ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/260134902

Belief Revision: an introduction

Chapter · May 1992

DOI: 10.1017/CBO9780511526664.001

CITATIONS
183

READS
821

1 author:

Peter Gärdenfors
Lund University

278 PUBLICATIONS **13,734** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Conceptual Spaces View project

Project The evolution of teaching View project

All content following this page was uploaded by Peter Gärdenfors on 08 January 2016.

The user has requested enhancement of the downloaded file.

Belief Revision: An Introduction

PETER GÄRDENFORS

Cognitive Science, Department of Philosophy,
Lund University, S-223 50 Lund, Sweden

1 THE PROBLEMS OF BELIEF REVISION

1.1 An Example

Suppose that you have a database that contains, among other things, the following pieces of information (in some form of code):

α: All European swans are white.
β: The bird caught in the trap is a swan.
γ: The bird caught in the trap comes from Sweden.
δ: Sweden is part of Europe.

If your database is coupled with a program that can compute logical inferences in the given code, the following fact is derivable from α - δ:

ε: The bird caught in the trap is white.

Now suppose that, *as a matter of fact*, the bird caught in the trap turns out to be black. This means that you want to add the fact $\neg\varepsilon$, i.e., the negation of ε, to the database. But then the database becomes *inconsistent*. If you want to keep the database consistent, which is normally a sound methodology, you need to *revise* it. This means that some of the beliefs in the original database must be retracted. You don't want to give up all of the beliefs since this would be an unnecessary loss of valuable information. So you have to *choose* between retracting α, β, γ or δ.

The problem of belief revision is that logical considerations alone do not tell you which beliefs to give up, but this has to be decided by some other means. What makes things more complicated is that beliefs in a database have *logical consequences*, so when giving up a belief you have to decide as well which of the consequences to retain and which to retract. For example, if you decide to retract α in the situation described here, α has as logical consequences, among others, the following two:

α': All European swans except the one caught in the trap are white

and

α'': All European swans except some of the Swedish are white.

Do you want to keep any of these sentences in the revised database?

1.2 The Methodological Problems of Belief Revisions

When trying to handle belief revisions in a computational setting, there are three main methodological questions to settle:

(1) How are the beliefs in the database *represented*?

Most databases work with elements like *facts* and *rules* as primitive forms of representing information. The code used to represent the beliefs may be more or less closely related to standard logical formalism. A mechanism for belief revision is sensitive to the formalism chosen to represent the beliefs.

(2) What is the relation between the elements explicitly represented in the database and the beliefs that may be *derived* from these elements?

This relation is to a large extent dependent on the *application area* of the database. In some cases the elements explicitly formulated in the database have a special status in comparison to the logical consequences of these beliefs that may be derived by some inference mechanism. In other cases, the formulation of the beliefs in the database is immaterial so that any representation that has the same logical consequences, i.e., the same set of implicit beliefs, is equivalent. As will be seen in several papers in this volume, the nature of the relation between explicit and implicit beliefs is of crucial importance for how the belief revision process is attacked.

(3) How are the choices concerning what to retract made?

Logic alone is not sufficient to decide between which beliefs to give up and which to retain when performing a belief revision. What are the extralogical factors that determine the choices? One idea is that the information lost when giving up beliefs should be kept minimal. Another idea is that some beliefs are considered more important or entrenched than others and the beliefs that should be retracted are the least important ones. Within computer science the use of *integrity constraints* is a common way of handling the problem. Again, the methodological rules chosen here are dependent on the application area.

1.3 Three Kinds of Belief Changes

A belief revision occurs when a new piece of information that is *inconsistent* with the present belief system (or database) is added to that system in such a way that the result is a new consistent belief system. But this is not the only kind of change that can occur in a belief system. Depending on how beliefs are represented and what kinds of inputs are accepted, different typologies of belief changes are possible.

In the most common case, when beliefs are represented by *sentences* in some code, and when a belief is either *accepted* or *rejected* in a belief system K (so that no degrees of belief are considered), one can distinguish three main kinds of belief changes:

(i) *Expansion*: A new sentence ϕ is added to a belief system K together with the logical consequences of the addition (regardless of whether the larger set so formed is consistent). The belief system that results from expanding K by a sentence ϕ will be denoted $K+\phi$.

(ii) *Revision*: A new sentence that is inconsistent with a belief system K is added, but, in order to maintain consistency in the resulting belief system, some of the old sentences in K are deleted. The result of revising K by a sentence ϕ will be denoted $K\dotplus\phi$.

(iii) *Contraction*: Some sentence in K is retracted without adding any new facts. In order for the resulting system to be closed under logical consequences some other sentences from K must be given up. The result of contracting K with respect to ϕ will be denoted $K\dot{-}\phi$.

Expansions of belief systems can be handled comparatively easily. $K+\phi$ can simply be defined as the logical closure of K together with ϕ:

(Def +) $\qquad K+\phi = \{\psi: K \cup \{\phi\} \vdash \psi\}$

As is easily shown, $K+\phi$ defined in this way will be closed under logical consequences and will be consistent when ϕ is consistent with K.

It is not possible to give a similar explicit definition of revisions and contractions in logical and set-theoretical notions only. The problems for revisions were presented in the introductory example. There is no purely logical reason for making one choice rather than the other among the sentences to be retracted, but we have to rely on additional information about these sentences. Thus, from a logical point of view, there are several ways of specifying the revision $K\dotplus\phi$. Though $K\dotplus\phi$ cannot be characterized uniquely in logical terms, the *general properties* of a revision function can be investigated, and – in some cases, at least – *algorithms* can be found for computing revision functions. These two goals will be handled technically by using the notion of a *revision function* "$\dotplus$" which has two arguments, a belief system K and a sentence ϕ, and which has as its value the revised belief system $K\dotplus\phi$.

The contraction process faces parallel problems. To give a simple example, consider a belief system K which contains the sentences ϕ, ψ, $\phi \wedge \psi \rightarrow \chi$ and their logical consequences (among which is χ). Suppose that we want to contract K by deleting χ. Of course, χ must be deleted from K when forming $K \dot{-} \chi$, but also at least one of the sentences ϕ, ψ, or $\phi \wedge \psi \rightarrow \chi$ must be given up in order to maintain consistency. Again, there is no purely logical reason for making one choice rather than the other. Another concrete example is provided by Fagin, Ullman and Vardi (1983, p. 353).

The common denominator in both this example and the introductory one is that the database is not viewed merely as a collection of logically independent facts, but rather as a collection of axioms from which other facts can be derived. It is the interaction between the updated facts and the derived facts that is the source of the problem.

In parallel with revision we can introduce the concept of a *contraction function* "$\dot{-}$" which has the same two arguments as before, i.e., a belief system K and a sentence ϕ (to be retracted from K), and which produces as its value the belief system $K \dot{-} \phi$. In Section 3.3, I shall show that the problems of revision and contraction are closely related – being two sides of the same coin.

1.4 Two Approaches to Describing Belief Revisions

When tackling the problem of belief revision there are two general strategies to follow, namely, to present explicit *constructions* of the revision process and to formulate *postulates* for such constructions. For a computer scientist the ultimate solution to the problem about belief revision is to develop *algorithms* for computing appropriate revision and contraction functions for an arbitrary belief system. In this volume several proposals for constructions of revision methods will be presented. These methods are not presented as pure algorithms, but on a slightly more general level.

However, in order to know whether an algorithm is successful or not it is necessary to determine what an 'appropriate' revision function is. Our standards for revision and contraction functions will be various *rationality postulates*. The formulations of these postulates are given in a more or less equational form. One guiding idea is that the revision $K \dot{+} \phi$ of K with respect to ϕ should represent the minimal change of K needed to accommodate ϕ consistently. The consequences of the postulates will also be investigated.

Much of the theoretical work within belief revision theory consists of connecting the two approaches. This is done via a number of *representation theorems*, which show that the revision methods that satisfy a particular set of rationality postulates are exactly those that fall within some computationally well defined class of methods.[1]

[1] For further discussion of the two strategies cf. Makinson (1985, pp. 350-351).

2 MODELS OF BELIEF STATES

2.1 Preliminaries

Before we can start discussing models of belief revision, we must have a way of modelling belief states since a revision method is defined as a function from one belief state into another. The most common models of belief states in computational contexts are *sentential* or *propositional,* in the sense that the elements constituting the belief systems are coded as formulas representing sentences. This kind of model will be the focus of this introduction, but some alternative types of models will be encountered in the volume.

But even if we stick to propositional models of belief systems, there are many options. First of all, we must choose an appropriate *language* to formulate the belief sentences. For example, databases include some form of *rules*, and there are many ways of formalizing these: as quantified sentences in first order logic, as PROLOG rules (corresponding to Horn-clauses), as default statements (e.g., in the style of Reiter (1980)), as probability statements, etc.

In this introduction, I shall work with a language L which is based on first order logic. The details of L will be left open for the time being. It will be assumed that L is closed under applications of the *boolean operators* $\neg$ (negation), $\wedge$ (conjunction), $\vee$ (disjunction) and $\rightarrow$ (implication). We will use ϕ, ψ, χ, etc. as variables over sentences in L. It is also convenient to introduce the symbols $\top$ and $\perp$ for the two sentential constants "truth" and "falsity."

What is accepted in a formal model of a belief state are not only the sentences that are explicitly put into the database, but also the *logical consequences* of these beliefs. Hence, the second factor which has to be decided upon when modelling a belief state is what *logic* governs the beliefs. In practice this depends on which theorem-proving mechanism is used in combination with the database. However, when doing a theoretical analysis, one wants to abstract from the idiosyncracies of a particular algorithm for theorem proving and start from a more general description of the logic. If the logic is undecidable, further complications will arise, but we will ignore these for the time being.

I shall assume that the underlying logic includes *classical propositional logic* and that it is compact.[2] If K logically entails ϕ we will write this as $K \vdash \phi$. Where K is a set of sentences, we shall use the notation Cn(K) for the set of all logical consequences of K, i.e., $Cn(K) = \{\phi: K \vdash \phi\}$. All papers in this volume presume classical logic, except the one by Cross and Thomason where a four-valued logic is used instead.

---

[2] A logic is compact iff whenever A is a logical consequence of a set of sentence K, then there is a *finite* subset K' of K such that A is a logical consequence of K'.

2.2 Belief Sets

The simplest way of modelling a belief state is to represent it by a *set* of sentences from L. Accordingly, we define a *belief set* as a set K of sentences in L which satisfies the following *integrity constraint*:[3]

(I) If K logically entails ψ, then $\psi \in K$.

In logical parlance, (I) says that K is *closed under logical consequences*. The interpretation of such a set is that it contains all the sentences that are *accepted* in the modelled belief state. Consequently, when $\phi \in K$ we say that ϕ is accepted in K and when $\neg\phi \in K$ we say that ϕ is rejected in K. It should be noted that a sentence being accepted does not imply that it has any form of justification or support.[4] A belief set can also be seen as a *theory* which is a partial description of the world. "Partial" because in general there are sentences ϕ such that neither ϕ nor $\neg\phi$ are in K.

By classical logic, whenever K is *inconsistent*, then $K \vdash \phi$ for every sentence ϕ of the language L. This means that there is exactly one inconsistent belief set under our definition, namely, the set of all sentences of L. We introduce the notation $K_\perp$ for this belief set.

2.3 Belief Bases

Against modelling belief states as belief sets it has been argued (Makinson 1985, Hansson 1990, 1991, Nebel 1990, Fuhrmann 1991) that some of our beliefs have no independent standing but arise only as inferences from our more basic belief. It is not possible to express this distinction in a belief set since there are no markers for which beliefs are basic and which are derived. Furthermore, it seems that when we perform revisions or contractions we never do it to the belief set itself which contains an infinite number of elements, but rather on some finite *base* for the belief set.

Formally, this idea can be modelled by saying that B_K is a *base for a belief set* K iff B_K is a finite subset of K and $Cn(B_K) = K$. Then instead of introducing revision and contraction functions that are defined on belief sets it is assumed that these functions are defined on bases. Such functions will be called *base revisions* and *base contractions* respectively. This approach introduces a more finegrained structure since we can have two bases B_K and C_K such that $Cn(B_K) = Cn(C_K)$ but $B_K \neq C_K$. The papers by Nebel and Hansson in this volume concern base revisions. They will be presented in Section 3.5.

---

[3]Belief sets were called *knowledge sets* in Gärdenfors and Makinson (1988).

[4]For further discussion of the interpretation of belief sets cf. Gärdenfors (1988).

There is no general answer to the question of which model is the best of full belief sets or bases, but this depends on the particular application area. Within computer science applications, bases seem easier to handle since they are explicitly finite structures. However, it has been argued in Gärdenfors (1990) that much of the computational advantages of bases for belief sets can be modelled by belief sets together with the notion of *epistemic entrenchment* of beliefs (cf. Section 4.1).

2.4 Possible Worlds Models

An obvious objection to using sets of sentences as models of belief states is that the *objects* of belief are normally not sentences but rather the *contents* of sentences, that is, propositions. The characterization of propositions that has been most popular among philosophers during recent years is to identify them with *sets of possible worlds*. The basic semantic idea connecting sentences with propositions is then that a sentence expresses a given proposition if and only if it is true in exactly those possible worlds that constitute the set of worlds representing the proposition.

By taking beliefs to be beliefs in propositions, we can then model a belief state by a set W_K of possible worlds. The epistemic interpretation of W_K is that it is the narrowest set of possible worlds in which the individual being in the modelled belief state is certain to find the actual world. This kind of model of a belief state has been used by Harper (1977), Grove (1988), among others and in a generalized form by Spohn (1988) (also cf. the comparisons in Gärdenfors (1978)). In this volume, Katsuno and Mendelzon, and Morreau use this way of modelling belief states.

There is a very close correspondence between belief sets and possible worlds models. For any set W_K of possible worlds we can define a corresponding belief set K as the set of those sentences that are true in all worlds in W_K (assuming that the set of propositional atoms is finite). It is easy to verify that K defined in this way satisfies the integrity constraint (I) so that it is indeed a belief set. Conversely, for any belief set K, we can define a corresponding possible worlds model W_K by identifying the possible worlds in W_K with the *maximal consistent extensions* of K. Then we say that a sentence ϕ is *true* in such an extension w iff $\phi \in w$. Again it is easy to verify that this will generate an appropriate possible worlds model (for details cf. Grove (1988)).

From a computational point of view, belief sets are much more tractable than possible worlds models. So even though possible worlds models are popular among logicians, the considerations here show that the two kinds of models are basically equivalent. And if we want to implement belief revision systems, sentential models like belief sets, and in particular bases for belief sets, are much easier to handle.

2.5 Justifications vs. Coherence Models

Another question that has to be answered when modelling a state of belief is whether the *justifications* for the beliefs should be part of the model or not. With respect to this question there are two main approaches. One is the *foundations* theory which holds that one should keep track of the justifications for one's beliefs: Propositions that have no justification should not be accepted as beliefs. The other is the *coherence* theory which holds that one need not consider the pedigree of one's beliefs. The focus is instead on the *logical* structure of the beliefs – what matters is how a belief coheres with the other beliefs that are accepted in the present state.[5] The belief sets presented above clearly fall into the latter category.

It should be obvious that the foundations and the coherence theories have very different implications for what should count as rational *changes* of belief systems. According to the foundations theory, belief revision should consist, first, in giving up all beliefs that no longer have a *satisfactory justification* and, second, in adding new beliefs that have become justified. On the other hand, according to the coherence theory, the objectives are, first, to maintain *consistency* in the revised epistemic state and, second, to make *minimal changes* of the old state that guarantee sufficient overall coherence. Thus, the two theories of belief revision are based on conflicting ideas of what constitutes rational changes of belief. The choice of underlying theory is, of course, also crucial for how a computer scientist will attack the problem of implementing a belief revision system.

Doyle's paper in this volume deals with the relations between justification theories and coherence theories of belief revision. In an earlier paper (Gärdenfors 1990), I presented some arguments for preferring the coherence approach to the foundations approach. Doyle argues that I have overemphasized the differences between the two approaches. He also wants to show that the foundations approach represents the most direct way of making the coherence approach computationally accessible.

Galliers' theory of autonomous belief revision, also in this volume, suggests in another way that the choice between coherence and foundational theories may not be exclusive; her theory in fact represents a blend between the two approaches. In a sense, also the belief base models presented in Section 2.3 show traces of justificationalism – the beliefs in the base are thought of as more foundational than the derived beliefs.

[5] Harman (1986) presents an analysis of the epistemological aspects of the two approaches.

3 RATIONALITY POSTULATES FOR BELIEF REVISION

3.1 The AGM Postulates for Revision

In this section, it will be assumed that belief sets (that is sets of sentences closed under logical consequences) are used as models of belief states. The goal is now to formulate postulates for rational revision and expansion functions defined over such belief sets.

The underlying motivation for these postulates (which are taken from Alchourrón, Gärdenfors, and Makinson (1985), hence the name) is that when we change our beliefs, we want to retain as much as possible from our old beliefs – we want to make a *minimal change*. Information is in general not gratuitous, and unnecessary losses of information are therefore to be avoided. This heuristic criterion may be called the criterion of *informational economy*.

However, it turns out to be difficult to give a precise quantitative definition of the loss of information (see, e.g., the discussion of minimality in Gärdenfors 1988, pp. 66-68). Instead we shall follow another line of specifying 'minimal change': We assume that the sentences in a belief set have different degrees of *epistemic entrenchment*, and when we give up sentences when forming a revision or a contraction, we give up those with the lowest degree of entrenchment. The idea of epistemic entrenchment will be presented in greater detail in Section 4.1.

It is assumed that for every belief set K and every sentence ϕ in L, there is a *unique* belief set $K\dot{+}\phi$ representing the revision of K with respect to ϕ. In other words $\dot{+}$ is a *function* taking a belief set and a sentence as arguments and giving a belief set as a result. This is admittedly a strong assumption, since in many cases, the information available is not sufficient to determine a unique revision. However, from a computational point of view this assumption is gratifying. In Doyle (1991) and Galliers' paper in this volume this assumption is not made.

The first postulate requires that the outputs of the revision function indeed be belief sets:

($K\dot{+}1$) For any sentence ϕ and any belief set K, $K\dot{+}\phi$ is a belief set.

The second postulate guarantees that the input sentence ϕ is accepted in $K\dot{+}\phi$:

($K\dot{+}2$) $\phi \in K\dot{+}\phi$.

The normal application area of a revision process is when the input ϕ contradicts what is already in K, that is $\neg\phi \in K$. However, in order to have the revision function defined for all arguments, we can easily extend it to cover the case when $\neg\phi \notin K$. In this case, revision is identified with expansion. For technical reasons, this identification is divided into two parts:

($K\dot{+}3$) $K\dot{+}\phi \subseteq K+\phi$.

($K\dot{+}4$) If $\neg\phi \notin K$, then $K+\phi \subseteq K\dot{+}\phi$.

The purpose of a revision is to produce a new *consistent* belief set. Thus $K\dot{+}\phi$ should be consistent, unless ϕ is logically impossible:

($K\dot{+}5$) $K\dot{+}\phi = K_{\perp}$ if and only if $\vdash \neg\phi$.

It should be the *content* of the input sentence ϕ rather than its particular linguistic formulation that determines the revision. In other words, belief revisions should be analysed on the *knowledge level* and not on the syntactic level. This means that logically equivalent sentences should lead to identical revisions:

($K\dot{+}6$) If $\vdash \phi \leftrightarrow \psi$, then $K\dot{+}\phi = K\dot{+}\psi$.

The postulates ($K\dot{+}1$) - ($K\dot{+}6$) are elementary requirements that connect K, ϕ and $K\dot{+}\phi$. This set will be called the *basic* set of postulates. The final two conditions concern *composite* belief revisions. The idea is that, if $K\dot{+}\phi$ is a revision of K and $K\dot{+}\phi$ is to be changed by a further sentence ψ, such a change should be made by expansions of $K\dot{+}\phi$ whenever possible. More generally, the minimal change of K to include both ϕ and ψ, that is, $K\dot{+}\phi\wedge\psi$, ought to be the same as the expansion of $K\dot{+}\phi$ by ψ, so long as ψ does not contradict the beliefs in $K\dot{+}\phi$. For technical reasons the precise formulation is split into two postulates:

($K\dot{+}7$) $K\dot{+}\phi\wedge\psi \subseteq (K\dot{+}\phi)+\psi$.

($K\dot{+}8$) If $\neg\psi \notin K\dot{+}\phi$, then $(K\dot{+}\phi)+\psi \subseteq K\dot{+}\phi\wedge\psi$.

When $\neg\psi \in K$, then $(K\dot{+}\phi)+\psi$ is $K_{\perp}$, which is why the proviso is needed in ($K\dot{+}8$) but not in ($K\dot{+}7$).

We turn next to some consequences of the postulates. It can be shown (Gärdenfors, 1988, p. 57) that in the presence of the basic set of postulates ($K\dot{+}7$) is equivalent to:

(1) $K\dot{+}\phi \cap K\dot{+}\psi \subseteq K\dot{+}\phi\vee\psi$.

Another principle that is useful is the following 'factoring' condition:

(2) $K\dot{+}\phi\vee\psi = K\dot{+}\phi$ or $K\dot{+}\phi\vee\psi = K\dot{+}\psi$ or $K\dot{+}\phi\vee\psi = K\dot{+}\phi \cap K\dot{+}\psi$.

It can be shown that, given the basic postulates, (2) is in fact equivalent to the conjunction of ($K\dot{+}7$) and ($K\dot{+}8$).

Furthermore ($K\dot{+}7$) and ($K\dot{+}8$) together entail the following identity criterion:

(3) $K\dot{+}\phi = K\dot{+}\psi$ if and only if $\psi \in K\dot{+}\phi$ and $\phi \in K\dot{+}\psi$.

The postulates ($K\dotplus 1$) - ($K\dotplus 8$) do not uniquely characterise the revision $K\dotplus\phi$ in terms of only K and ϕ. This is, however, as it should be. I believe it would be a mistake to expect that only logical properties are sufficient to characterise the revision process.

3.2 The AGM Postulates for Contraction

The postulates for the contraction function '$\dot-$' will, to an even larger extent than for revisions, be motivated by the princple of informational economy. The first postulate is of a familiar kind:

($K\dot-1$) For any sentence ϕ and any belief set K, $K\dot-\phi$ is a belief set.

Because $K\dot-\phi$ is formed from K by giving up some beliefs, it should be required that no new beliefs occur in $K\dot-\phi$:

($K\dot-2$) $K\dot-\phi \subseteq K$.

When $\phi \notin K$, the criterion of informational economy requires that nothing be retracted from K:

($K\dot-3$) If $\phi \notin K$, then $K\dot-\phi = K$.

We also postulate that the sentence to be contracted not be a logical consequence of the beliefs retained in $K\dot-\phi$ (unless ϕ is logically valid in which case it can never be retracted because of the integrity constraint (I)):

($K\dot-4$) If not $\vdash \phi$, then $\phi \notin K\dot-\phi$.

From ($K\dot-1$) to ($K\dot-4$) it follows that

(4) If $\phi \notin K$, then $(K\dot-\phi)+\phi \subseteq K$.

In other words, if we first retract ϕ and then add ϕ again to the resulting belief set $K\dot-\phi$, no beliefs are accepted that were not accepted in the original belief set. The criterion of informational economy demands that as many beliefs as possible should be kept in $K\dot-\phi$. One way of guaranteeing this is to require that expanding $K\dot-\phi$ by ϕ should take us back to exactly the same state as before the contraction, that is K:

($K\dot-5$) If $\phi \in K$, then $K \subseteq (K\dot-\phi)+\phi$.

This is the so called *recovery postulate*, which enables us to 'undo' contractions. It has turned out to be the most controversial among the AGM postulates for contraction.

The sixth postulate is analogous to ($K\dotplus 6$):

($K\dot-6$) If $\vdash \phi \leftrightarrow \psi$, then $K\dot-\phi = K\dot-\psi$.

Postulates ($K\dot{-}1$) - ($K\dot{-}6$) are called the *basic set* of postulates for contractions. Again, two further postulates for contractions with respect to conjunctions will be added. The motivations for these postulates are much the same as for ($K\dot{+}7$) and ($K\dot{+}8$).

($K\dot{-}7$) $\quad K\dot{-}\phi \cap K\dot{-}\psi \subseteq K\dot{-}\phi\wedge\psi$.

($K\dot{-}8$) $\quad$ If $\phi \notin K\dot{-}\phi\wedge\psi$, then $K\dot{-}\phi\wedge\psi \subseteq K\dot{-}\psi$.

It is interesting to note that ($K\dot{-}7$) is in fact equivalent, given the basic postulates, to the seemingly weaker

(5) $\quad K\dot{-}\phi \cap Cn(\{\phi\}) \subseteq K\dot{-}\phi\wedge\psi$.

In parallel with (2) it can be shown that ($K\dot{-}7$) and ($K\dot{-}8$) are jointly equivalent to the following condition:

(6) $\quad K\dot{-}\phi\wedge\psi = K\dot{-}\phi$ or $K\dot{-}\phi\wedge\psi = K\dot{-}\psi$ or $K\dot{-}\phi\wedge\psi = K\dot{-}\phi \cap K\dot{-}\psi$.

A useful consequence of (6) is the following which says that $K\dot{-}\phi\wedge\psi$ is 'covered' either by $K\dot{-}\phi$ or by $K\dot{-}\psi$:

(7) $\quad$ Either $K\dot{-}\phi\wedge\psi \subseteq K\dot{-}\phi$ or $K\dot{-}\phi\wedge\psi \subseteq K\dot{-}\psi$.

The postulates for revision and contraction and their consequences are dicussed further in Chapter 3 of Gärdenfors (1988).

3.3 From Contractions to Revisions and vice versa

We turn next to a study of the connections between revision and contraction functions. In the previous two sections they were characterized by two sets of postulates. These postulates are *independent* in the sense that the postulates for revisions do not refer to contractions and vice versa. A natural question is now whether either contraction or revision can be defined in terms of the other. Here we shall present two positive answers to this question.

A revision of a knowledge set can be seen as a composition of a contraction and an expansion. More precisely: In order to construct the revision $K\dot{+}\phi$, one first contracts K with respect to $\neg\phi$ and then expands $K\dot{-}\neg\phi$ by ϕ. Formally, we have the following definition which is called the *Levi identity*:

(Def $\dot{+}$) $\quad K\dot{+}\phi = (K\dot{-}\neg\phi)+\phi$

That this definition is appropriate is shown by the following result:

Theorem 1: If a contraction function '$\dot{-}$' satisfies ($K\dot{-}1$) to ($K\dot{-}4$) and ($K\dot{-}6$), then the revision function '$\dot{+}$' obtained from (Def $\dot{+}$) satisfies ($K\dot{+}1$) - ($K\dot{+}6$). Furthermore, if ($K\dot{-}7$) also is satisfied, ($K\dot{+}7$) will be satisfied for the defined revision function; and if ($K\dot{-}8$) also is satisfied, ($K\dot{+}8$) will be satisfied for the defined revision function.

This result supports (Def $\dot{+}$) as an appropriate definition of a revision function. Note that the controversial recovery postulate ($K\dot{-}5$) is not used in the theorem.

Conversely, contractions can be defined in terms of revisions. The idea is that a sentence ψ is accepted in the contraction $K\dot{-}\phi$ if and only if ψ is accepted both in K and in $K\dot{+}\neg\phi$. Formally, this amounts to the following definition which has been called the *Harper identity*:

(Def $\dot{-}$) $\quad K\dot{-}\phi = K \cap K\dot{+}\neg\phi$.

Again, this definition is supported by the following result:

Theorem 2: If a revision function '$\dot{+}$' satisfies ($K\dot{+}$-1) to ($K\dot{+}6$), then the contraction function '$\dot{-}$' obtained from (Def $\dot{-}$) satisfies ($K\dot{-}1$) - ($K\dot{-}6$). Furthermore, if ($K\dot{+}7$) is satisfied, ($K\dot{-}7$) will be satisfied for the defined contraction function; and if ($K\dot{+}8$) is satisfied, ($K\dot{-}8$) will be satisfied for the defined contraction function.

The two theorems show that the defined revision and contraction functions have the right properties. Hence, the two sets of postulates for revision and contraction functions are interchangeable and a method for constructing one of the functions would automatically, via (Def $\dot{+}$) or (Def $\dot{-}$), yield a construction of the other function satisfying the desired set of postulates.

3.4 Representation Theorems

This section will introduce a first kind of explicit modelling of a contraction function for belief sets. Via the Levi identity (Def $\dot{+}$) and Theorem 1, such a model can be used to define a revision function as well.

The problem in focus is how to define the contraction $K\dot{-}\phi$ with respect to a belief set K and a proposition ϕ. A general idea is to start from K and then give some recipe for choosing which propositions to delete from K so that $K\dot{-}\phi$ does not contain ϕ as a logical consequence. According to the criterion of informational economy we should look at as large a subset of K as possible.

The following notion is useful: A belief set K' is a *maximal subset of K that fails to imply* ϕ if and only if (i) $K' \subseteq K$, (ii) $\phi \notin Cn(K')$, and (iii) for any K" such that $K' \subset K'' \subseteq K$, $\phi \in Cn(K'')$. The last clause entails that if K' were to be expanded by some sentence from K-K' it would entail ϕ. The set of all belief sets that fail to imply ϕ will be denoted $K\perp\phi$. Using the assumption that $\vdash$ is compact it is easy to show that this set is nonempty, unless ϕ is logically valid.

A first tentative solution to the problem of constructing a contraction function is to identify $K\dot{-}\phi$ with one of the maximal subsets in $K\perp\phi$. Technically, this can be done with the aid of a *selection function* γ that picks out an element $\gamma(K\perp\phi)$ of $K\perp\phi$ for any K and any ϕ whenever $K\perp\phi$ is nonempty. We then define $K\dot{-}\phi$ by the following rule:

(*Maxichoice*) $K\dot{-}\phi = \gamma(K\perp\phi)$ when not $\vdash\phi$, and $K\dot{-}\phi = K$ otherwise.

Contraction functions determined by some such selection function were called *maxichoice contraction functions* in Alchourrón, Gärdenfors and Makinson (1985).

A first test for this construction is whether it has the desirable properties. It is easy to show that any maxichoice contraction function satisfies ($K\dot{-}1$) - ($K\dot{-}6$). But it will also satisfy the following *fullness* condition:

($K\dot{-}F$) If $\psi \in K$ and $\psi \notin K\dot{-}\phi$, then $\psi \rightarrow \phi \in K\dot{-}\phi$ for any belief set K.

We can now show that ($K\dot{-}1$) - ($K\dot{-}6$) and ($K\dot{-}F$) characterizes maxichoice contraction function in the sense of the following *representation theorem*. Let us say that a contraction function '$\dot{-}$' can be *generated* by a maxichoice contraction function iff there is some selection function γ such that '$\dot{-}$' is identical with the function obtained from γ by the maxichoice rule above.

Theorem 3: Any contraction function that satisfies ($K\dot{-}1$) - ($K\dot{-}6$) and ($K\dot{-}F$) can be generated by a maxichoice contraction function.

However, in a sense, maxichoice contraction functions in general produce contractions that are *too large*. A result from Alchourrón and Makinson (1982) is applicable here: Let us say that a belief set K is *maximal* iff for every sentence ψ, either $\psi \in K$ or $\neg\psi \in K$. One can now show the following discomforting result:

Theorem 4: If a revision function '$\dot{+}$' is defined from a maxichoice contraction function '$\dot{-}$' by means of the Levi identity, then, for any ϕ such that $\neg\phi \in K$, $K\dot{+}\phi$ will be maximal.

In a sense, maxichoice contraction functions create maximal belief sets. So a second tentative idea is to assume that $K\dot{-}\phi$ contains only the propositions that are *common to all* of the maximal subsets in $K\perp\phi$:

(*Meet*) $K\dot{-}\phi = \cap(K\perp\phi)$ whenever $K\perp\phi$ is nonempty and $K\dot{-}\phi = K$ otherwise.

This kind of function was called *full meet contraction function* in Alchourrón, Gärdenfors, and Makinson (1985). Again, it is easy to show that any full meet contraction function satisfies ($K\dot{-}1$) - ($K\dot{-}6$). They also satisfy the following*intersection* condition:

($K\dot{-}I$) For all ϕ and ψ, $K\dot{-}\phi\wedge\psi = K\dot{-}\phi \cap K\dot{-}\psi$.

We have the following representation theorem:

Theorem 5: A contraction function satisfies (K$\dot{-}$1) - (K$\dot{-}$6) and (K$\dot{-}$I) iff it can be generated as a full meet contraction function.

The drawback of of full meet contraction is the opposite of maxichoice contraction – in general it results in contracted belief sets that are far *too small*. The following result is proved in Alchourrón and Makinson (1982):

Theorem 6: If a revision function '$\dot{+}$' is defined from a full meet contraction function '$\dot{-}$' by means of the Levi identity, then, for any ϕ such that $\neg\phi \in K$, $K\dot{+}\phi = Cn(\{\phi\})$.

In other words, the revision will contain only ϕ and its logical consequences.

A third attempt is to use only *some* of the maximal subsets in $K\perp\phi$ when defining $K\dot{-}\phi$. Technically, a *selection function* γ can be used to pick out a nonempty *subset* $\gamma(K\perp\phi)$ of $K\perp\phi$, if the latter is nonempty, and that puts $\gamma(K\perp\phi) = K$ in the limiting case when $K\perp\phi$ is empty. The contraction function can then be defined as follows:

(*Partial meet*): $K\dot{-}\phi = \cap\gamma(K\perp\phi)$.

Such a contraction function was called a *partial meet contraction function* in Alchourrón, Gärdenfors, and Makinson (1985). The following representation theorem shows that (K$\dot{-}$1) - (K$\dot{-}$6) indeed characterizes the class of partial meet contraction functions:

Theorem 7: For every belief set K, '$\dot{-}$' is a partial meet contraction function *iff* '$\dot{-}$' satisfies postulates (K$\dot{-}$1) - (K$\dot{-}$6).

So far we have put no constraints on the selection function γ. The idea of γ picking out the 'best' elements of $K\perp\phi$ can be made more precise by assuming that there is an *ordering* of the maximal subsets in $K\perp\phi$ that can be used to pick out the top elements. Technically, we do this by introducing the notation M(K) for the *union* of the family of all the sets $K\perp\phi$, where ϕ is any proposition in K that is not logically valid. Then it is assumed that there exists a *transitive and reflexive* ordering relation $\leq$ on M(K). When $K\perp\phi$ is nonempty, this relation can be used to *define* a selection function that picks out the top elements in the ordering:

(Def γ) $\gamma(K\perp\phi) = \{K' \in K\perp\phi: K'' \leq K' \text{ for all } K'' \in K\perp\phi\}$

A contraction function that is determined from $\leq$ via the selection function γ given by (Def γ) will be called a *transitively relational partial meet contraction function*. This way of defining the selection function constrains the class of partial meet contraction functions that can be generated:

Theorem 8: For any belief set K, '$\dot{-}$' satisfies (K$\dot{-}$1) - (K$\dot{-}$8) iff '$\dot{-}$' is a transitively relational partial meet contraction function.

Thus we have found a way of connecting the rationality postulates with a general way of modelling contraction functions. The drawback of the construction is that the computational costs involved in determining the content of the relevant maximal subsets of a belief set K are so overwhelming that we should take a look at some other possible solutions to the problem of constructing belief revisions and contractions.

3.5 Contraction and Revision of Bases

As a generalization of the AGM postulates several authors have suggested postulates for revisions and contractions of *bases* for belief sets rather than the belief sets themselves. In this volume the papers by Hansson and Nebel (see also Fuhrmann 1989, Hansson 1989, 1991, Makinson 1987, Nebel 1990) use this kind of model. As Hansson writes in his paper, "this model is based on the intuition that some of our beliefs have no independent standing but arise only as inferences from our more basic beliefs."

Hansson and Nebel analyse various forms of base revision and base contractions. Nebel evaluates his models, in a number of theorems, in relation to the AGM postulates, but Hansson also introduces some postulates that are special for base revision. For example, his postulate of *relevance* can (slightly simplified) be written as follows in my terminology:

(8) If $\psi \in H$, but $\psi \notin H \dot{-} \phi$, then there is some H' such that $H \dot{-} \phi \subset H' \subseteq H$ and $\phi \notin Cn(H')$, but $\phi \in Cn(H' \cup \{\psi\})$.

Here H denotes a *finite base* for a belief state consisting of sentences from L. The logical closure of H, that is Cn(H), will be a belief set. The intuition behind this postulate is that if a sentence ψ is retracted from H when ϕ is rejected, then ψ plays some role in the fact that H but not $H \dot{-} \phi$ logically entails ϕ. On the basis of relevance and other postulates for base contraction, Hansson proves several representation theorems (and further results can be found in Hansson (1991)).

An interesting feature of Nebel's paper is that he investigates the *computational complexity* of different belief revision procedures. As far as I know, he is the first one to attack these issues. An initial problem is that already the trivial case of deciding whether $\psi \in Cn(\emptyset) \dot{+} \phi$ is co-NP-complete so that a more finegrained set of complexity classes are needed than just saying that belief revision is NP-hard. Nebel solves this problem by using the polynomial hierarchy of complexity classes (Garey and Johnson 1979). On the basis of this hierarchy, he is then able to prove a number of results concerning the complexity of various revision methods. The analysis shows that all base revision methods analyzed in his paper that satisfy the full set of AGM postulates turn out to be no harder than ordinary propositional derivability.

4 CONSTRUCTIVE MODELS

4.1 Epistemic Entrenchment

Even if all sentences in a belief set are accepted or considered as facts (so that they are assigned maximal probability), this does not mean that all sentences are are of equal value for planning or problem-solving purposes. Certain pieces of our knowledge and beliefs about the world are more important than others when planning future actions, conducting scientific investigations, or reasoning in general. We will say that some sentences in a belief system have a higher degree of *epistemic entrenchment* than others. This degree of entrenchment will, intuitively, have a bearing on what is abandoned from a belief set, and what is retained, when a contraction or a revision is carried out. This section begins by presenting a set of postulates for epistemic entrenchment which will serve as a basis for a constructive definition of appropriate revision and contraction functions.

The guiding idea for the construction is that when a belief set K is revised or contracted, the sentences in K that are given up are those having the lowest degrees of epistemic entrenchment. Fagin, Ullman and Vardi (1983), pp. 358 ff., introduce the notion of "database priorities" which is closely related to the idea of epistemic entrenchment and is used in a similar way to update belief sets. However, they do not present any axiomatization of this notion. Section 5 of Nebel's paper in this volume provides a precise characterization of the relationship between epistemic entrenchment and database priorities.

We will not assume that one can quantitatively measure degrees of epistemic entrenchment, but will only work with *qualitative* properties of this notion. One reason for this is that we want to emphasise that the problem of uniquely specifying a revision function (or a contraction function) can be solved, assuming only very little structure on the belief sets apart from their logical properties.

If ϕ and ψ are sentences in L, the notation $\phi \leq \psi$ will be used as a shorthand for "ψ is at least as epistemically entrenched as ϕ." The strict relation $\phi < \psi$, representing "ψ is epistemically more entrenched than ϕ," is defined as $\phi \leq \psi$ and not $\psi \leq \phi$.

Postulates for epistemic entrenchment:

(EE1)	If $\phi \leq \psi$ and $\psi \leq \chi$, then $\phi \leq \chi$	(transitivity)
(EE2)	If $\phi \vdash \psi$, then $\phi \leq \psi$	(dominance)
(EE3)	For any ϕ and ψ, $\phi \leq \phi \wedge \psi$ or $\psi \leq \phi \wedge \psi$	(conjunctiveness)
(EE4)	When $K \neq K_\perp$, $\phi \notin K$ iff $\phi \leq \psi$, for all ψ	(minimality)
(EE5)	If $\psi \leq \phi$ for all ψ, then $\vdash \phi$	(maximality)

The justification for (EE2) is that if ϕ logically entails ψ, and either ϕ or ψ must be retracted from K, then it will be a smaller change to give up ϕ and retain ψ rather than to give up ψ, because then ϕ must be retracted too, if we want the revised belief set to satisfy the integrity constraint (I). The rationale for (EE3) is as follows: If one wants to retract $\phi \wedge \psi$ from K, this can only be achieved by giving up either ϕ or ψ and, consequently, the informational loss incurred by giving up $\phi \wedge \psi$ will be the same as the loss incurred by giving up ϕ or that incurred by giving up ψ. (Note that it follows already from (EE2) that $\phi\wedge\psi \leq \phi$ and $\phi\wedge\psi \leq \psi$.) The postulates (EE4) and (EE5) only take care of limiting cases: (EE4) requires that sentences already not in K have minimal epistemic entrenchment in relation to K; and (EE5) says that only logically valid sentences can be maximal in $\leq$. (The converse of (EE5) follows from (EE2), since if $\vdash \phi$, then $\psi \vdash \phi$, for all ψ.)

It should be noted that the relation $\leq$ is only defined *in relation to a given K* – different belief sets may be associated with different orderings of epistemic entrenchment.[6]

We mention the following simple consequences of these postulates:

Lemma: Suppose the ordering $\leq$ satisfies (EE1) - (EE3). Then it also has the following properties:

(i) $\phi \leq \psi$ or $\psi \leq \phi$ (connectivity);

(ii) If $\psi\wedge\chi \leq \phi$, then $\psi \leq \phi$ or $\chi \leq \phi$;

(iii) $\phi < \psi$ iff $\phi\wedge\psi < \psi$.

(iv) If $\chi \leq \phi$ and $\chi \leq \psi$, then $\chi \leq \phi\wedge\psi$.

(v) If $\phi \leq \psi$, then $\phi \leq \phi\wedge\psi$.

The main purpose of this section is to show the connections between orderings of epistemic entrenchment and the AGM contraction and revision functions presented in Sections 3.1 and 3.2. We will accomplish this by providing two conditions, one of which determines an ordering of epistemic entrenchment assuming a contraction function and a belief set as given, and the other of which determines a contraction function assuming an ordering of epistemic entrenchment and a belief set as given. The first condition is:

(C$\leq$) $\phi \leq \psi$ if and only if $\phi \notin K \dot{-} \phi\wedge\psi$ or $\vdash \phi\wedge\psi$.

The idea underlying this definition is that when we contract K with respect to $\phi\wedge\psi$ we must give up ϕ or ψ (or both) and ϕ should be retracted just in case ψ is at least as epistemically entrenched as ϕ. In the limiting case when both ϕ and ψ are logically valid, they are of equal epistemic entrenchment (in conformity with (EE2)).

The second, and from a constructive point of view most central, condition gives an explicit definition of a contraction function in terms of the relation of epistemic entrenchment:

(C$\dot{-}$) $\psi \in K \dot{-} \phi$ if and only if $\psi \in K$ and either $\phi < \phi \vee \psi$ or $\vdash \phi$.

---

[6]Rott (1992) has developed a generalized notion of epistemic entrenchment which is not dependent on a particular K.

Condition ($C\dot{-}$) provides us with a tool for explicitly defining a contraction function in terms of the ordering $\leq$. An encouraging test of the appropriateness of such a definition is the following theorem proved in Gärdenfors and Makinson (1988):

Theorem 9: If an ordering $\leq$ satisfies (EE1) - (EE5), then the contraction function which is uniquely determined by ($C\dot{-}$) satisfies ($K\dot{-}1$) - ($K\dot{-}8$) as well as the condition ($C\leq$).

Conversely, we can show that if we start from a given contraction function and determine an ordering of epistemic entrenchment with the aid of condition ($C\leq$), the ordering will have the desired properties:

Theorem 10: If a contraction function '$\dot{-}$' satisfies ($K\dot{-}1$) - ($K\dot{-}8$), then the ordering $\leq$ that is uniquely determined by ($C\leq$) satisfies (EE1) - (EE5) as well as the condition ($C\dot{-}$).

These results suggest that the problem of constructing appropriate contraction and revision functions can be *reduced* to the problem of providing an appropriate ordering of epistemic entrenchment. Furthermore, condition ($C\dot{-}$) gives an explicit answer to which sentences are included in the contracted belief set, given the initial belief set and an ordering of epistemic entrenchment. From a computational point of view, applying ($C\dot{-}$) is trivial, once the ordering $\leq$ of the elements of K is given.

The comparison $\phi < \phi \vee \psi$ in ($C\dot{-}$) is somewhat counterintuitive. Rott (1991) has investigated the following more natural version of the condition:

($C\dot{-}R$) $\quad \psi \in K\dot{-}_R\phi$ if and only if $\psi \in K$ and either $\phi < \psi$ or $\vdash\phi$.

He then shows that the contraction function '$\dot{-}_R$' defined in this way has the following properties:

Theorem 11: Let '$\dot{-}_R$' be the contraction function defined in ($C\dot{-}R$). If $\leq$ satisfies (EE1) - (EE5), then '$\dot{-}_R$' satisfies ($K\dot{-}1$) - ($K\dot{-}4$) and ($K\dot{-}6$) - ($K\dot{-}8$), but not ($K\dot{-}5$).

Since '$\dot{-}_R$' does not satisfy the controversial 'recovery' postulate ($K\dot{-}5$), it follows that '$\dot{-}_R$' defined by ($C\dot{-}R$) is in general not identical to '$\dot{-}$' defined by ($C\dot{-}$).[7] However, let '$\dot{+}_R$' and '$\dot{+}$' be the revision functions defined from '$\dot{-}_R$' and '$\dot{-}$' by the Levi identity. Rott proves:

Theorem 12: '$\dot{+}_R$' and '$\dot{+}$' are identical revision functions.

A consequence of this theorem is that if we are only interested in modelling revisions and not contractions, we can use the extremely simple test ($C\dot{-}R$) when computing the revision functions, without having to bother about the disjunctions in ($C\dot{-}$).

4.2 Safe Contraction

[7]'$-_R$' is a 'withdrawal function' in the sense of Makinson (1987).

Yet another approach to the problem of constructing contraction functions was introduced by Alchourrón and Makinson (1985) and is called *safe contraction*. Their contraction procedure can be described as follows: Let K be a belief set, and suppose that we want to contract K with respect to ϕ. Alchourrón and Makinson postulate a "hierarchy" < over K that is assumed to be acyclical (that is, for no ϕ_1, ϕ_2 ..., ϕ_n in K is it the case that $\phi_1 < \phi_2 < \ldots < \phi_n < \phi_1$). Given such a hierarchy, we say that an element ψ is *safe* with respect to ϕ iff ψ is not a minimal element (under <) of any *minimal* subset K' of K such that K' $\vdash \phi$. Equivalently, every minimal subset K' of K such that K' $\vdash \phi$ either does not contain ψ or else contains some χ such that $\chi < \psi$. Intuitively, the idea is that ψ is safe if it can never be "blamed" for the implication of ϕ. Note that, in contrast to the earlier constructions, this definition uses *minimal* subsets of K that *entail* ϕ rather than maximal subsets of K that do *not* entail ϕ.

Rott's paper in this volume concerns the relation between orderings of epistemic entrenchment and the hierarchies over K used in the definition of safe contraction. He presents ways of translating between the types of orderings and proves that they are equivalent. This is in contrast to what was conjectured by Alchourrón and Makinson (1985) and Gärdenfors (1988). In this way, he completes the map of correspondences between (1) the AGM postulates for contractions, (2) the AGM postulates for revisions, (3) relational partial meet contractions functions, (4) epistemic entrenchment contractions, and (5) safe contractions.

4.3 Possibility Theory

Apart from the five areas mentioned in the previous paragraph, *possibility theory* can be added as a sixth area which can be connected to epistemic entrenchment relations, in the first place, and thereby also indirectly with belief revisions and contraction. This is the topic of Dubois and Prade's contribution to the volume.

Perhaps the best way of relating possibility theory to the theory of belief revision is to start with the *qualitative necessity relation* (Dubois 1986) which is an ordering $\geq_c$ (and the corresponding strict relation $>_c$) of sentences satisfying the following axioms:

(A0) $\top >_c \bot$

(A1) $\phi \geq_c \psi$ or $\psi \geq_c \phi$

(A2) $\phi \geq_c \psi$ and $\psi \geq_c \chi$ imply $\phi \geq_c \chi$

(A3) $\phi \geq_c \bot$

(C) if $\phi \geq_c \psi$, then, for all χ, $\phi \wedge \chi \geq_c \psi \wedge \chi$

A qualitative necessity relation can be generated from a *necessity measure* N so that $\phi \geq_c \psi$ if and only if $N(\phi) \geq N(\psi)$, where N satisfies the following characteristic property:

(9) $N(\phi \wedge \psi) = \min(N(\phi),N(\psi))$

Dually, one can define a possibility measure Π as a function satisfying:

(10) $\Pi(\phi \vee \psi) = \max(\Pi(\phi),\Pi(\psi))$

Indeed, possibility measures are related to necessity measures through the relationship $N(\phi) = 1 - \Pi(\neg\phi)$. The dual qualitative possibility ordering $\geq_\Pi$ can be related to $\geq_C$ by the following equivalence:

(11) $\phi \geq_\Pi \psi$ if and only if $\neg\psi \geq_C \neg\phi$.

It has been shown by Dubois and Prade (1991) that a qualitative necessity ordering is almost identical to an epistemic entrenchment relation. The exception is that for epistemic entrenchment it is requested that $\top >_C \phi$ instead of (A0).

This connection between possibility theory and epistemic entrenchment forms the starting point for the paper by Dubois and Prade in this volume. They represent belief states by necessity measures and rewrite the rationality postulates for revision and contraction accordingly. They also show how this model of a belief state can be used to describe updating with *uncertain* pieces of evidence. They make some interesting comparisons with Spohn´s (1988) *ordinal conditional functions*, which form a different way of introducing degrees of belief.

4.4 Updates vs. Revisions

Katsuno and Mendelzon present an interesting alternative to revision in their contribution to this volume. This alternative method is called *updating* and is also used by Morreau in his paper on planning. The basic idea is that one needs to make a distinction between two kinds of information and the corresponding changes. On the one hand there is new information about a *static* world. For this kind of information the revision process, as it has been described is appropriate. On the other hand, there is new information about *changes* in the world brought about by some agent. For both types, the new piece of information may be inconsistent with the current state of belief. However Katsuno, Mendelzon and Morreau argue that the revision process is inadequate as a model of rational belief change caused by the second type of information. For this kind of change an *updating* procedure is appropriate

To illustrate their argument, let me borrow an example from Winslett (1988). Suppose that all we know in K about a particular room is that there is a table, a book and a magazine in it, and that either (β) the book is on the table, or (μ) the magazine is on the table, but not both, i.e., the belief state K is essentially $Cn((\beta\wedge\neg\mu) \vee (\mu\wedge\neg\beta))$. A robot is then ordered to put the book on the table, and as a consequence, we learn that β. If we change our beliefs by revision we should, according to ($K\dot{+}4$), end up in a belief state that contains $\beta\wedge\neg\mu$ since β is consistent with K. But why should we conclude that the magazine is not on the table?

In order to describe how updating works, we must present their version of a possible worlds model for belief states. Let L be the language of standard propositional logic and let P be the set of propositional letters in L. An *interpretation* of L is a function I from P to the set {T,F} of truth values. This function is extended to L recursively in the standard way, so that $I(\phi \wedge \psi) = T$ iff $I(\phi) = T$ and $I(\psi) = T$, etc. A *model* of a sentence ϕ is an interpretation I such that $I(\phi) = T$. A model of a set of sentences K is an interpretation I such that $I(\phi) = T$, for all $\phi \in K$. Mod(K) denotes the set of all models of K.

Instead of using an ordering of $K\perp\neg\phi$ (the maximal consistent subsets of K that don't entail $\neg\phi$) when determining $K\dot{+}\phi$, Katsuno and Mendelzon (and several other researchers; see Katsuno and Mendelzon's (1989) survey) have proposed to look at an *ordering of the set of all interpretations* and then use this ordering to decide which interpretations should constitute *models of* $K\dot{+}\phi$, and thus indirectly determine $K\dot{+}\phi$ in this way. The intended meaning of such an ordering is that some interpretations that are models of ϕ (but not of K) are *closer* to models of K than other interpretations. Such an ordering of interpretations should, of course, be dependent on K.

Technically, we assign to each belief set K a pre-ordering $\leq_K$ over the set of interpretations of L (and a corresponding strict ordering $<_K$).[8] Following Katsuno and Mendelzon we say that $\leq_K$ is *faithful*[9] if these conditions hold:

(i) $\leq_K$ is transitive and reflexive.

(ii) If $I, J \in Mod(K)$, then $I <_K J$ does not hold.

(iii) If $I \in Mod(K)$ and $J \notin Mod(K)$, then $I <_K J$.

If M is a set of interpretations of L, we let $Min(M,\leq_K)$ denote the set of interpretations I which are minimal in M with respect to $\leq_K$. $K\dot{+}\phi$ can now be determined from $\leq_K$ as the belief set which has exactly $Min(Mod(\{\phi\}),\leq_K)$ as its set of models. Katsuno and Mendelzon (1989) prove the following general result:

Theorem 13: A revision function $\dot{+}$ satisfies ($K\dot{+}1$) - ($K\dot{+}8$) if and only if there exists a total faithful ordering $\leq_K$ such $Mod(K\dot{+}\phi) = Min(Mod(\{\phi\}),\leq_K)$.

This theorem gives a representation of belief revision in their terminology. Using this terminology the difference between revising and updating can be described as follows: Methods for revising K by ϕ that satisfy ($K\dot{+}1$) - ($K\dot{+}8$) are exactly those that select from the models of ϕ that are 'closest' to the models of K. In contrast, update models select, for

[8]To be precise, Katsuno and Mendelson only consider belief sets that can be represented by a *single* sentence from L (i.e., the conjunction of all the beliefs in K).

[9]This condition is called 'persistent' in Katsuno and Mendelzon (1989).

each model I of K, the set of models of ϕ that are *closest to I*.[10] The update $K\dot{+}\phi$ is then characterized by the union of all such models.

The difference between the methods may seem marginal at a first glance, but the properties of updating are, in general, quite different from those of revision. In connection with the example above, we have already noted that updating violates ($K\dot{+}4$). On the other hand, updating satisfies the following postulate, which is violated by revision:

(12) $(K \vee K')\dot{+}\phi = (K\dot{+}\phi) \vee (K'\dot{+}\phi)$

This postulates presumes that belief states are modelled by single sentences so that disjunctions of belief states are well defined.

Morreau's paper in this volume is an application to *planning* of the updating procedure. Using this method he presents a framework for modelling reasoning about action. The language he uses includes a conditional operator > which is used to represent statements of the form 'If the agent were to do α, this would result in ϕ being true.' The conditionals are analyzed both semantically and axiomatically. By nesting conditionals of this kind he can describe the content of sequences of actions, and in this way obtain an elegant way of representing planning.

4.5 Autonomous Belief Revision and Communication

Postulate ($K\dot{+}2$) requires that an input ϕ given to a belief state K *must* be accepted in the revision $K\dot{+}\phi$. Galliers argues in her paper in this volume that this is in conflict with the *autonomy* of the agents having the various belief states. In *communication*, one agent informs another about something, aiming at changing the receiver's belief state. It remains, however, the decision of the autonomous receiver whether the information should be accepted or not. As Gallier's puts it, "Autonomous agents may or may not comply with the recognised intended effects of an utterance on their cognitive states. There are no specialised rules dictating what is a cooperative response. Rational communicative action must therefore be planned not only as purposive, but as *strategic*."

So in order to model belief revisions in a communicative setting one must be able to specify *whether* to accept the contents of an utterance from another agent, as well as *how* to perform the possible revisions caused by the utterance. The underlying principle for Galliers is that the acceptability of a new utterance is dependent on the degree of coherence of the belief state that would result if the utterance were added. Her model of revision is essentially determined by such a coherence ordering, where the degree of coherence is defined as maximal derivability of *core beliefs*. It is not required that there be a unique

---

[10] This method is essentially equivalent to 'imaging' as introduced in a probabilistic context by Lewis (1976) and generalized to the context of belief sets in Gärdenfors (1988).

revision, and, furthermore, it is not required that any preferred revision incorporate the communicated information.

Galliers then adds a foundational aspect to the belief revision model by working with *assumptions* of various kinds and *justifications* for the assumptions. For example, the endorsement of an assumption depends on whether it is communicated by a reliable source or a spurious source. In this way, her model model of autonomous belief revision is a mixture of coherence and foundationalism. The model has been implemented as a component of a strategic planner for cooperative dialogue.

4.6 Conditionals and the Ramsey Test

There is a close connection between belief revisions and the meaning of *conditional sentences*. The carrying hypothesis is that conditional sentences, in various forms, are about changes of states of belief. The form of conditional sentence that is central is "if ϕ were the case, then β would be the case" or "if ϕ is the case, then β is (will be) the case," where ϕ may or may not contradict what is already accepted in a given epistemic state K. If ϕ contradicts what is accepted in K, the conditional is called a *counterfactual* (relative to K), otherwise it is called an *open conditional* (relative to K).

The epistemic semantics for counterfactuals and open conditionals will be based on F. P. Ramsey's test for evaluating a conditional sentence. His test can be described as follows: In order to find out whether a conditional sentence is acceptable in a given state of belief, one first adds the antecedent of the conditional hypothetically to the given stock of beliefs. Second, if the antecedent together with the formerly accepted sentences leads to a contradiction, then one makes some adjustments, as small as possible without modifying the hypothetical belief in the antecedent, such that consistency is maintained. Finally, one considers whether or not the consequent of the conditional is accepted in this adjusted state of belief.

Given the analysis of belief revisions in Section 2, we see that it is very natural to reformulate the Ramsey test in a more condensed way:

(RT) $\phi > \beta \in K$ iff $\beta \in K\dot{+}\phi$

This test has attracted a great deal of attention as a possible starting point for a formal semantics of conditionals. Ginsberg (1986) argues that a formal semantics for counter-factuals is of great value for many problem areas within AI, in particular since they form the core of nonmonotonic inferences.

Note that the formulation of (RT) presupposes that sentences of the form $\phi > \beta$ belong to the object language and that they can be elements of the belief sets in a belief revision model. Let us call this extended object language L'.

Some results in Gärdenfors (1978) seem to justify the claim that the Ramsey test can be used as a basis for an epistemic semantics of conditionals. However, the list of conditions that were used to generate the logic of conditionals does not include $(K\dotplus 4)$ (or the full strength of $(K\dotplus 8)$). An interesting question is whether it is possible to use (RT) together with $(K\dotplus 4)$ when analysing the logic of conditionals. With minor qualifications, the answer turns out to be *no*. In order to put the result as strongly as possible, one can start from the following *preservation* condition:

$(K\dotplus P)$ If $\neg\phi \notin K$ and $\psi \in K$, then $\psi \in K\dotplus\phi$

It is easy to show that the preservation criterion is essentially equivalent to $(K\dotplus 4)$.

The Ramsey test and the preservation criterion are each of considerable interest for the analysis of the dynamics of belief. Unfortunately, it can be proved that, on the pain of triviality, the Ramsey test and the preservation criterion are inconsistent with each other (Gärdenfors 1986).

Let us formulate this result with some care. A background assumption is that the revision function is defined for all belief sets.[11] Note that in L', sentences containing the conditional connective '>' will be treated on a par with sentences without this operator. The Ramsey test (RT) is, of course, dependent on this assumption. A consequence of (RT) that is crucial is the following *monotonicity* criterion:

$(K\dotplus M)$ For all belief sets K and K' and all ϕ, if $K \subseteq K'$, then $K\dotplus\phi \subseteq K'\dotplus\phi$.

The conditions on the revision function that will be needed for the proof are $(K\dotplus 2)$ and the following very weak criterion, which is one half of $(K\dotplus 5)$:

$(K\dotplus 5w)$ If $K \neq K_\perp$ and $K\dotplus\phi = K_\perp$, then $\vdash \neg\phi$.

The final assumption that will be needed for the inconsistency result is that the belief revision system is non-trivial. As usual, two propositions ϕ and ψ are said to be disjoint iff $\vdash \neg(\phi \wedge \psi)$. A belief revision system will be said to be non-trivial iff there are at least three pairwise disjoint sentences ϕ, ψ, and χ and some belief set K which is consistent with all three sentences, i.e., $\neg\phi \notin K$, $\neg\psi \notin K$, and $\neg\chi \notin K$.

Theorem 14: There is no non-trivial belief revision system that satisfies all the conditions $(K\dotplus 2)$, $(K\dotplus 5w)$, $(K\dotplus M)$ and $(K\dotplus P)$.

It should be noted that the conditional connective '>' is used neither in the formulation of the theorem nor in its proof. If $(K\dotplus P)$ is replaced by $(K\dotplus 4)$, then $(K\dotplus 2)$ is not needed for the proof of the theorem.

---

[11] What is needed for the proof is only the assumption that if K is in the domain of the revision function, so are are all *expansions* $K+\phi$.

Corollary: There is no non-trivial belief revision system that satisfies all the conditions $(K\dotplus 2)$, $(K\dotplus 5w)$, (RT) and $(K\dotplus P)$.

The theorem and its corollary show that the Ramsey test (RT) and the preservation condition $(K\dotplus P)$ (or, equivalently, $(K\dotplus 4)$) cannot both be rational criteria for belief revisions. However, the Ramsey test has a great deal of appeal, and several ways of getting around the impossibility result have been tried. In his paper in this volume, Morreau uses *updating* instead of revision in formulating the Ramsey test (as was noted above, updating does not satisfy $(K\dotplus 4)$) and it is easy to show that this combination is consistent.

Another approach is taken by Cross and Thomason in their paper in this volume. They also retain the Ramsey test, but they work with a different logical framework than what has been used above. Firstly, they use a four-valued logic which changes the accompanying proof theory. Secondly, they restrict revision to *atomistic* inputs, i.e., the only sentences for which the revision process is defined are either atomic sentences or negated atomic sentences.

By restricting the revision procedure in this way, Cross and Thomason show that it is now possible to work out a theory of conditionals that satisfies the Ramsey test. Furthermore, they show that the theory of nonmonotonic *inheritance* from Horty, Thomason, and Touretzky (1990) can be interpreted as a special case of their logic of conditionals. In this way we find yet another connection between the theory of belief revision and other areas of computer science.

Acknowledgements

I wish to thank Hirofumi Katsuno, Michael Morreau, Bernhard Nebel, Richmond Thomason and Hans Rott for helpful comments on an earlier version of this paper.

REFERENCES

Alchourrón, C. E., P. Gärdenfors, and D. Makinson (1985): "On the logic of theory change: Partial meet contraction and revision functions," *The Journal of Symbolic Logic 50*, 510-530.

Alchourrón, C. E. and Makinson, D. (1985): "On the logic of theory change: Safe contraction," *Studia Logica 44*, 405-422.

Doyle, J. (1991): "Rational belief revision (Preliminary report)," *Principles of Knowledge Representation and Reasoning*, ed. by J. Allen, R. Fikes, and E. Sandewall, Los Altos, CA: Morgan Kaufmann, 163-174.

Dubois, D. (1986): "Belief structures, possibility theory and decomposable confidence measures on finite sets," *Computers and Artificial Intelligence 5*, 403-416.

Dubois, D. and H. Prade (1991): "Epistemic entrenchment and possibility logic," to appear in *Artificial Intelligence*.

Fagin, R., Ullman, J. D., and Vardi, M. Y. (1983): "On the semantics of updates in databases," *Proceedings of Second ACM SIGACT-SIGMOD*, Atlanta, pp. 352-365.

Fuhrmann, A. (1991): "Theory contraction through base contraction," *Journal of Philosophical Logic 20*, 175-203.

Gärdenfors, P. (1978): "Conditionals and changes of belief," in *The Logic and Epistemology of Scientific Change*, ed. by I. Niiniluoto and R. Tuomela, *Acta Philosophica Fennica 30*, 381-404.

Gärdenfors, P. (1986): "Belief revisions and the Ramsey test for conditionals," *The Philosophical Review 95*, 81-93.

Gärdenfors P. (1988): *Knowledge in Flux: Modeling the Dynamics of Epistemic States*, Cambridge, MA: The MIT Press, Bradford Books.

Gärdenfors, P. and D. Makinson. (1988): "Revisions of knowledge systems using epistemic entrenchment," in *Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge*, M. Vardi ed., Los Altos, CA: Morgan Kaufmann.

Garey, M. R. and Johnson, D. S. (1979): *Computers and Intractability – A Guide to the Theory of NP-Completeness*, San Francisco: Freeman.

Ginsberg, M. L. (1986): "Counterfactuals," *Artificial Intelligence 30*, 35-79.

Grove, A. (1988): "Two modellings for theory change," *Journal of Philosophical Logic 17*, 157-170.

Hansson, S. O. (1989): "In defense of base contraction," to appear in *Synthese*.

Hansson, S. O. (1991): *Belief Base Dynamics*, Uppsala: Acta Universitatis Upsaliensis.

Harman, G. (1986): *Change in Viev: Principles of Reasoning*, Cambridge, MA: The MIT Press, Bradford Books.

Harper, W. L. (1977): "Rational conceptual change," in *PSA 1976*, East Lansing, Mich: Philosophy of Science Association, vol. 2, 462-494.

Horty, J. F., Thomason R. H., and Touretsky, D. S. (1990): "A skeptical theory of inheritance in nomonotonic semantic networks," *Artificial Intelligence 42*, 311-348.

Katsuno, H. and Mendelzon, A. O. (1989): "A unified view of propositional knowledge base updates," *Proceedings of the 11th Internationsl Joint Conference on Artificial Intelligence*, San Mateo, CA: Morgan Kaufmann, 269-276.

Lewis, D. K. (1976): "Probabilities of conditionals and conditional probabilities," *The Philosophical Review 85*, 297-315.

Makinson, D. (1985): "How to give it up: A survey of some formal aspects of the logic of theory change," *Synthese 62*, 347-363.

Makinson, D. (1987): "On the status of the postulate of recovery in the logic of theory change," *Journal of Philosophical Logic 16*, 383-394.

Nebel, B. (1990): *Reasoning and Revision in Hybrid Representation Systems*, *Lecture Notes in Computer Science*, vol. 422, Berlin: Springer

Reiter R. (1980): "A logic for default reasoning," *Artificial Intelligence 13*, 81-132.

Rott, H. (1991):"Two methods of constructing contractions and revisions of knowledge systems," *Journal of Philosophical Logic 20*, 149-173.

Rott, H. (1992): ”Preferential belief change using generalized epistemic entrenchment”, to appear in *Journal of Logic, Language and Information*.

Spohn, W. (1988): "Ordinal conditional functions: A dynamic theory of epistemic states," in W. L. Harper and B. Skyrms, eds, *Causation in Decision, Belief Change, and Statistics, vol. 2*, Dordrecht: Reidel, 105-134.

Winslett, M. (1988): "Reasoning about action using a possible models approach," *Proceedings of the Seventh National Conference on Artificial Intelligence*, 89-93.