# Paraconsistent Belief Revision

by
GRAHAM PRIEST
University of Melbourne

The rational person apportions their beliefs according to the evidence. This is, no doubt, a fundamental feature of rationality. But there is more to rationality concerning belief than that. The rational person is also one who *changes* their beliefs in the light of new information in an appropriate way. Sometimes, this will merely require the addition of the new information to the belief-set. At other times, this may require more drastic change, since old beliefs may have to be dropped. Though the topic of how one ought, rationally, to change one's beliefs is a venerable one in the history of Western philosophy, the application of formal methods to it is relatively recent. Yet these applications suffice to throw into relief various fundamental problem concerning belief-revision. It is one of these, concerning the role of consistency in the process, which I will address here. I will explain the problem in a moment. Let me first set the scene by saying something about one of the most common formal representations of belief-revision, the AGM account.

## *1. Background*

In the AGM account, a set of beliefs is represented by a set of sentences, $K$, of some fixed language, $\mathscr{L}$. $K$ is closed under deducibility. Hence, it is perhaps better to look on it as representing the set of *commitments* of an agent, rather than as their set of beliefs, since a real agent's explicit beliefs are not, as a matter of fact, closed under deduction. Given some information, $\alpha$, expressed in $\mathscr{L}$, the AGM theory concerns three operations that may be effected on $K$ in virtue of this.

The first is adding $\alpha$ to $K$. The result, written as $K+\alpha$ (and usually called 'expansion'), is easily defined. We simply put $\alpha$ together with $K$, and then form their logical closure (since belief-sets are closed under deducibility). Thus, $K+\alpha=\{\beta\,;K\cup\{\alpha\}\vdash_{CL}\beta\}$ (where $\vdash_{CL}$, denotes classical logical consequence).

The second operation is the result of taking $\alpha$ away from $K$, written as $K-\alpha$ (and usually called 'contraction'). Of course, $\alpha$ may not be in $K$ at all, in which case, $K-\alpha$ is just $K$. But if it is, we cannot *simply* delete $\alpha$ from $K$. Since the result must be logically closed, we may have to delete other things as well. Depending on $K$, there may be no unique way of doing this. (Example: if $K$ is the logical closure of $\{p, p\rightarrow q\}$, then if we remove $q$ from $K$, one of these two will have to be removed. But removing either of them will do the trick.) Possibly, other considerations will determine which deletions ought to be made, but it is not at all clear that simply being rational *does* define the result uniquely: *prima facie*, at least, rational people may well disagree about what ought to be given up. At any rate, the AGM account does not give an explicit definition of contraction; instead, it gives a set of axioms that $K-\alpha$ must satisfy. The conditions are well known, and I will not reproduce them here.[1]

A third operation on $K$, the one we are really interested in, revision, is the general result of revising $K$, given new information, $\alpha$. This may be written as $K*\alpha$ . Since revision may involve rejecting things, the same sort of considerations that apply to contraction apply to revision. Thus, AGM give a set of conditions that $K*\alpha$ must satisfy. These are as follows:

K*1 $K*\alpha$ is logically closed
K*2 $\alpha\in K*\alpha$
K*3 $K*\alpha\subseteq K+\alpha$
K*4 $\neg\alpha\notin K\Rightarrow K+\alpha\subseteq K*\alpha$
K*5 $K*\alpha$ is inconsistent $\Rightarrow$ $\alpha$ is a logical contradiction
K*6 $\alpha$ and $\beta$ are logically equivalent $\Rightarrow K*\alpha=K*\beta$
K*7 $K*(\alpha\wedge\beta)=(K*\alpha)+\beta$
K*8 $\neg\beta\notin K*\alpha\Rightarrow(K*\alpha)+\beta\subseteq K*(\alpha\wedge\beta)$

K*2 is usually called the *success condition*: revising by $\alpha$ produces a belief-set that contains $\alpha$. K*5, which will be of particular importance in what follows, may be called the *consistency* condition. (Its converse is given by the success condition.)

A natural thought at this point is that revision may be defined in terms

[1] For details of these, and of all the features of the standard AGM theory mentioned in what follows, see Gärdenfors (1988).

of the other two operations: $K * \alpha = (K - \neg\alpha) + \alpha$ . This is often called the *Levi identity*. We will come back to it later, but at any rate, if revision is defined in this way, it can be shown that the AGM conditions for + and – entail those for $*$.

Note that one feature of the AGM account of revision is that $K * \alpha$ is a set of sentences of the same language as that with which we started. Thus, it provides no account of revision that involves *conceptual innovation*. This is a highly important form of revision, not perhaps a common form, but arguably the most profound. At any rate, AGM has nothing to tell us about this.

The AGM axioms themselves are rather abstract. There are many concrete models of them; these characterise revision (uniquely) in terms of other things. One of the nicest models is in terms of "spheres". A simple way of thinking of this is as follows. A set of beliefs, $K$, can be identified with the set of its models (in the model-theoretic sense). Let us write $K^+$ for the set of all interpretations that make $K$ true (its models). I will write $\alpha^+$ for $\{\alpha\}^+$. $K + \alpha$ is then simply the theory whose models are $K^+ \cap \alpha^+$. To define the other two operations, we suppose that $K$ comes furnished with a set of spheres: $K^+ = S_0 \subseteq S_1 \subseteq \ldots \subseteq S_n = I$ (the set of all interpretations of $\mathscr{L}$). To keep things simple, we suppose that there is a finite number of spheres, though this is inessential. Each sphere may be thought of as a "fallback" theory. Thus, if any $S_i$ is shown to be untenable, the agent's next choice is $S_{i+1}$. For any non-empty set, $X \subseteq I$, there must be a smallest sphere, $S_X$, that intersects it (and if $X$ is empty, let $S_X = I$). $K * \alpha$ may be defined as the theory whose models are $S_{\alpha^+} \cap \alpha^+$. That is, it is the largest portion of the most preferred fallback theory that entails $\alpha$. The definition of contraction need not concern here.[2] Let us just note that it can be shown that the operations, defined in terms of spheres, satisfy the AGM conditions.

## 2. *The Problem*

### 2.1 Inconsistent Belief

Now to the problem posed by inconsistency. In the standard AGM account, as we saw, it is assumed that the logical consequence relation

[2] For the record again, if we write $\alpha^-$, for the set of all interpretations that do not make $\alpha$ true $K-\alpha$ is defined as the theory whose models are $K^+ \cup (S_{\alpha^-} \cap \alpha^-)$.

employed is classical. In particular, then, an inconsistent belief-set is trivial. Thus, adding to a belief-set something that is inconsistent with it produces triviality. Moreover, revising a belief-set with something inconsistent also gives triviality, because of the success condition.

Now, here is the problem. People often have inconsistent commitments. The persons whose beliefs are consistent is, in fact, a rarity. Yet it is absurd to suppose that a person who has inconsistent commitments is thereby committed to everything. If, by oversight, I believe both that I will give a talk on campus at noon, but also that I will be in town at noon, this hardly commits me to believing that the Battle of Hastings was in 1939.

It might be suggested that the theory of belief revision is one of an ideally rational agent, and that such an agent never has inconsistent beliefs. But this is a confusion. The theory of belief revision is a theory of how an ideally rational agent *changes* their beliefs. It is quite possible that such an agent should find themself with inconsistent beliefs in the first place (maybe through their education). Indeed, one thing we should expect of a theory of belief revision is an account of what it is rational to do if we do find ourselves in this situation.

Worse, it is not at all clear that an ideally rational agent must have consistent beliefs. Sometimes there is overwhelming evidence for inconsistent beliefs. The evidence points mercilessly to the fact that either *a* committed the crime, or that *b* did. But I have known both *a* and *b* for years. Both are of impeccable moral characters; neither is the kind of person who would do such a thing. This sort of situation seems to be a not uncommon one in science. For example, in the late 19th century, the evidence that evolution had occurred, and that the age of the earth was hundreds of millions of years, was overwhelming. But the thermodynamic evidence concerning the age of the sun, showed that the earth could not possibly be that old. Nor is this simply a matter of history: it is well known that presently the general theory of relativity and quantum theory are mutually inconsistent. Indeed, as many philosophers of science have noted, most accepted scientific theories face contradictions and anomalies[3]. For this reason, we cannot simply *suspend* belief in theories facing inconsistency. If we did, we would have no science left.

---

[3] E.g., Lakatos (1970), Feyerabend (1975), ch. 5.

Further, there are, in fact, good grounds for supposing that an ideally rational agent *must* have inconsistent beliefs. Such an agent would not believe something unless the evidence supported its truth. Hence, every one of their beliefs, $\alpha_1, \ldots, \alpha_n$, is rationally grounded. But the rational agent also knows that no one is perfect, and that the evidence is overwhelming that everyone has false beliefs (rational agents included: rationality does not entail infallibility). Hence, they believe $\neg(\alpha_1 \wedge \ldots \wedge \alpha_n)$. So their beliefs are inconsistent.[4] And even if, for some reason, this argument fails, one would hardly want to rule out *a priori* the *possibility* that rational belief is fallible in this way.

Of course, a rational agent can hardly believe everything. Hence, whether one takes the agent of the theory to be simply a rational reviser, or to be a rational believer as well, a theory of belief revision must allow for the possibility of agents having inconsistent but non-trivial beliefs.

### 2.2 Paraconsistency

Possibly, there are a number of ways that one might try to solve this problem, but the most obvious is simply to employ a relation of logical consequence that is paraconsistent, that is, in which inconsistencies do not entail everything. An agent's commitments may then be inconsistent without being trivial. There are many paraconsistent logics, and we do not need to go into details here.[5] Using such a logic, pretty much the whole of the AGM theory goes over intact. The sphere modeling, too, works the same way. We just replace classical models with the models of the paraconsistent logic employed. There is only one major and inevitable casualty of the AGM postulates: the consistency postulate.[6] This must fail: $K * \alpha$ may be inconsistent, even if $\alpha$ is quite consistent, and for reasons that have nothing to do with $\alpha$. But this is entirely what one should expect in the current context.[7]

---

[4] This is a version of the "preface paradox". For references and further arguments to the effect that if may be rational to have inconsistent beliefs, see Priest (1987), 7.4.

[5] These can be found in, e.g., Priest *et al.* (1989).

[6] There are a number of somewhat sensitive issues that I am sliding over, but nothing that affects what I have to say here. Details can be found in Tanaka (1996).

[7] Since the standard AGM account employs classical logic, there is no difference between inconsistency and triviality. Hence, the consistency condition is sometimes expressed with 'trivial' replacing 'inconsistent'. Both the consistency condition and its converse may then fail. For the converse: revising by a logical inconsistency will certainly produce an inconsist-

But now we face another problem. If we are allowing for the possibility of inconsistent beliefs, why should revising our beliefs with new information ever cause us to reject anything from our belief-set at all? Why not simply add the belief to our belief-set, and leave it at that?

The problem comes out sharply in the sphere modeling of the AGM postulates. In many standard paraconsistent logics, there is a trivial interpretation, $\infty$, one that makes everything true. *A fortiori*, it is a model of every theory. Given such an interpretation, for any $\alpha$ and $K$, $\infty \in K^{+} \cap \alpha^{+}$. Thus, $S_{\alpha^{+}}$ is just $K^{+}$ itself, and so $S_{\alpha^{+}} \cap \alpha^{+} = K^{+} \cap \alpha^{+}$. Revision and expansion are exactly the same thing![8] $\infty$ is just a technical way of saying that when you can believe any contradiction, revision requires nothing but addition. Note, however, that in a paraconsistent context there is no reason to suppose the Levi identity holds (unless, of course, one uses it to *define* revision). For example, let $K$ be the set of logical consequences of $\neg p$. Then $\neg p$ is not in $(K - \neg p) + p$, but it is in $K + p$.

## *3. A Solution*

### 3.1 Multiple Criteria

But this too fast. Suppose that we use a paraconsistent logic. If we revise our beliefs in the light of new information, $\alpha$, there is nothing now in logic that will force us to delete $\neg\alpha$ (and some of the things that entail it) from our beliefs. But just because this is a *logical* possibility, it does not follows that it is a *rational* possibility. There is a lot more to rationality than consistency. Many quite consistent beliefs are irrational; for example, the belief that I am a fried egg. I may even hold this belief consistently with the rest of me beliefs, if I make suitable adjustments elsewhere (by jettisoning, e.g., the belief that I was born, and not laid). This makes it no more rational.

As epistemologists have often noted, there are many features of a set of beliefs that may speak against its rational acceptability. Inconsistency

---

ent set – assuming success – but this will not, in general, be trivial. The possible failure in the opposite direction is not so obvious, but follows from a result that I will mention in a moment.

[8] This shows that the consistency condition fails in this model when 'inconsistent' is replaced by 'trivial'. If $K$ is the trivial set, then, whatever $\alpha$ is, $K * \alpha = K + \alpha = K$, and so is trivial.

may well do this, but so, for example, does a high level of *ad-hoc*ness – which is presumably what goes wrong in the case of the person who gerrymanders their beliefs into a consistent set containing one to the effect that they are a fried egg. Conversely, there are many features of a set of beliefs that speak in favour of its rational acceptability. Simplicity is a traditional such virtue; so are: a low degree of *ad-hoc*ness, fruitfulness, explanatory power, unifying power.[9] What, exactly, all these criteria – except consistency – amount to, is a thorny issue: they are all notoriously difficult to spell out. It may even be the case that some of them, when properly understood, come to the same thing, but certainly not all of them. Moreover, how one is to justify these epistemic virtues is a very difficult question.[10] But it is also a different one: as with moral virtues, people can agree that something, say kindness, is a virtue, whilst disagreeing in their theoretical accounts of why this is so. Indeed, that kindness is a virtue is a datum much firmer than any moral theory will ever be.

We need not go into any of these questions further here, though. For it is simply the multiplicity of the criteria that is presently relevant. As long as there is a multiplicity, there is the possibility of conflict between them.[11] One set of beliefs may be the simplest, have the highest explanatory power, but be inconsistent. Another may be consistent and fruitful, but highly *ad hoc*. What is it rational to believe in such circumstances? There may be no determinate answer to this question. But it would seem clear that (vaguely) if one set of beliefs scores sufficiently better on a sufficient number of these criteria than another, it is rationally preferable. This is how an inconsistent set of beliefs can be rationally acceptable: it scores highly on many of the other criteria. Conversely, and to return to the problem at hand, this is why an inconsistent set of beliefs may not be rationally acceptable. Its inconsistency may speak against it; and so may many other criteria. In particular, if we always revise simply by adding on the new information, we are likely to lose in simplicity, economy, unity. So the result may be quite irrational.

Looking at things in this way, consistency is no longer a necessary condition for rational belief, merely one of a list of (potentially conflicting)

---

9 See, e.g., Quine and Ullian (1970), ch. 5, Kuhn (1977), Lycan (1988), ch. 7.

10 For one account, see Lycan (1988), ch. 7.

11 The point is made in Kuhn (1977) and Lycan (1988), p. 130.

desiderata. But we conceded this once we conceded the possibility that it might be rational to have inconsistent beliefs anyway. The important point is that simply tacking on new information to our beliefs, rendering the whole belief-set inconsistent, though it may be a logical option, may not be the rational course of action.[12]

### 3.2 Formal Models

Let me now outline a formal model of belief revision incorporating this insight. Suppose, as before, that our belief-set is $K$, and that new information, $\alpha$, arrives. What is the new rational set of beliefs? There are a number of possibilities. One is simply the addition, $K+\alpha$. Another is something obtained by the Levi operation: $(K-\neg\alpha) + \alpha$, where $-$ is some suitable notion of contraction.[13] Another is a theory obtained by reversing the operations involved: $(K+\alpha) - \neg\alpha$. (There is no reason to suppose that, in general, this gives the same result. For example, it may not satisfy the AGM success condition.) Recall that the AGM conditions do not specify contraction uniquely; there may therefore even be many theories of the forms specified by the Levi identity and its reversal. And there may well be other possibilities as well. For example, one rational course of action may simply be to reject the new information, and write if off on the grounds of experimental error, or whatever (in which case, again, the success condition will clearly fail).[14] More profoundly, the new information may occasion conceptual innovation, and thus there may be sets of beliefs in a new language. Let us call the collection of all options, whatever they are, $K^{\alpha}$.

We know that there is a set of criteria, $C$, which can be used to evaluate sets of belief. These criteria are not all-or-nothing matters. (Given a paraconsistent logic, even inconsistency comes by degrees.) So let us suppose that for each $c \in C=\{c_1,\ldots,c_n\}$ there is a scale of how well a belief-set fares according to that criterion. Specifically, for each belief-set, $k \in K^{\alpha}$,

---

[12] The matter is discussed further in Priest (1987), ch. 7, and Priest (1998).

[13] Fuhrmann (1991) defines such a notion in terms of base contraction (i.e., revising a theory via revising its axioms). This approach does not obviate the need for the use of a paraconsistent logic, as Fuhrmann notes. Note, also, that the existence of an appropriate $(K-\neg\alpha) + \alpha$ does nothing on its own to solve the problem of why one should not revise simply by expanding.

[14] Further on this, see Hansson (1997).

there is a real number, $\mu_c(k)$ measuring how good that set is. (The higher the number, the better the set.) One might prefer some other scale, perhaps some subset of the real numbers, such as [0,1]. But since the whole matter is conventional, this is not a crucial issue.

Next, we need a way of amalgamating the various criteria. A simple way of doing this is by taking their weighted average.[15] Let $\rho(k)$, the "rationality index" of $K$, be defined thus:

$$\rho(k)=w_1\mu_{c_1}(k)+\ldots+w_n\mu_{c_n}(k)$$

The weights, $w_i$, reflect the relative importance of each of the criteria. I will have a little more to say about these later. $K*\alpha$ can now be defined as the member of $K^{\alpha}$ with the largest rationality index.[16] If there is more than one, $K*\alpha$ is one of these, non-deterministically. (Believers have a free choice.)

The model I have just described is simple, but obviously unrealistic in a number of ways. One of these is that, given a criterion, $c$, the assignment of a unique real number, $\mu_c(k)$, to a theory, $k$, is not really to be expected. About the best we can hope for is to assign $k$ a range of numbers. (And probably a vague range, at that. But how to handle vagueness poses a whole new set of problems that would take us far away from the present problem.)

A second feature of the model is that, according to it, all theories are comparable as to the rationality of their acceptability. The ordering is a linear one.[17] This seems quite unrealistic. Rational people may well disagree about the best thing to believe when different criteria rank theories radically differently.[18] Thus, suppose that there are two criteria, $c_1$ and $c_2$; $c_1$ ranks $k_1$ high, but $k_2$ low, and $c_2$ vice versa. Then there may just be

---

[15] A similar proposal is mooted by Levi (1967), p. 106.

[16] In certain traditions, it is common to talk about the coherence of a set of beliefs. It is usually assumed that consistency is a necessary condition for coherence, but not a sufficient one. The issue of what *are* sufficient conditions, though, is a tough one, and answers are hard to find. One way of understanding the notion is simply to define the degree of coherence of a theory to be its rationality index. If one does this then consistency is, of course, no longer a necessary condition for (a high degree of) coherence

[17] In many ways, it might be more natural to take the range of each $\mu_c$ to be some partial (non-numerical) ordering. One then has to face the question of how to amalgamate all these orderings into a single partial ordering with the appropriate properties. There may well be ways of doing this, but I do not presently know of any.

[18] For an excellent discussion, see Kuhn (1977).

no comparing the two theories. The best we can hope for is a determinate answer when most of the criteria give similar results (and that's vague too). Thus, the ordering *ought* to be a partial one.

By a small modification of the construction, we can solve both of these problems at once. We now take each $\mu_c$ to assign each $k$, not a single value, but a non-empty range of values, $[\mu_c^-(k), \mu_c^+(k)]$. (The first figure is the lower bound; the second figure is the upper bound.) When we amalgamate these, the result is also going to be a range of values. The highest value a belief-set can hope for is clearly realised when it obtains the highest value under each component, and similarly for the lowest. So we can take $\rho(k)$ to be the range $[\rho^-(k), \rho^+(k)]$, where:

$$\rho^-(k) = \sum_{1\leq i\leq n} w_i \mu_{c_i}^-(k)$$

$$\rho^+(k) = \sum_{1\leq i\leq n} w_i \mu_{c_i}^+(k)$$

An over-all ranking, $\sqsubset$, may now be defined on belief-sets. One belief-set, $k_1$, is rationally preferable to another, $k_2$, if it is clearly better, that is, if any value that $k_1$ can have is better than any value $k_2$ can have:

$$k_1 \sqsupset k_2 \Leftrightarrow \rho^-(k_1) > \rho^+(k_2)$$

It is easy to check that this is a partial ordering. It is not, in general, a total ordering. For example, if $\rho(k_1) = [3,7]$ and $\rho(k_2) = [2,4]$, then neither $k_1 \sqsupset k_2$ nor $k_2 \sqsupset k_1$. Thus, it may be that the rationality ordering, $\sqsupset$, provides a clear judgment concerning two theories sometimes, but not others. Rational disagreement is possible.

We can now define $K*\alpha$ to be any belief-set in $K^\alpha$ maximal in the ordering. If there is more than one, this is non-deterministic. For it is easy to see that there may be more than one distinct maximum. For example, suppose that $\rho(k_1) = [3,7]$ and $\rho(k_2) = [4,6]$, but for every other $k \in K^\alpha$, $\rho(k) = [1,2]$. Then $k_1$ and $k_2$ are rationally preferable to all the other $k$s, except each other, and neither is preferable to the other.

Let me illustrate this definition of revision with a toy example. Suppose that $K$ is the set of logical consequences of $\{p,q\}$ and that $\alpha$ is $\neg p$. We suppose that the only members of $K^\alpha$ are $k_1 = K+\alpha$, the set of logical consequences of $\{p,q,\neg p\}$, and the Levi revision $k_2 = (K-p) + \neg p$, the set of logical consequences of $\{q,\neg p\}$. Suppose, now, that there are two criteria, consistency, $c_1$, and explanatory power, $c_2$, and that the values of $\mu$ are (for the sake of illustration) as follows:

$$\mu_{c_1}(k_1) = [2,2] \qquad \mu_{c_1}(k_2) = [8,8]$$
$$\mu_{c_2}(k_1) = [4,5] \qquad \mu_{c_2}(k_2) = [3,4]$$

There is no room for disagreement about which theory is the more consistent! Finally, suppose that $w_1$=1, and $w_2$=2. Thus, we think that explanatory power is more important than consistency. Computing, we determine that:

$$\rho(k_1) = [10,12] \qquad \rho(k_2) = [14,16]$$

Thus, $k_1 \sqsubset k_2$, and $K * \alpha = k_2$. In particular, note, revision has occasioned the dropping of a belief.

This more complex model is clearly more realistic, but it is still a very idealised model. For a start, is there any reason to suppose that the criterion-weights, $w_i$, are uniquely and precisely determined? None that I can see. Moreover, it might well be the case that there could be rational disagreement about what their values ought to be.[19] A solution here is to treat weights as we treated the single-criterion values. That is, instead of taking them to be single values, we take them to be a range of values, broad enough to encompass differences of opinion. When we amalgamate, $\rho^+(k)$ is computed using the maxima of the weight-values; $\rho^-(k)$ is computed using the minima.

The account, thus modified, is still just a model. If it were ever to be applied in practice, the question of how to determine the values of the $w$s and the $\mu$ -values of each theory would become an important one. Of course, the *absolute* values are unimportant: the scales are almost completely arbitrary. It is the *relative* values (within scales and across scales) that are important. It can be hoped that a community of investigators could reach consensus on appropriate figures, particularly since it is only a *range* of values that has to be agreed upon. We may be able to set limits that keep everybody happy.

### 3.3 The AGM Conditions Revisited

Since we started with a discussion of the AGM axioms, let me finish by briefly discussing to what extent $K * \alpha$, as defined here, satisfies them. Take them one by one.

---

[19] This is, again, well illustrated in Kuhn (1977). More radically, there might even be grounds for revising the weights themselves in the light of scientific development.

K*1: $K * \alpha$ is logically closed.

$K * \alpha$ is a theory, and so logically closed; but under which logic? As argued, we may expect it to be a paraconsistent logic. But note that the construction does not require the logic to be any particular logic. Indeed, the different members of $K^{\alpha}$ may well be closed under logics that are different from each other, and from that of $K$ too. For one thing that we may wish to revise under the influence of recalcitrant information is exactly our logical theory. Thus, e.g., under the weight of counterexamples to the material conditional, it might be (indeed it is!) rational to move to a relevant logic. Quite unlike standard AGM constructions, then, this account shows us how, and in what way, our logic may be revised too. The rational choice, $K * \alpha$, simply has a different logic from $K$.

K*2: $\alpha \in K * \alpha$ .

The success postulate need not be satisfied by $K * \alpha$, as has already been discussed.

K*3: $K * \alpha \subseteq K + \alpha$ .

This postulate may not be satisfied either. For $\alpha$ may occasion conceptual revision. And if it does, there may be beliefs in $K * \alpha$ that cannot even be expressed in the language of $K$, and so are not in $K+\alpha$ at all. For this reason, the revision of phlogiston chemistry under the recalcitrant results obtained by weighing, occasioned beliefs employing quite new concepts, like *oxygen*.

K*4: $\neg\alpha \notin K \Rightarrow K + \alpha \subseteq K * \alpha$ .

For similar reasons, this postulate fails. Old beliefs, such as that there is a substance called phlogiston, may get ditched in the process of revision.

K*5: $K * \alpha$ is inconsistent $\Rightarrow$ $\alpha$ is a logical contradiction.

The failure of this postulate needs no further discussion.

K*6: $\alpha$ and $\beta$ are logically equivalent $\Rightarrow K * \alpha = K * \beta$ .

There is no reason to suppose that this postulate holds either. Nor should it. For $\alpha$ and $\beta$ may be logically equivalent under the logic of $K$, but they may occasion different revisions of logic. For example, the existence of a counterexample to the law of excluded middle is equivalent to the existence of one to distribution (of $\wedge$ over $\vee$) in classical logic, since these are both tautologies. Yet the first may occasion a rational move to a three-valued logic (in which distribution holds), whilst the second might occasion a rational move to a quantum logic (in which the law of excluded middle holds).

K*7: $K*(\alpha \wedge \beta) = (K * \alpha) + \beta$ .

This postulate fails too, as it should. Suppose that you revise given the (*prima facie*) information that $\alpha \wedge \neg\alpha$ – maybe it is a new logical paradox. Perhaps the rational thing to do is to move to a belief-set that contains neither $\alpha$ nor $\neg\alpha$ (since $\alpha$ is neither true nor false). Certainly, then $\neg\alpha \notin K*(\alpha \wedge \neg\alpha)$. But $\neg\alpha \in (K*\alpha) + \neg\alpha$.

K*8: $\neg\beta \notin K*\alpha \Rightarrow (K*\alpha) + \beta \subseteq K*(\alpha \wedge \beta)$.

This postulate fails for similar reasons. Suppose that $\beta$ is $\neg\alpha$ ; then as for K*7.

In summary, then, given the construction of this section, all the AGM postulates may fail. They do so because the model shows how to operate in a much more general class of revisions than the simple ones envisaged by AGM. This, indeed, is one of its strengths.[20]

## *4. Conclusion*

The model of belief revision give here is, as we have seen, a much more general and flexible one than AGM. I do not think that, in real life, disputes over what belief-set (theory) it is rational to adopt proceed explicitly in the way prescribed by the model. Rational belief change is a much less cleanly articulated business than this. But I do think that the model captures, at least roughly, the qualitative features of what goes on implic-

---

[20] There is another solution to the main problem of this paper that is worth noting here. In (1997), ch. 4, and (1999) Fuhrmann gives ways of defining an operation which merges two sets of information, $A$ and $B$, to form a consistent set A ∘ B. (A ∘ B is offered as the way to revise $A$ given a whole set of new information, $B$.) Fuhrmann observes that if we have premise sets $A$ and $B$, which may be inconsistent, and define $A,B \vdash \alpha$ as $A \circ B \vdash_{CL}$ – or, alternatively, define $A \vdash \alpha$ as $A \circ A \vdash_{CL} \alpha$ – then $\vdash$ is a paraconsistent notion of inference. Given that it is paraconsistent, it could well be the notion of inference employed in the construction of this paper. But there is a way of applying the construction more directly (though not one mooted by Fuhrmann himself). Suppose that *whenever* we acquire new information, $\alpha$, we simply add it; but we allow the consequences to be determined by $\vdash$. Thus, the inference engine itself determines the resolution of contradictions. This construction is less satisfactory than the one described here. The question to ask is: which set it is that represents our beliefs (comittments)? Is it our total information, or is it its $\vdash$-consequences. If it is the first, then we never drop any beliefs at all, which certainly seems wrong. If it is the second, then our set of beliefs would never be inconsistent; as argued, this is wrong too. Another drawback of the suggestion is that it is much less general than the account given here: it allows for neither conceptual innovation nor for revision of the underlying logic of a belief set.

itly in such contexts. And it suffices to give the lie to an objection often raised against those who think that it may be rational to believe a contradiction.[21] This is to the effect that there can be no rational debate with a person who will accept contradictions, since they may accept anything. They may do nothing of the kind.[22]

## *References*

[1] P. FEYERABEND (1975), *Against Method*, London: New Left Books.

[2] A. FUHRMANN (1991), 'Theory Contraction through Base Contraction', *Journal of Philosophical Logic* 20, 175–203.

[3] A. FUHRMANN (1997), *An Essay on Contraction*, Stanford, CA: CSLI Publications.

[4] A. FUHRMANN (1999), 'When Hyperpropositions Meet', *Journal of Philosophical Logic* 28, 559–74.

[5] P. GÄRDENFORS (1988), *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, Boston, MA: MIT Press.

[6] S. O. HANSSON (1997), 'Semi-Revision', *Journal of Applied and Non-Classical Logics* 7, 151–75.

[7] T. S. KUHN (1977), 'Objectivity, Value Judgment and Theory Choice', ch. 13 of *The Essential Tension*, Chicago, IL: University of Chicago Press.

[8] I. LAKATOS (1970), 'Falsification and the Methodology of Scientific Research Programmes', in I. Lakatos and A. Musgrave (eds.), *Criticism and the Growth of Knowledge*, Cambridge: Cambridge University Press.

[9] I. LEVI (1967), *Gambling with the Truth*, London: Routledge and Kegan Paul.

[10] D. LEWIS (1982), 'Logic for Equivocators', *Noûs* 14, 431–41.

[11] W. G. LYCAN (1988), *Judgment and Justification*, Cambridge: Cambridge University Press.

[12] K. POPPER (1963), *Conjectures and Refutations*, London: Routledge and Kegan Paul.

[13] G. PRIEST (1987), *In Contradiction*, Dordrecht: Kluwer Academic Publishers.

[14] G. PRIEST (1998), 'What's so Bad about Contradictions?', *Journal of Philosophy* 95, 410–26.

---

21 Versions of it can be found in Lewis (1982), p. 434, and Popper (1963), p. 316f.

22 I am grateful for comments on earlier drafts of this essay to an anonymous referee, Koji Tanaka and, especially, Hans Rott.