mod)** are satisfied when the consistency in Definition 7 is understood relative to **B**.[26] This is another sense in which the two notions of iterated revision are captured by this representation of the new information.

Acknowledgements. Early versions of this paper were presented at Logica 2006 and the Studia Logica Trends in Logic IV conference. The author wishes to thank the respective audiences, as well as two anonymous referees, for their useful comments.

References

[1] Alchourron, Carlos E., Peter Gärdenfors, and David Makinson, 'The logic of theory change: Partial meet contraction and revision functions', *The Journal of Symbolic Logic* 50 (1985), 510–530.

[2] Baltag, Alexandru, and Lawrence S. Moss, 'Logic for epistemic programs', *Synthese* 60 (2004), 1–59.

[3] Baltag, Alexandru, Lawrence S. Moss, and Slawomir Solecki, 'The logic of common knowledge, public announcements and private suspicions', in I. Gilboa, (ed.), *Proceedings of the 7th Conference on Theoretical Aspects of Rationality and Knowledge (TARK'98)*, 1998, pp. 43–56.

[4] Baltag, Alexandru, and Sonja Smets, 'Dynamic belief revision over multi-agent plausibility models', in Giacomo Bonanno, Wiebe van der Hoek, and Michael Wooldridge, (eds.), *Proceedings of the 7th Conference on Logic and the Foundations of Game and Decision Theory (LOFT06)*, 2006, pp. 11–24.

[5] Baltag, Alexandru, and Sonja Smets, 'The logic of conditional doxastic actions: A theory of dynamic multi-agent belief revision', in S. Artemov, and R. Parikh, (eds.), *Proceedings of the Workshop on Rationality and Knowledge ESSLLI 2006*, 2006, pp. 13–30.

[6] Blackburn, Patrick, Maarten de Rijke, and Yde Venema, *Modal Logic*, Cambridge University Press, Cambridge, 2001.

[7] Darwiche, Adnan, and Judea Pearl, 'On the logic of iterated belief revision', *Artificial Intelligence* 89 (1997), 1–29.

---

[26] The restriction on ψ is required given the flexibility of the representation of new information. It is not required in the case of **(Rad)**, since point algebras are being used (see the remarks following Theorem 1).

[8] Fagin, Ronald, and Joseph Y. Halpern, 'Belief, awareness, and limited reasoning', *Artificial Intelligence* 34 (1988), 39–76.

[9] Friedman, Nir, and Joseph Y. Halpern, 'Belief revision: a critique', *Journal of Logic, Language, and Information* 8 (1999), 401–420.

[10] Gerbrandy, Jelle, and Willem Groeneveld, 'Reasoning about information change', *Journal of Logic, Language, and Information* 6 (1997), 147–169.

[11] Gärdenfors, Peter, *Knowledge in Flux : Modeling the Dynamics of Epistemic States*, MIT Press, Cambridge, MA, 1988.

[12] Gärdenfors, Peter, and Hans Rott, 'Belief revision', in Gabbay, Hogger, and Robinson, (eds.), *Handbook of Logic in AI and Logic Programming*, vol. 4, Oxford University Press, Oxford, 1995.

[13] Grove, Adam, 'Two modelings for theory change', *Journal of Philosophical Logic* 17 (1988), 157–170.

[14] Halpern, Joseph Y., and Leandro Chaves Rêgo, 'Interactive awareness revisited', in *Proceedings of Tenth Conference on Theoretical Aspects of Rationality and Knowledge*, 2005, pp. 78–91.

[15] Hansson, Sven Ove, 'Ten philosophical problems in belief revision', *Journal of Logic and Computution* 13 (2003), 37–49.

[16] Heifetz, A., M. Meier, and B. Schipper, 'Interactive unawareness', *Journal of Economic Theory* 130 (2006), 78–94.

[17] Heifetz, A., M. Meier, and B. Schipper, 'Awareness, beliefs and games', in *Theoretical Aspects of Rationality and Knowledge XI*, vol. XI, 2007.

[18] Henkin, Leon, J. Donald Monk, and Alfred Tarski, *Cylindrical Algebras. Part I*, North-Holland, 1985.

[19] Hill, Brian, *Jouer avec le Faux. Recherches sur les processus mentaux à l'œuvre dans la lecture des textes de fiction*, Doctorate thesis, University Paris 1 Panthéon-Sorbonne, 2006.

[20] Hill, Brian, 'Living without state-independence of utilities', Tech. rep., GREGHEC, 2007. 874/2007.

[21] Hill, Brian, 'The logic of awareness change', in *Proceedings of ILCLI International Worskop on Logic and Philosophy of Knowledge, Communication and Action*, University of the Basque Country Press, 2007.

[22] Hill, Brian, 'Logicality: from a local point of view', *Yeditepe'de Felsefe Yearbook*, 6 (2007).

[23] Hintikka, Jaakko, 'Impossible possible worlds vindicated', *Journal of Philosophical Logic*, 4 (1975), 475–484.

[24] Kahneman, Daniel, and Amos Tversky, 'Prospect theory: An analysis of decision under risk', *Econometrica* 47 (1979), 263–291.

[25] Konieczny, Sébastien, and Ramón Pino Pérez, 'A framework for iterated revision', *Journal of Applied Non-Classical Logics* 10 (2000), 3-4, 339–367.

[26] Koppelberg, S., 'General theory of boolean algebras', in J. D. Monk, and R. Bonnet, (eds.), *Handbook of Boolean Algebras*, vol. 1, North Holland, 1989.

[27] Lewis, David K., *Counterfactuals*, Harvard University Press, Cambridge, Massachusetts, 1973.

[28] Nayak, Abhaya C., 'Iterated belief change based on epistemic entrenchment', *Erkenntnis* 41 (1994), 353–390.

[29] Rabinowicz, Wlodzimierz, 'Stable revision, or is preservation worth preserving', in André Fuhrmann, and Hans Rott, (eds.), *Logic, Action and Information: Essays on Logic in Philosophy and Artificial Intelligence*, De Gruyter, Berlin, 1995, pp. 101–128.

[30] Rott, Hans, *Change, Choice and Inference*, vol. 42 of *Oxford Logic Guides*, Clarendon Press, Oxford, 2001.

[31] Rott, Hans, 'Coherence and conservatism in the dynamics of belief. part ii: Iterated belief change without dispositional coherence', *Journal of Logic and Computation* 13 (2003), 111–145.

[32] Rott, Hans, 'A counterexample to six fundamental principles of belief formation', *Synthese* 139 (2004), 225–240.

[33] Rott, Hans, 'A counterexample to six fundamental principles of belief formation', *Synthese* 139 (2004), 225–240. Retrieved February 20, 2006, from http://www.uni-regensburg.de/Fakultaeten/phil_Fak_I/Philosophie/theo_neu/RottV/Index_HRott.htm.

[34] Rott, Hans, 'Shifting priorities: Simple representations for twenty-seven iterated theory change operators', in *Modality Matters: Twenty-Five Essays in Honour of Krister Segerberg*, vol. 53, Uppsala Philosophical Studies, 2006, pp. 359–384.

[35] Savage, Leonard, *The Foundations of Statistics*, Dover, New York, 1954. 2nd edn 1971.

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

[37] Spohn, Wolfgang, 'Ordinal conditional functions', in William L Harper, and Brian Skyrms, (eds.), *Causation in Decision, Belief Change and Statistics*, vol. II, Reidel, Dordrecht, 1988, pp. 105–134.

[38] Stalnaker, Robert C., *Inquiry*, MIT Press, Cambridge, MA, 1984.

[39] Tversky, Amos, and Daniel Kahneman, 'Rational choice and the framing of decisions', *Journal of Business* 59 (1986), 251–278.

[40] van Benthem, Johan, 'Dynamic logic for belief change', *Journal of Applied Non-classical Logics* 17 (2007), 2007.

[41] van Ditmarsch, Hans P., 'Knowledge games', *Bulletin of Economic Research* 53 (2001), 249–273.

[42] van Ditmarsch, Hans P., 'Prolegomena to dynamic logic of belief revision', *Synthese* 147 (2005), 229–275.

Brian Hill
GREGHEC and IHPST
HEC Paris
1 rue de la Libération
78351 Jouy-en-Josas, France
hill@hec.fr