Sequences of Revisions: On the Semantics of Nested Conditionals

by
Craig Boutilier

Technical Report 92-23
September 1992

Department of Computer Science
University of British Columbia
Rm 333 - 6356 Agricultural Road
Vancouver, B.C.
CANADA V6T 1Z2

Sequences of Revisions: On the Semantics of Nested Conditionals

Craig Boutilier
Department of Computer Science
University of British Columbia
Vancouver, British Columbia
CANADA, V6T 1Z2
email: cebly@cs.ubc.ca

September 20, 1992

Abstract

The truth conditions for conditional sentences have been well-studied, but few compelling attempts have been made to define means of evaluating iterated or nested conditionals. We start with a semantic account of subjunctive conditionals based on the AGM model of revision, and extend this model in a natural fashion to account for right-nesting of conditionals, describing a process called *natural revision*. These sentences capture sequences of propositional revisions of a knowledge base. We examine the properties of this model, demonstrating that the maximum amount of conditional information in a belief set is preserved after revision. Furthermore, we show how any sequence of revisions can be reduced to natural revision by a single sentence. This demonstrates that any arbitrarily nested sentence is equivalent to a sentence without nesting of the conditional connective. We show cases where revision models, even after the processing of an arbitrary sequence of revisions, can be described purely propositionally, and often in a manner that permits tractable inference. We also examine a form of revision known as *paranoid revision* which appears to be the simplest form of belief revision that fits within the AGM framework, and captures semantically the notion of full meet revision.

1 Introduction

Subjunctive conditionals have recently attracted much attention in the knowledge representation community. It has been pointed out that counterfactuals may play a large role in planning and diagnostic systems (Ginsberg 1986), that subjunctives may be used to capture knowledge base update and revision (Katsuno and Mendelzon 1990; Katsuno and Mendelzon 1991; Boutilier 1992d), and that they are intimately related to the conditionals used in default reasoning (Boutilier 1992g; Makinson and Gärdenfors 1990).

We denote by $A > B$ the subjunctive conditional "If A were the case then B would be true." Various subjunctive and counterfactual logics have been proposed to account for properties of the connective $>$, the most influential accounts being those of Stalnaker (1968) and Lewis (1973). The truth conditions for $A > B$ are typically presented in terms of possible worlds and a *similarity relation* over these worlds. Loosely, $A > B$ is true at a particular world if B is true at those worlds satisfying A that are *most similar* to the world in question. Different properties of the similarity relation correspond to different conditional logics.

From the point of view of knowledge representation, *acceptance* conditions for $A > B$ are especially important. Under what conditions should an agent assent to the conditional? Unlike "ordinary" objective beliefs, an agent cannot simply "see" if $A > B$ is true since its truth conditions are phrased in terms of other situations, to which an agent does not have direct access. For this reason, the *Ramsey test* provides a natural model for acceptance of a conditional:

> First add the antecedent (hypothetically) to your stock of beliefs; second make whatever adjustments are required to maintain consistency (without modifying the hypothetical belief in the antecedent); finally, consider whether or not the consequent is true. (Stalnaker 1968, p.44)

The key step in the Ramsey test is the revision of the belief set. The notion of revision adopted will determine which conditionals are accepted and rejected. Conversely, given a fixed (complete) set of accepted conditionals, the revision function adopted by an agent will also be determined: revising by A is simply a matter of believing those B such that $A > B$ is accepted. Consequently, the study of revision and conditional logic are virtually the same if one accepts the Ramsey test.

It should be clear that a semantic model of *sequences* of revisions of a belief set is also vital. Consider some diagnostic task where tests are being performed over time to discover some fault. As test results become available, these have to be incorporated into a knowledge base that has been revised by earlier results. One may argue that we simply need to revise our knowledge base once, using the conjunction of all results. However, it may be the case that certain results suggest the implausibility, or even contradict, earlier results. One test result may cause one to believe that component A is faulty, followed by a result suggesting fault with component B rather than A.[1] The conjunction of the two "observations" is not a reasonable candidate for revision. But neither should the first observation be thrown away (at least, in many circumstances); for it may influence the manner in which the second is incorporated into the knowledge base. Similar comments apply to other domains. For example, while planning, an agent must change its belief state as it discovers facts about the world that help it make decisions, and as it acts upon the world, changing the truth of various propositions. This is a natural application of a theory of revision sequences.[2]

The most prominent theory of belief revision is that put forth by Alchourrón, Gärdenfors and Makinson (1985) and expounded by Gärdenfors (1988). Within this framework many people have explored the connection to conditionals.[3] The *AGM theory*, which we describe in

[1]Indeed, in a diagnosis task of this type it is likely that some "degrees of belief" or probabilities are attached to these potential conflicts. A conflicting observation of the second type may simply reduce confidence in the first hypothesis. In the last section, we discuss generalizations of the model presented here that can deal with degrees of *possibility* or degrees of *probability*.

[2]It has been observed by Winslett (1988) that belief revision is inappropriate for changing a knowledge base in response to a changing world. She has proposed an operation known as *update* that captures this type of change in belief. Update has been formalized and studied by Kastuno and Mendelzon (1991). In contrast, belief *revision*, which we describe below, seems more suitable for modeling change in belief about a static world (correcting mistakes, so to speak). It is clear that both types of belief change are necessary for planning. An agent requires an update facility to realize the effects of its actions, while it must be able to revise its beliefs as it learns new information, perhaps as the result of various "information-gathering actions" (Haddawy and Hanks 1992). We do not discuss update in this paper, though the techniques used here could be extended to update. We will occasionally lapse and use the term "update" to refer to a new piece of information to be added to a belief set. This is not to be taken as an update in the sense of Katsuno and Mendelzon. Rather, the reader uncomfortable with the term should substitute "revision," as that is the only type of belief change we examine here.

[3]However, the connections have not always been smooth due to the limited expressive power of the logics in question, the confusion between truth and acceptance conditions, and the celebrated triviality results of Gärdenfors (1986). See (Gärdenfors 1988; Gärdenfors 1986; Rott 1989; Fuhrmann 1989; Boutilier 1992d; Boutilier 1992e).

the next section, imposes various constraints on acceptable revision functions. Roughly, the revision function preserves as much information as possible. Unfortunately, the AGM theory has little to say about revision sequences. Gärdenfors (1988) proposes *belief revision systems* that map one belief set to another when a revision is applied to the first. These mappings must respect the AGM theory, but the AGM postulates do not constrain the behavior of the revision of this *new* belief set, given its origins from the first.[4] This means the conditionals accepted in the new belief set need not be related to those in the first set (except for certain obvious antecedents). Thus, while information content is "preserved" with respect to objective beliefs, the information content of conditionals is ignored.

It is this problem of preserving conditionals under revision that we investigate here. The semantic model of AGM revision (for propositional belief sets) we describe in the next section orders possible worlds according to their plausibility, or "degree of consistency" with a fixed belief set K. While this ordering guides the selection of a revised belief set K_A^* (which incorporates A into K), this AGM model fails to provide a new ordering suitable for the revision of K_A^*. The goal of this paper is to show how one might use the original ordering to constrain the new ordering, hence the revision of the *revised* belief set K_A^*. We propose that the new ordering retain as much of the old ordering as possible, consistent with the AGM postulates. This *minimal change* in the ordering is precisely defined a provides the semantic basis for our model of revision sequences. This approach, dubbed *natural revision* ensures that a maximal subset of one's conditional beliefs is retained during the revision process. Of course, this type of revision is a specific instance of the more general phenomenon of iterated revision captured by general belief revision systems. However, we shall argue that our model provides a very natural way of extending the concept of "minimal change" or "informational economy," the hallmark of the AGM theory, to sequences of revisions and conditional beliefs. Indeed, natural revision provably retains the maximum amount of conditional information consistent with the constraints of the AGM theory.

In the next section we present the details of the AGM theory and review a semantic model and logical characterization due to Boutilier (1992d). We adopt this as the basic model of

[4]Strictly speaking, this is not true; for *consistent* revisions are specified (see the next section), providing one exception.

revision and describe in general terms our natural model of nesting. In Section 3, we present a very simple subclass of our general revision models and show how iterated conditionals are captured in this subclass. This form of *paranoid revision* uniquely determines the truth of a nested conditional based solely on the propositional content of a belief set, and provides a semantic account of *full meet revision* (Alchourrón, Gärdenfors and Makinson 1985). These considerations, however, cannot be extended in a natural way to the more general case. This drawback leads us to postulate a characterization of arbitrary revision sequences (or arbitrary right-nested conditionals) based on a certain "intended" model of revision. In Section 4 we describe this model, examine the behavior of natural revision with respect to the preservation of conditionals under revision. Of particular importance is Theorem 13 and its corollaries, which demonstrate that natural revision preserves as many conditional beliefs as possible. We also show that any revision sequence can be captured by a single revision on the original belief set (Theorem 26). We also describe how such a sentence is computed (Theorems 28 and 30). Given the Ramsey test, this implies that any arbitrary right-nested conditional is equivalent to a conditional without nesting. Because this model preserves information in a strong sense, we can show that no belief set that occurs later in a revision sequence can be smaller than an earlier belief state (Corollary 32). In Section 5 we describe how nested conditionals can be evaluated in our logic such that our "intended" model is captured, and circumstances under which this can be reduced to propositional reasoning. Proofs of the main results can be found in the appendix.

2 A Logic for Revision

An important and well-studied problem in philosophical logic, database theory and artificial intelligence is that of modeling theory change or belief revision. Suppose K is a deductively closed set of beliefs. Revising K is required when new information must be integrated with these beliefs. If $K \not\models \neg A$, learning A is relatively unproblematic as the new belief set $Cn(K \cup \{A\})$ seems adequate for modeling this change. This process is known as *expansion*. More troublesome is the revision of K by A when $K \models \neg A$. Some beliefs in K must be given up before A can be accommodated. The problem is in determining which part of K to give up, as there are a multitude of choices. Furthermore, in general, there are no logical grounds for choosing which of these alternative revisions is acceptable (Stalnaker 1984), the issue depending largely on context. Fortunately, there are some logical criteria for reducing this set of possibilities.

The main criterion for discarding some revisions in deference to others is that of *minimal change*. Informational economy dictates that as "few" beliefs as possible from K be discarded in order to facilitate belief in A (Gärdenfors 1988). By "few" we intend that, as much as possible, the informational content of K is kept intact. In particular, if KB is a finite representation of K, we do not require that as few sentences as possible from KB be given up, only that the information implicit in these sentences is minimal.[5] While pragmatic considerations will often enter into these deliberations, the main emphasis of the work of Alchourrón, Gärdenfors and Makinson (1985) is in logically delimiting the scope of acceptable revisions. To this end, the AGM postulates below, are maintained to hold for any reasonable notion of revision (Gärdenfors 1988). We will use K^*_A to denote revision of K by A. K^+_A to denote expansion, and $\perp$ to denote the identically false proposition.

(R1) K^*_A is a belief set (i.e. deductively closed).

(R2) $A \in K^*_A$.

(R3) $K^*_A \subseteq K^+_A$.

(R4) If $\neg A \notin K$ then $K^+_A \subseteq K^*_A$.

(R5) $K^*_A = Cn(\perp)$ iff $\models \neg A$.

(R6) If $\models A \equiv B$ then $K^*_A = K^*_B$.

(R7) $K^*_{A \wedge B} \subseteq (K^*_A)^+_B$.

(R8) If $\neg B \notin K^*_A$ then $(K^*_A)^+_B \subseteq K^*_{A \wedge B}$.

[5]Indeed, we do not require that the revised set be equal to the closure of *any* subset of KB (contrast (Nebel 1989) where the syntax, and not the semantic content, of KB is crucial).

2.1 The Logic CO*

The bimodal logics CO and CO* were first presented in (Boutilier 1991) for reasoning about conditional statements of normality. The modal logic CO is based on a standard propositional modal language (over variables $\mathbf{P}$) augmented with an additional modal operator $\overleftarrow{\Box}$. We denote by $\mathbf{L}_B$ this bimodal language and by $\mathbf{L}_{CPL}$ its classical propositional sublanguage. The sentence $\overleftarrow{\Box}\alpha$ is read "α is true at all *inaccessible* worlds" (in contrast to the usual $\Box\alpha$ that refers to truth at accessible worlds). A CO-model is a triple $M = \langle W, R, \varphi \rangle$, where W is a set of worlds with valuation φ and R is an accessibility relation over W. We insist that R be transitive and connected.[6]

The notion of *cluster* will play a large role in future developments. In any reflexive, transitive Kripke frame, a cluster is any maximal mutually accessible set of worlds (Segerberg 1970). In other words, a set $C \subseteq W$ is a cluster just when wRv for all $v, w \in C$ and no extension $C' \supset C$ has this property. We note that CO-structures consist of a totally-ordered set of clusters of mutually accessible worlds. We use $\|A\|$ to denote the set of worlds w in a model M (usually understood) satisfying sentence A, that is, such that $M \models_w A$. We refer to such worlds w as A-worlds. We also us this notation for sets of sentences K, twK denoting those worlds satisfying each element of K, and somewhat loosely use $\neg K$-worlds to refer to worlds falsifying some element of K.

Satisfaction is defined in the usual way, with the truth of a modal formula at a world defined as:

1. $M \models_w \Box\alpha$ iff for each v such that wRv, $M \models_v \alpha$.

2. $M \models_w \overleftarrow{\Box}\alpha$ iff for each v such that not wRv, $M \models_v \alpha$.

We define several new connectives as follows: $\Diamond\alpha \equiv_{\mathrm{df}} \neg\Box\neg\alpha$; $\overleftarrow{\Diamond}\alpha \equiv_{\mathrm{df}} \neg\overleftarrow{\Box}\neg\alpha$; $\overleftrightarrow{\Box}\alpha \equiv_{\mathrm{df}} \Box\alpha \wedge \overleftarrow{\Box}\alpha$; and $\overleftrightarrow{\Diamond}\alpha \equiv_{\mathrm{df}} \Diamond\alpha \vee \overleftarrow{\Diamond}\alpha$. It is easy to verify that these connectives have the following truth conditions: $\Diamond\alpha$ ($\overleftarrow{\Diamond}\alpha$) is true at a world if α holds at some accessible (inaccessible) world; $\overleftrightarrow{\Box}\alpha$ ($\overleftrightarrow{\Diamond}\alpha$) holds iff α holds at all (some) worlds. CO is captured axiomatically as follows.

[6] R is (totally) connected if wRv or vRw for any $v, w \in W$ (this implies reflexivity). This restriction is relaxed in (Boutilier 1992a), where we develop a weaker logic CT4O based on a reflexive, transitive accessibility relation.

Definition (Boutilier 1991) The conditional logic CO is the smallest $S \subseteq \mathbf{L}$ such that S contains CPL (and its substitution instances) and the following axiom schemata, and is closed under the following rules of inference:

K $\Box(A \supset B) \supset (\Box A \supset \Box B)$

K′ $\overleftarrow{\Box}(A \supset B) \supset (\overleftarrow{\Box} A \supset \overleftarrow{\Box} B)$

T $\Box A \supset A$

4 $\Box A \supset \Box\Box A$

S $A \supset \overleftarrow{\Box}\Diamond A$

H $\overleftrightarrow{\Diamond}(\Box A \wedge \overleftarrow{\Box} B) \supset \overleftrightarrow{\Box}(A \vee B)$

Nes From A infer $\overleftrightarrow{\Box} A$.

MP From $A \supset B$ and A infer B.

Provability and derivability are defined in the standard fashion, in terms of theoremhood.

Theorem 1 (Boutilier 1991) $\vdash_{CO} \alpha$ *iff* $\models_{CO} \alpha$.

In many circumstances we want to ensure that all logically possible worlds are taken into consideration. This is crucial for the solution of irrelevance in (Boutilier 1991), and will be required for our logical account of belief revision. For this purpose we introduce the logic CO*, which is based on the class of CO-models in which all propositional valuations are represented (see also (Levesque 1990)). For all $w \in W$, let w^* be the map from $\mathbf{P}$ into $\{0,1\}$ such that $w^*(A) = 1$ iff $w \in \varphi(A)$; in other words, w^* is the valuation associated with w.

Definition (Boutilier 1991) A *CO*-model* is any CO-model $M = \langle W, R, \varphi \rangle$, such that

$$\{f : f \text{ maps } \mathbf{P} \text{ into } \{0,1\}\} \subseteq \{w^* : w \in W\}$$

Definition (Boutilier 1991) CO* is the smallest extension of CO closed under all rules of CO and containing the following axioms:

NB $\overleftarrow{\Box}\alpha \supset \neg\Box\alpha$ for all falsifiable propositional α.[7]

[7] Alternatively, we could use $\overleftrightarrow{\Diamond}\alpha$ for all satisfiable α.

Theorem 2 (Boutilier 1991) $\vdash_{CO*} \alpha$ *iff* $\models_{CO*} \alpha$.

2.2 Representation of Revision Functions

We can use CO*-models to represent the revision of a theory K (for a more detailed account we refer to (Boutilier 1992d)). The interpretation of R is as follows: wRv iff v is as at least as *plausible* a state of affairs as w given an agent's belief state K. As usual, v is *more plausible* than w iff wRv but not vRw. Plausibility is a pragmatic measure that reflects the degree to which one would accept w as a possible state of affairs given that belief in K may have to be given up. If v is more plausible than w, loosely speaking v is "more consistent" with the belief set K than w. We can think of the clusters of worlds in a CO*-model as equivalence classes of equally plausible worlds. Thus each cluster "assigns" a unique "plausibility ranking" to its members. If we have two distinct clusters C and C' such that C sees C' (i.e., wRv for some $w \in C$ and $v \in C'$), then the degree of plausibility C' is greater than that of C.

Another requirement is that those worlds consistent with our belief set K should be exactly those minimal in R. That is, vRw for all $v \in W$ iff $M \models_w K$. This condition ensures that no world is more plausible than any world consistent with K, and that all K-worlds are equally plausible. It is just the epistemically possible worlds (those consistent with K) that should be most plausible. Such models are called *K-revision models* and have as their minimal cluster the set $\|K\|$. This constraint can be expressed in our language as

$$\overleftrightarrow{\Box}(KB \supset (\Box KB \wedge \overleftarrow{\Box}\neg KB)) \tag{1}$$

for any K that is finitely expressible as *KB*. This ensures that any *KB*-world sees every other *KB*-world ($\overleftarrow{\Box}\neg KB$), and that it sees only *KB*-worlds ($\Box KB$). All statements about revision are implicitly evaluated with respect to *KB*. We abbreviate sentence (1) as $O(KB)$ and intend it to mean we "only know" *KB*.

When we use a model $M = \langle W, R, \varphi \rangle$ to represent the revision of some belief set K, we intend that the minimal worlds in the relation R are exactly those an agent considers epistemically possible. To believe (each sentence in) K is to consider only those worlds satisfying K to be consistent with the agent's beliefs. Thus, only K-worlds should be minimal in R. However, when we assert K (or *KB*) as a belief set we usually have in mind that those

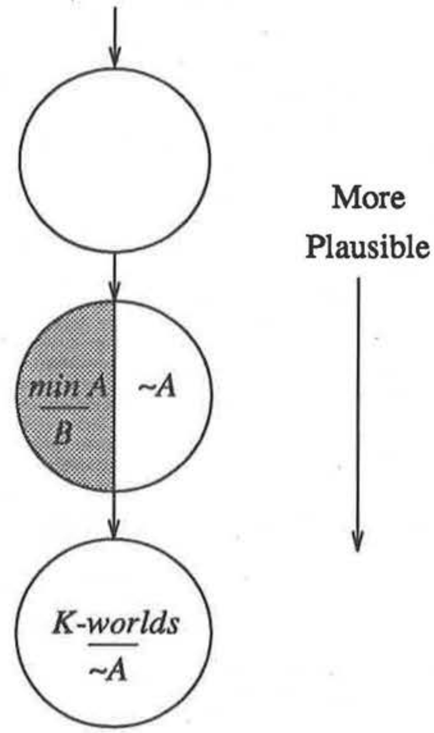


Figure 1: Truth conditions for the conditional

sentences in K are *all* that is believed. If a certain world w is not considered possible, there should be some knowledge in K that forces its exclusion, some sentence $\alpha \in K$ such that $w \not\models \alpha$. If K is all that is believed, no K-world should be epistemically impossible, for otherwise the agent would know more than K, having some additional knowledge corresponding to this exclusion. When *all and only* K-worlds are minimal in R, the agent only knows K (see Levesque (1990) for further details regarding only knowing). This is precisely the constraint imposed by $O(KB)$.

Given this structure, we want the set of minimal A-worlds to represent the state of affairs believed when K is revised by A, since these are the most plausible worlds, the ones we are most willing to adopt, given A. In Figure 1, we have a typical K-revision model with each large circle representing a cluster of mutually accessible (equally plausible) worlds, with arrows indicating accessibility between clusters. The minimal cluster consists of all K-worlds, and we have that $K \vdash \neg A$. The set of minimal A-worlds is indicated by the shaded region, and this set forms the set of "newly possible" worlds when K is revised by A. Thus $A \xrightarrow{\text{KB}} B$ should hold just when B is true at each world in the shaded region.

Of course such a *minimal* set of A-worlds may not exist (e.g., consider an infinite chain of more and more plausible A-worlds). Still $A \xrightarrow{\text{KB}} B$ should still be true if, at any point on this chain, B holds at all more plausible A-worlds (hence, B is true at some hypothetical limit of this chain). We can define the connective as follows:

$$A \xrightarrow{\text{KB}} B \equiv_{\text{df}} \overleftrightarrow{\Box}\neg A \vee \overleftrightarrow{\Diamond}(A \wedge \Box(A \supset B)). \tag{2}$$

This sentence is true in the trivial case when A is impossible. It also holds when the second disjunct is satisfied, which states that there is some world w such that A holds and $A \supset B$ holds at all worlds still closer than w. Thus B holds at the most plausible A-worlds (whether this is a "hypothetical" or actual limit). [8] It is important to note that $\xrightarrow{\text{KB}}$ does not describe a family of related connectives indexed by KB. It is a conditional connective in the usual sense. "KB" is used to emphasize the fact that $\xrightarrow{\text{KB}}$ is typically used for the revision of some intended knowledge base. The connective is perfectly well-defined and meaningful when KB is left unspecified.

We define for any $A \in \mathbf{L}_{CPL}$ the belief set resulting from revision of K by A as follows:

$$K_A^{*^{\text{M}}} = \{B \in \mathbf{L}_{CPL} : M \models A \xrightarrow{\text{KB}} B\}. \tag{3}$$

We can show that $*$ satisfies the AGM postulates for belief revision and any AGM revision operator has an equivalent formulation as such a $*$.

Theorem 3 (Boutilier 1992d) *Let M be a K-revision model and $*^{\text{M}}$ the revision function determined by M. Then $*^{\text{M}}$ satisfies postulates (R1) through (R8).*

Theorem 4 (Boutilier 1992d) *Let $*$ be a revision function satisfying postulates (R1) through (R8). Then for any theory K there exists a K-revision model M such that $K_A^* = K_A^{*^{\text{M}}}$ for all A.*

[8] In this manner we avoid the *Limit Assumption* (Lewis 1973; Stalnaker 1968). This has certain practical implications for revision as well (Boutilier 1992a).

Thus, we can use the logic CO* to represent the revision of a theory *KB*, and reason about the results of this revision, in a manner respecting the AGM postulates. In fact, CO* appears to be the first purely logical characterization of AGM revision (in the sense of providing an explicit language and consequence operator for revision).

2.3 Beliefs and Epistemic Logic

The distinguishing feature of CO*, which permits this characterization of AGM revision, is its ability to express that a belief set can be *only known.* This enables consistency with a belief set to be represented directly in the object level theory. Indeed, CO* is a reasonable epistemic logic, as well. We can define a modality for belief $\boxtimes$, reading $\boxtimes\alpha$ as "α is believed." This sentence will hold just when α is true at each epistemically possible world, those minimal in the plausibility ordering R. Hence, we define belief as:

$$\boxtimes\alpha \equiv_{\mathrm{df}} \overleftrightarrow{\Diamond}\Box\alpha$$

We will have occasion to use this modality here, and we note that it behaves according to the usual weak S5 interpretation of belief.

For any CO*-model, we can define the *objective belief set* associated with it to be those propositional sentences that are "believed" in the model.

Definition 1 Let M be a CO*-model. The *objective belief set* associated with M is

$$\{\alpha \in \mathbf{L}_{CPL} : M \models \boxtimes\alpha\}$$

We will sometimes refer to this as the propositional belief set or simply the belief set for M. Naturally, the belief set for any K-revision model is just K. We will be more interested in "subjective beliefs" associated with a revision model, those sentences that are believed and involve some modal operators. Of particular concern are those conditionals that are believed by an agent. We therefore extend the notion of belief set to cover arbitrary sentences.

Definition 2 Let M be a CO*-model. The *extended belief set* associated with M is

$$\{\alpha \in \mathbf{L}_B : M \models \boxtimes\alpha\}$$

For any CO*-model with belief set K and extended belief set E, we have $K \subseteq E$. While we have clear characterization of the revised belief set K_A^* from M, it is less clear what form the revised extended belief, denoted E_A^* should take.

An alternative model of revision is based on the notion of *epistemic entrenchment* (Gärdenfors 1988). Given a belief set K, we can characterize the revision of K by ordering beliefs according to our willingness to give them up when some contradictory information requires such. If one of two beliefs must be retracted in order to accommodate some new fact, the least entrenched belief will be relinquished, while the most entrenched persists. Gärdenfors (1988) presents five postulates for such an ordering and shows that these orderings determine exactly the space of revision functions satisfying the AGM postulates. We let $B \leq_E A$ denote the fact that A is at least as entrenched as B in theory K. We note that such a statement can be faithfully represented in CO* as $\overleftrightarrow{\Box}(\neg A \supset \Diamond\neg B)$ (see (Boutilier 1992b) for further details). Thus a complete set of sentences of this form is sufficient to specify a revision function.

We note that the dual of an entrenchment ordering is a plausibility ordering on sentences. A sentence A is more plausible than B just when $\neg A$ is less entrenched than $\neg B$, and means that A would be more readily accepted than B if the opportunity arose. Grove (1988) studied this relationship and its connection to the AGM theory.

2.4 The Problem of Iterated Revision

Typical semantic accounts of revision (and conditional logics generally) take $B \in K_A^*$ to hold just when B is true at all *minimal* A-worlds (Grove 1988; Katsuno and Mendelzon 1990; Kraus, Lehmann and Magidor 1990). Notice that our models do not require that such minimal, or most plausible, A-worlds exist and our definition of $A \xrightarrow{\text{KB}} B$ reflects this (in much the same manner as Lewis's (1973) counterfactual semantics). However, the representation result of Grove (1988) can be applied directly to our CO*-model of revision. Therefore, we can legitimately restrict our attention to the class of models whose members satisfy the following well-foundedness condition.

Definition 3 Let $M = \langle W, R, \varphi \rangle$ be a CO-model. For any $A \in \mathbf{L}_{CPL}$ we define

$$\min(M, A) = \{w \in W : M \models_w A, \text{ and } M \models_v A \text{ implies } vRw \text{ for all } v \in W\}$$

A model is *well-founded* iff, for all A, $\min(A, M) \neq \emptyset$.

Grove shows that such models are adequate for the representation of AGM revision functions and in what follows we will assume our CO* revision models are well-founded: for each $A \in \mathbf{L}_{CPL}$ there will be a nonempty set of minimal A-worlds. According to our definitions of $*^{\mathrm{M}}$ and $\xrightarrow{\mathrm{KB}}$, this ensures that $K_A^{*^{\mathrm{M}}}$ will consist of exactly those objective sentences true at each world in this minimal set.

If we are curious about the status of proposition B after an arbitrary propositional revision A of belief set K, we may simply ask if $B \in K_A^*$. Naturally, because of the Ramsey test, this holds iff $A \xrightarrow{\mathrm{KB}} B$ is in the extended belief set E containing K. The new belief set K_A^* is adequately represented by the set of minimal A-worlds in the K-revision model (say M) being used. While this characterization of K_A^* is clear, it is less obvious just what form the revised *extended* belief set, E_A^*, should take. At first glance, one might think that the conditional and other subjective sentences contained in E_A^* should (by analogy) be just those true at each world in $\min(M, A)$. However, reflection on the truth conditions for $\xrightarrow{\mathrm{KB}}$ shows this is impossible. A sentence $\alpha \xrightarrow{\mathrm{KB}} \beta$ is true in model M just when it is true at each world in M. This remark applies to belief sentences $\boxtimes\alpha$ and sentences of entrenchment or plausibility $\alpha \leq_E \beta$ as well. The definitions of these connectives are *global*. The features of the world at which such sentences are evaluated is immaterial. Their truth is determined by the ordering of relative plausibility of worlds. As in standard epistemic logics, the truth of sentence $\boxtimes A$ at a world w is not determined by w itself, but by the set of epistemically possible worlds relative to w (typically, those *accessible* from w), of which w may or may not be a member. A K-revision model is suitable for a fixed belief set only (indeed, a fixed extended belief set). It represents one ordering, hence one set of conditionals and one set of beliefs. If $A \in K$ then $\boxtimes A$ is true at *every world* in the structure.

We can speak of K_A^* being represented by the set $\min(M, A)$ because it contains only new objective beliefs. This view is tenable because the truth of such sentences is determined by this set alone, regardless of the ordering of other worlds. Notice, however, that each conditional, belief sentence, entrenchment sentence, etc. true in the extended set E associated with M is true at each world in $\min(M, A)$. Thus,

$$\{\alpha \in \mathbf{L}_B : M \models_w \alpha \text{ for each } w \in \min(M, A)\}$$

is *not* the correct extended belief set E^*_A. If it were, one could never give up conditionals or other subjective beliefs, though the objective component K can change drastically. When E is revised by A, the CO*-model used to represent E becomes inadequate. It is an E-revision model and the representation of E^*_A requires (of course) and E^*_A-revision model.

What are the natural requirements on this new model? To reiterate, we are assuming a revision model M that captures extended belief set E with objective component $K \subseteq E$. When a propositional revision A is received, we want the revision function $*$ to map model M into a new model M^*_A that captures the revised belief set E^*_A. Clearly, K^*_A is uniquely determined by M (in particular, by $\min(M, A)$). Naturally we insist that $K^*_A \subseteq E^*_A$ and that K^*_A form the *entire* objective component of E^*_A. This simply means that M^*_A should be a K^*_A-revision model, or that K^*_A is *only known* in M^*_A. The minimal cluster of worlds in M^*_A should be exactly $\min(M, A)$. This is illustrated in Figure 2. Let us dub this constraint the *Basic Requirement* on revision functions as applied to models.

The Basic Requirement: If M is a K-revision model determining revision function $*$, then the revision model M^*_A must be such that $\min(M^*_A, \top) = \min(M, A)$.

In fact, from a purely logical perspective, this is probably all we want to say about M^*_A. If one changes an objective belief, it is impossible in general to predict what becomes of one's conditionals. This model of iterated revision is captured by Gärdenfors's (1988) *belief revision systems*, although not in this semantic fashion. Such a system consists of a set $\mathbf{K}$ of belief sets and a revision function $*$ that maps $\langle K, A\rangle$, where $K \in \mathbf{K}$ and $A \in \mathbf{L}_{CPL}$, into $K^*_A \in \mathbf{K}$. The function $*$ must satisfy the AGM postulates, but the behavior of $*$ on the extended belief set E associated (by the Ramsey test) with K is left unspecified. Only postulates (R7) and (R8) constrain how revision of K^*_A should take place, and these constraints are quite mild. This model provides no guidance as to what conditionals in E should be accepted or rejected in E^*_A. Put another way, no hints are provided on how to revise K^*_A, given its "origination" as a revision of K by A. The ordering of worlds in the nebulous region of Figure 2 (b) is

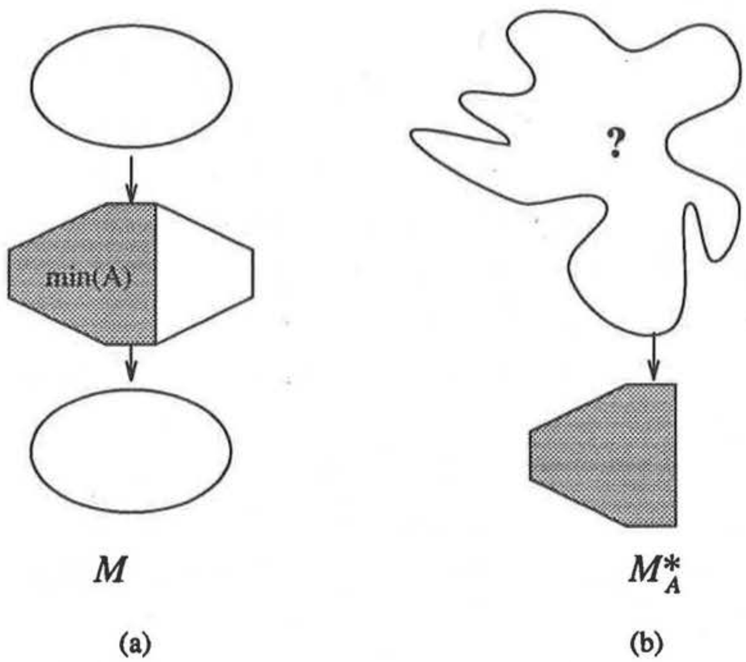


Figure 2: General constraints on the mapping to a new revision model

completely unspecified.

This model possesses two unattractive features, one logical and one pragmatic. First, states of belief are distinguished solely by their objective component. For any K, the extended belief set E associated with K (via the Ramsey test) is fixed, since K_A^* is fixed. There can be no two distinct extended belief sets that share the same objective part. In other words, there can only be one way of revising a belief set. This is certainly an unnatural restriction for which there is no logical justification. Notice that our mapping, though not yet precisely defined, applies to models, hence to *extended* sets, so this situation is rectified. We take the view that a belief state is not uniquely determined by its objective component, that conditionals or information of entrenchment forms an integral part of an agent's belief state. An agent can be in two different belief states, but have the same propositional beliefs in each and differ on its accepted conditionals. The CO*-models representing each belief state could agree on the minimal cluster, but differ on the rest of the model.[9] This is a natural extension of revision systems, but it is also quite straightforward. The logical "characterization" of such a system is completely captured by Figure 2.

The second criticism is of a more pragmatic nature. Everything above points to a mapping of M to M_A^* satisfying one condition, that the objective belief set associated with M_A^* be precisely the set K_A^* determined by M. As Figure 2 illustrates, just about all of the ordering information, capturing an agent's conditional beliefs and judgements of entrenchment, is (potentially) lost in this mapping. There is something unsatisfying about this model. The ordering relation R is intended to reflect the informational content or importance of beliefs. When certain beliefs must be given up, it seems very natural to try to keep not only important beliefs, but as much of the *ordering* a possible. An revision should not generally change one's opinion of the relative importance of most sentences.

Instead of arbitrary mappings from M to M_A^*, we will propose a class of natural mappings that preserve as much ordering information as can be expected. This determines the class of *natural revision functions*, that tend to preserve the entrenchment information and conditional beliefs found in an extended belief set. It is important to note that the model we propose is *not*

[9]Indeed, both are acceptable, and an agent should be able to move from one to the other by revising only its conditional beliefs. We do not address the problem of revising by conditionals in this paper, but see (Boutilier 1992c).

completely general, for it permits only a subset of those revision functions (on extended sets) allowed by the arbitrary mappings described above. These arbitrary mappings are explored in (Boutilier 1992c). However, it is a very natural subset, suitable for determining the result of propositional revision sequences, or the truth of right-nested conditionals, when the general model has little to offer.

3 Paranoid Revision

Before proceeding to changes to arbitrary revision models, it will prove instructive to examine properties of a subclass of these models. The logic CO* characterizes a very natural class of structures, those consisting of a totally-order set of clusters of worlds. This is the same class of models characterized by the modal logic S4.3. Somehow, even more basic than this partitioning of worlds into an indefinite number of classes is the notion of a bipartition. Instead of distinguishing "degrees" of plausibility, we can simply have two categories of worlds: plausible and implausible. Plausible worlds are those consistent with an agent's beliefs, while implausible worlds falsify at least one belief. This is a special case of a CO*-structure in which the (non-minimal) clusters of implausible worlds are collapsed into one cluster, thus eliminating "degrees" of implausibility. This observation has lead Boutilier (1992b) to show that CO* generalizes autoepistemic logic (Moore 1985; Levesque 1990).

The idea of a bipartition is quite natural and is implicit in the possible worlds semantics of most epistemic logics, though only the plausible worlds are explicitly represented in most instances (Halpern and Moses 1985; McArthur 1988). This single cluster approach typically leads to the modal logics S5 or weak S5. In the standard model framework, the specialization of S4.3 where two clusters are explicitly represented was studied by Segerberg (1971) and is dubbed S4F (or S4.3.2). Its applicability to knowledge representation has been demonstrated by Truszczyński (1991), who used it to capture the semantics of Reiter's (1980) default logic, and by Schwarts and Truszczyński (1992) in their characterization of minimal knowledge.

Let us denote by CO.2* the logic obtained by restricting attention to CO*-structures with exactly two clusters.[10] Suppose our revision structures for any belief set K (or extended set

[10] We can think of this as the extension of S4F with the second modality $\overleftarrow{\Box}$. In the full paper, we provide a

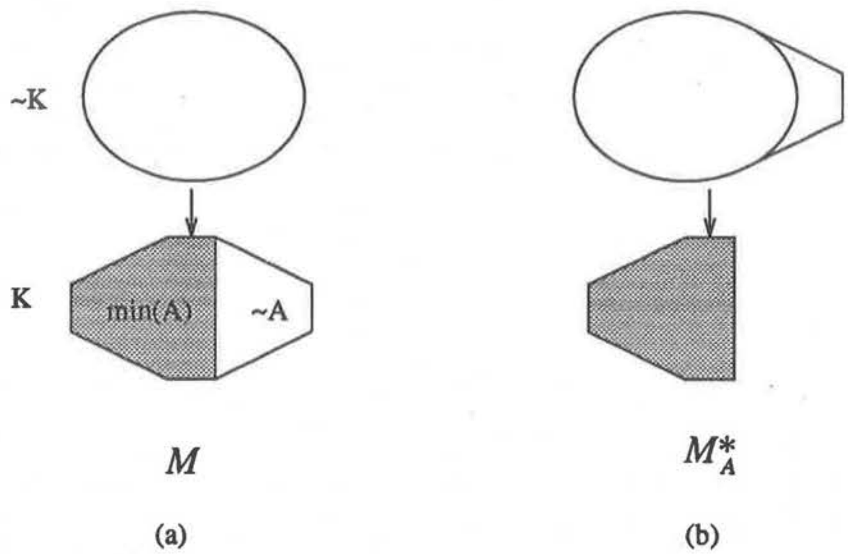


Figure 3: Paranoid revision with a consistent revision.

E) are drawn from this class. The behavior of revision in this case is quite interesting and remarkably simple to characterize.

Consider the K-revision model M in Figure 3 (a), where sentence A is consistent with belief set K. When K is revised by A, as discussed previously, M must be mapped to a new model M^*_A, where $\min(M, A)$ (the shaded region) forms the minimal cluster in M^*_A. Since $K \nvdash \neg A$, we have $\min(M, A) \subseteq \|K\|$. Clearly, the worlds in $\|K\| - \min(M, A)$, plausible in M, must become implausible in M^*_A, while the worlds in $\|\neg K\|$ must remain implausible. However, since we are discussing CO.2* models, there can only be one degree of implausibility, and M^*_A must have the structure illustrated in Figure 3 (b). Given the Basic Requirement on $*$, we see that the logic CO.2* uniquely determines a new revision model M^*_A.

More interesting is the case of an inconsistent revision, where $K \vdash \neg A$. This situation is depicted in Figure 4 (a). As before, the new revision model is uniquely determined by the revision A, but it has the unusual property that $K^*_A = Cn(A)$. To see this, notice that no A-worlds exist in the minimal cluster of M (since $K \vdash \neg A$). This implies that *all logically possible* A-worlds are situated in the non-minimal cluster (the shaded region), therefore,

simple axiomatization and completeness proof.

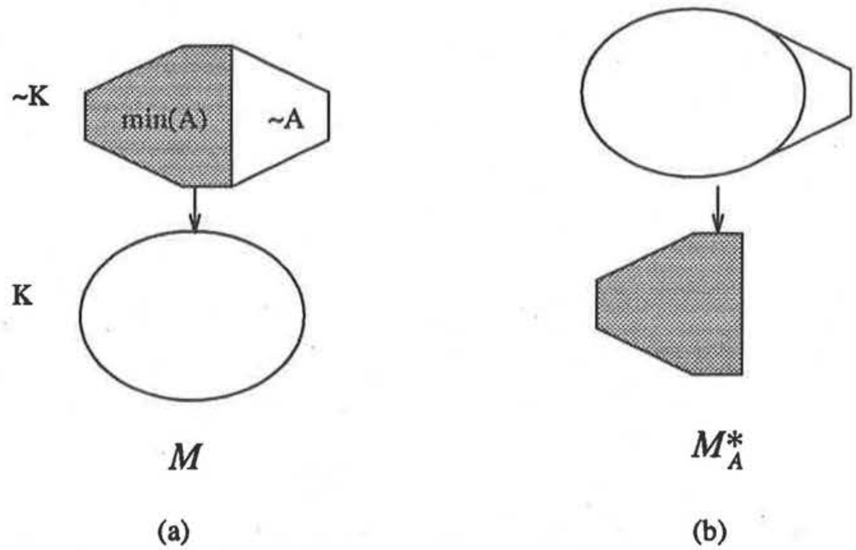


Figure 4: Paranoid revision with an inconsistent revision.

$K^*_A = Cn(A)$. In other words, any inconsistent revision results in the abandonment of all beliefs and adopting only logical consequences of the update. We call this *paranoid revision* since it seems to suggest an agent that, upon learning something inconsistent with its beliefs, gives up all of its beliefs for fear that this inconsistency could lurk anywhere in the original belief set.

Definition 4 A *paranoid K-revision model* is any K-revision model consisting of exactly two clusters. (If $\|K\| = W$ we assume an empty top cluster.)

Proposition 5 *Let $M = \langle W, R, \varphi\rangle$ be a paranoid revision model for K. Then wRv iff: a) $M \models_w \neg\alpha$ for some $\alpha \in K$; or b) $M \models_v \alpha$ for each $\alpha \in K$.*

Thus there is a unique paranoid revision model M for K.

Definition 5 Let $M = \langle W, R, \varphi\rangle$ be a paranoid revision model. For any $A \in \mathbf{L}_{CPL}$, the revised model $M^*_A = \langle W, R', \varphi\rangle$ is such that $wR'v$ iff: a) $v \in \min(M, A)$; or b) $w \notin \min(M, A)$.

Definition 6 Let M be a paranoid revision model for K. The *paranoid revision function* associated with M is $*$, defined for all $A \in \mathbf{L}_{CPL}$ as follows:

$$K_A^* = \{B \in \mathbf{L}_{CPL} : M_A^* \models \boxtimes B\}$$

Proposition 6 *Let M be a paranoid revision model for K, $*^{M}$ be the AGM revision function determined by M (as given by Definition 3), and $*$ be the paranoid revision function determined by M. For any $A \in \mathbf{L}_{CPL}$, we have $K_A^{*^{M}} = K_A^*$.*

Notice that the paranoid revision function $*$ is an AGM operator since paranoid revision models form a subclass of CO*-models. Of course, these revision functions respect the definition of the general revision functions on CO*-models (Proposition 6). We define the revised belief set in terms of the revised model in order to investigate the properties of non-objective sentences in the revised belief set. As we've seen, the earlier Definition 3 can only capture the newly accepted objective sentences.

The class of paranoid revision functions is a rather natural class of AGM revision operators. In fact, it appears to be the "simplest" possible model of AGM revision. If we were to reduce our class of models to those with one cluster, we could not distinguish plausible from implausible worlds. We might make all worlds plausible and capture a fixed belief set K (as in standard epistemic logics), but we would be unable to model a revised state of belief since a single cluster provides no way to separate the newly "implausible" worlds from this model. We could map to a "smaller" model with a single cluster. But in no circumstance could we characterize an inconsistent revision by A as a move to the set of minimal A-worlds. If A is inconsistent with K, there can be no A-worlds in the model, and revision by A leads to a contradiction, generally in violation of the AGM postulates. For this reason, we claim that CO.2* is the "maximal" logic that can be used for revision and satisfy the AGM postulates. It represents an upper bound on the restrictions that can be placed on a revision function of the AGM variety.

Sequences of propositional updates in this paranoid framework are very easy to characterize. Suppose we have a belief set K, that we assume to be finitely specified by *KB*, and a revision sequence $A_1, A_2, \ldots A_n$. For each revision A_i that remains consistent with the previous revisions (along with *KB*), the resulting belief set is captured by simply conjoining A_i to the knowledge base (already "revised" by conjoining the previous updates).

Proposition 7 *Let M be a paranoid revision model for K and $A_1, \ldots A_i$ be a revision sequence such that $K \not\models \neg(A_1 \wedge \cdots A_i)$. Then $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_i} = Cn(K \cup \{A_1, \cdots, A_i\})$, where $*$ is the paranoid revision function.*

Clearly, if K is finitely specifiable as *KB*, the resulting belief set is captured by the sentence $KB \wedge A_1 \wedge \cdots A_i$. Furthermore, nothing is lost by considering this single "large" revision.

Proposition 8 *Let M be a paranoid revision model for K and $A_1, \ldots A_i$ be a revision sequence such that $K \not\models \neg(A_1 \wedge \cdots A_i)$. Then $((M^*_{A_1})^*_{A_2} \cdots)^*_{A_i} = M^*_{A_1 \wedge \cdots A_i}$. Consequently, $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_i} = K^*_{A_1 \wedge \cdots A_i}$.*

Given a sequence of consistent revisions, one may chose to revise K by each in turn, or simply revise K by their conjunction. While the resulting objective belief set is the same in each case, more importantly the resulting revision model is identical. Thus, future revisions of the belief set need not worry about the order in which the revisions were applied or the strategy chosen (we will see shortly that *arbitrary* revision models do not enjoy this property).

Once an inconsistent revision is received, all previously held beliefs are discarded. All that remains is the set of consequences of the update (we assume A is logically satisfiable).

Proposition 9 *Let $M = \langle W, R, \varphi \rangle$ be a paranoid revision model for K, where $K \vdash \neg A$. If $*$ is the paranoid revision function then $M^*_A = \langle W, R', \varphi \rangle$ where $vR'w$ iff $M \models_w A$ or $M \models_v \neg A$. Consequently, $K^*_A = Cn(\{A\})$.*

Paranoid revision of this new belief set proceeds as usual, simply "adding" consistent revisions to the knowledge base and "starting over" when an inconsistent revision is encountered.

Theorem 10 *Let M be a paranoid revision model for K, let $A_1, \ldots A_n$ be a sequence of revisions, and let $*$ be the paranoid revision function. If $K \not\vdash \neg(A_1 \wedge \cdots A_n)$ then $((M^*_{A_1})^*_{A_2} \cdots)^*_{A_n}$ is the (unique) paranoid revision model for*

$$K' = ((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n} = Cn(K \cup \{A_1, \cdots, A_i\})$$

If $K \vdash \neg(A_1 \wedge \cdots A_n)$ *then* $((M^*_{A_1})^*_{A_2} \cdots)^*_{A_n}$ *is the (unique) paranoid revision model for*

$$K' = ((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n} = Cn(\{A_i, A_{i+1}, \cdots, A_n\})$$

where $1 \leq i \leq n$ *is such that* $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_{i-1}} \vdash \neg A_i$ *and* $A_i \not\vdash \neg(A_{i+1} \wedge \cdots A_n)$.

Thus we have a complete characterization of paranoid revision sequences that translates into an obvious algorithm for obliging such a revision sequence, that of "adding" sentences and "starting over". Notice that when a sequence contains an inconsistent revision, A_i in the theorem above is the *last* inconsistent revision in the sequence, the last revision that requires all previous beliefs be abandoned.

Paranoid revision is quite simple conceptually and easy to describe computationally. This is due to the fact that an objective belief set K uniquely determines its paranoid revision model. Regardless of how the belief set K is created, no matter what revision sequence lead to its acceptance, its revision model cannot vary. Its "history" is irrelevant. Indeed, since its revision model is fixed, K uniquely determines the extended belief set E of which it is part. In particular, the objective part K of an extended belief set determines all conditional beliefs. If $K \not\vdash \neg A$ then $A \xrightarrow{\text{KB}} B \in E$ iff $K \cup \{A\} \vdash B$. If $K \vdash \neg A$ then $A \xrightarrow{\text{KB}} B \in E$ iff $A \vdash B$. In moving from K' to K via some revision, there is no attempt to preserve the conditional beliefs in E' (where $K' \subseteq E'$). The ordering of worlds captured by E' is completely irrelevant to the ordering of worlds determined by the extended belief set E, as illustrated by Figures 3 and 4.

It is not hard to see that paranoid revision of an objective belief set K is precisely the *full meet revision* of K, as described in (Alchourrón, Gärdenfors and Makinson 1985; Gärdenfors 1988). Full meet revision is specified by considering all maximal subsets of K consistent with the new fact A, and adding A to their intersection. A well-known property of such a model of revision is that $K^*_A = Cn(A)$ whenever $K \vdash \neg A$. Thus, the semantics of paranoid revision provide a possible worlds characterization of full meet revision within our modal framework.[11]

The manner in which information is lost during paranoid revision is quite straightforward. For example, given a consistent revision A of K, the "newly implausible" $K \cup \{\neg A\}$-worlds

[11] As an aside, we note that *maxi-choice revision* corresponds to the requirement that each non-minimal cluster in our model contain exactly one world.

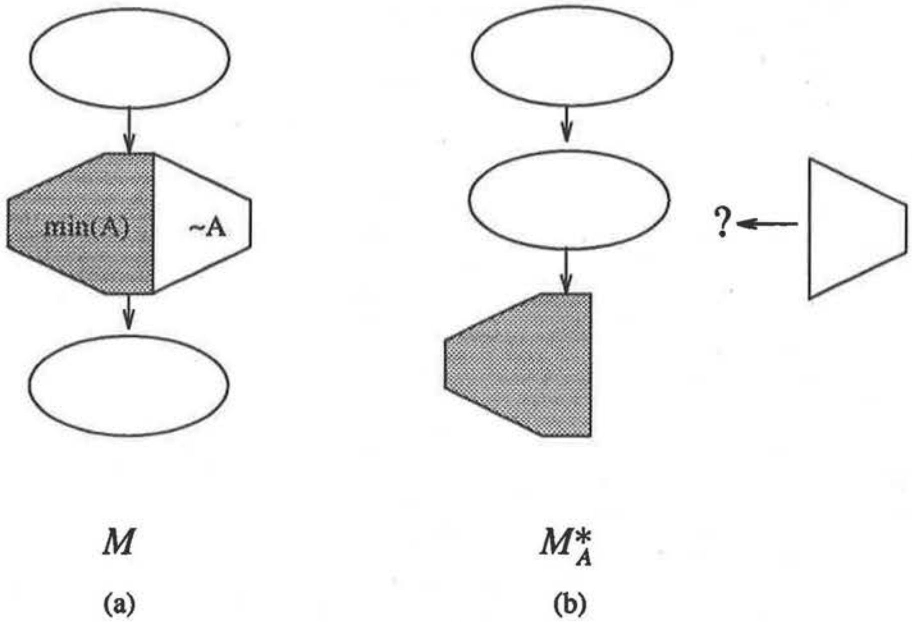
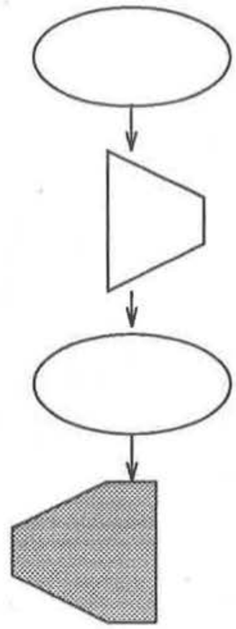


Figure 5: How to revise a three cluster model?

are simply moved into the cluster of "previously implausible" $\neg K$-worlds. There is no other choice if we are to move to a K_A^*-revision model and remain within the class of two-clustered models. In order to extend this idea, we might try to use this technique for models with any *fixed* number of clusters. However, even once we allow three clusters, it becomes clear that revised models are not uniquely determined, and furthermore there are no compelling principles for choosing from the alternatives.

Consider the three-cluster model M in Figure 5. Since M_A^* must be a revision model for K_A^*, clearly the shaded area of minimal A-worlds must form the lowest cluster in M_A^*. This again is our Basic requirement. What becomes of the partial cluster that was attached to $\min(M, A)$? If we are to remain within the class of models with three clusters, the two remaining clusters and this third partial cluster must be merged into two clusters. There will generally be an infinite number of ways to do this, but none seem very compelling. Fusing the partial cluster an either of the other other clusters seems rather arbitrary, for revising by A should intuitively not make the worlds in the partial cluster either more or less plausible with respect to these two clusters. We could chose to redistribute worlds from the partial cluster



M^*_A

Figure 6: The *natural* revision of a three cluster model.

among both clusters, coalesce the two "whole" clusters, or even break up the two clusters.

The most natural choice seems to involve none of these options. Rather, we ought to keep things as they are by leaving the clusters and the partial cluster intact, with the same relative ordering. This is illustrated in Figure 6. This moves us from a three-cluster model to a four-cluster model, and is rather different in nature from the paranoid revision model. However, while the paranoid model left us no options in choosing a revised model M^*_A, the same technique applied to three-clustered models leaves too many choices. The most natural technique will allow a revision of a model to (typically) *increase* the number of clusters. There is no compelling reason to adhere to some fixed number of clusters. Naturally, this idea applies to arbitrary revision structures, not just those with three clusters.

4 Natural Revision Functions

4.1 The Semantics of Revisions

As in the previous section, we must define a mapping of a K-revision model M to a new model M_A^*. We now remove the restriction to paranoid revision models and consider arbitrary (well-founded) CO*-models. Our new mapping must satisfy the Basic Requirement that $\|K_A^*\|$, the minimal cluster in M_A^*, be identical to $\min(M, A)$. This "preserves information" in the AGM sense, ensuring that the most entrenched propositional beliefs in K are retained when A is accommodated. However, we also want to preserve as much conditional and entrenchment information as possible. The principle of informational economy must be applied to non-objective beliefs as well.

The conditionals accepted by an agent are determined by its ordering of plausibility. If we insist that revision preserve as much of this ordering as possible, then, for the most part, the relative entrenchment and plausibility of sentences (hence conditional beliefs) will remain intact. Let $M = \langle W, R, \varphi \rangle$ be the revision model reflecting some extended belief set E. Given a propositional revision A of E (or the associated K), we must find a revision model $M_A^* = \langle W, R', \varphi \rangle$ such that R' reflects the minimal mutilation of R.

If $w \in \min(M, A)$, by the Basic Requirement w must be minimal in R', and these must be the only minimal worlds in R'. For any such w the relationships $wR'v$ and $vR'w$ are completely determined by membership of v in $\min(M, A)$, independently of their relationship in R. Figure 2 illustrates this. For w, v not in $\min(M, A)$, this picture leaves wRv completely unspecified. If R is to be left intact to the largest possible extent then the most compelling specification is to insist that $vR'w$ iff vRw. This has the effect of leaving R unaltered except as indisputably required by the Basic Requirement. Such a move is illustrated in Figure 6. We dub such a mapping on revision models the *natural revision operator*, and no describe the revision function it induces on the associated belief and extended belief sets.

Definition 7 Let $M = \langle W, R, \varphi \rangle$ be a revision model. The *natural revision operator* $*$ maps M into M_A^*, for any $A \in \mathbf{L}_{CPL}$, where $M_A^* = \langle W, R', \varphi \rangle$, and: a) if $v \in \min(M, A)$ then $wR'v$ for all $w \in W$ and $vR'w$ iff $w \in \min(M, A)$; and b) if $v, w \notin \min(M, A)$ then $vR'w$ iff vRw.

Definition 8 Let E be the extended belief set associated with the revision model M. The *natural revision function* associated with M is $*$, defined for all $A \in \mathbf{L}_{CPL}$ as follows:

$$E^*_A = \{B \in \mathbf{L}_B : M^*_A \models \boxtimes B\}$$

Let $K \subseteq E$ be the objective component of E. The natural revision of K, denoted K^*_A, is the restriction of E^*_A to $\mathbf{L}_{CPL}$; that is

$$K^*_A = \{B \in \mathbf{L}_{CPL} : M^*_A \models \boxtimes B\}$$

Clearly, the natural revision function is simply the AGM operator determined by M, when restricted to K.

Proposition 11 *Let M be a revision model for K, $*^{\mathrm{M}}$ be the AGM revision function determined by M (as given by Definition 3), and $*$ be the natural revision function determined by M. For any $A \in \mathbf{L}_{CPL}$, we have $K^{*^{\mathrm{M}}}_A = K^*_A$.*

Notice that this extends the AGM model of revision. The revised extended set E^*_A is defined using the updated revision model M^*_A and incorporates non-objective beliefs, such as conditionals, statements of entrenchment and plausibility, and nested belief sentences. Had we simply defined E^*_A to be those sentences true in M at the minimal A-worlds, we would have run into the problem discussed in Section 2, namely the fact that a model M can only model a fixed set of beliefs and conditionals.

If we are to extend the Ramsey test to include nested conditionals, the truth conditions for statements $A \xrightarrow{\text{KB}} \beta$ must be recast in this framework. For M to satisfy $A \xrightarrow{\text{KB}} \beta$, we must have $\beta \in E^*_A$ for the natural revision function $*$. For $\beta \in \mathbf{L}_{CPL}$, given Proposition 11, these truth conditions will be identical to those provided in Section 2. Thus, our new truth conditions for $\xrightarrow{\text{KB}}$ based on the Ramsey test will form a "conservative extension" of the old definition. However, for arbitrary $\beta \in \mathbf{L}_B$, especially sentences like $B \xrightarrow{\text{KB}} C$, the definition of Section 2 using modal operators is inadequate since it refers to truth at worlds in $\min(M, A)$. To evaluate $A \xrightarrow{\text{KB}} (B \xrightarrow{\text{KB}} C)$ we must test $B \xrightarrow{\text{KB}} C$ in M^*_A, not at $\min(M, A)$. This means we cannot equate the truth of $A \xrightarrow{\text{KB}} \beta$ with the truth of the modal sentence used to define $A \xrightarrow{\text{KB}} \beta$ for non-objective β. Therefore the connective $\xrightarrow{\text{KB}}$ must be introduced as a primitive connective. We let the *conditional language* $\mathbf{L}_C$ be the extension of $\mathbf{L}_B$ with the connective $\xrightarrow{\text{KB}}$ (no longer defined). Its truth conditions are as follows:

Definition 9 Let $M = \langle W, R, \varphi \rangle$ be a CO*-model, $A \in \mathbf{L}_{CPL}$, and $B \in \mathbf{L}_C$. M satisfies $A \xrightarrow{\text{KB}} B$ at $w \in W$ (written $M \models_w A \xrightarrow{\text{KB}} B$) iff $M_A^* \models \boxtimes B$.

Now we have a conditional connective whose truth conditions are specified directly by the Ramsey test. While this certainly provides us with a new logic, requiring a new axiomatization, we will see in Section 5 that these truth conditions can, in fact, be captured in the bimodal language using only the defined version of the connective $\xrightarrow{\text{KB}}$.

Notice that the truth of $A \xrightarrow{\text{KB}} B$ is unspecified for $A \notin \mathbf{L}_{CPL}$. Natural revision functions are suitable only for sequences of *propositional* updates. The nesting of conditionals sanctioned in a meaningful way in this framework is *right-nesting*, for instance, $A \xrightarrow{\text{KB}} (B \xrightarrow{\text{KB}} C)$ where $A, B, C \in \mathbf{L}_{CPL}$. A sentence $(A \xrightarrow{\text{KB}} B) \xrightarrow{\text{KB}} C$ has an unspecified truth value for it asks if C is believed when a knowledge base is revised to include $A \xrightarrow{\text{KB}} B$. This framework does not specify how to revise a knowledge base with non-objective sentences, though this problem is addressed in (Boutilier 1992c).

4.2 Properties of Single Revisions

In this section, we investigate some of the properties of single changes to a revision model or belief set using the natural revision function $*$. We assume throughout some revision model M capturing the extended belief set E and the belief set $K \subseteq E$. When M (or K or E) is revised by A the properties of K_A^* are obvious: $K_A^* = K_A^{*^{\mathrm{M}}}$, where $*^{\mathrm{M}}$ is the AGM operator determined by M. Of more interest is the characterization of of the new extended set E_A^*. Since minimal mutilation of the ordering relation R is intended to preserve as many conditional beliefs as possible, we must determine precisely which conditionals in E remain in E_A^* and which are sacrificed. Since we are interested in single revisions, we restrict our attention (for the time being) to *simple conditionals* of the form $A \xrightarrow{\text{KB}} B$ where both A and B are objective.

One important property to note is the following:

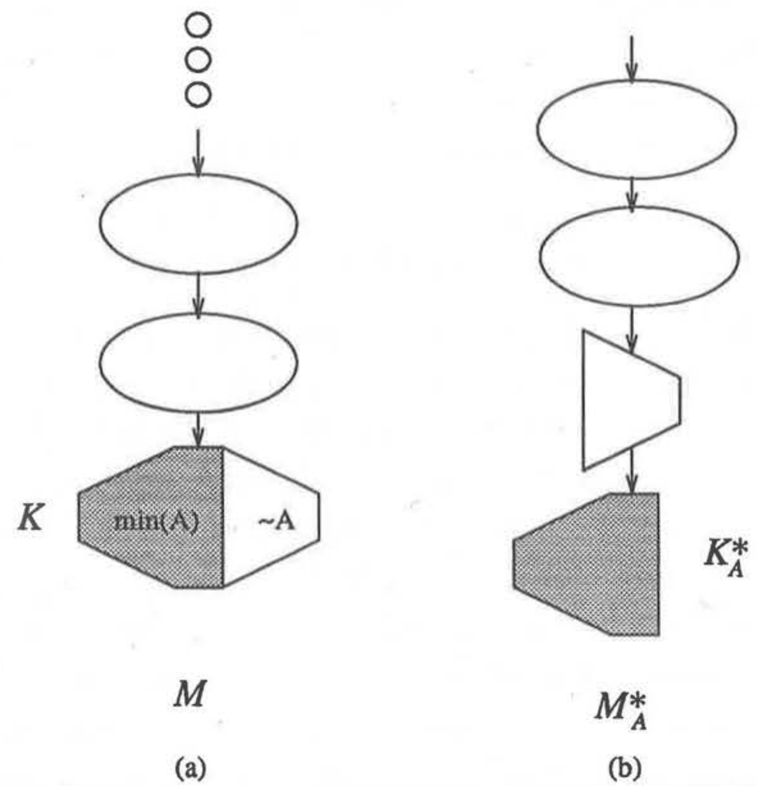


Figure 7: Natural revision with a consistent revision.

Proposition 12 *Let M be a K-revision model where $A \in K$. Then $M_A^* = M$.*

Updating by a sentence already in a belief set not only causes no change in the belief set K, as required by the AGM postulates, but also leaves the revision model M intact. More generally, consider the two scenarios, illustrated by Figures 7 and 8, that might arise when M is revised by A. Figure 7 shows the situation where A is a consistent revision, $K \nvdash \neg A$, while Figure 8 demonstrates the behavior of an inconsistent revision, $K \vdash \neg A$. We are interested in those simple conditionals $B \xrightarrow{\text{KB}} C$ that are true in the model M_A^*. In either of these two scenarios there are two cases to consider: $\neg B \notin K_A^*$ and $\neg B \in K_A^*$.

Consider the first situation where $\neg B \notin K_A^*$. This can be phrased equivalently as $M_A^* \not\models \boxtimes\neg B$, as $\neg(A \xrightarrow{\text{KB}} \neg B) \in E$, or as $M \models \neg(A \xrightarrow{\text{KB}} \neg B)$, and simply means that $\neg B$ is not accepted in the new belief state represented by K_A^*. In other words, B is consistent with K_A^*.

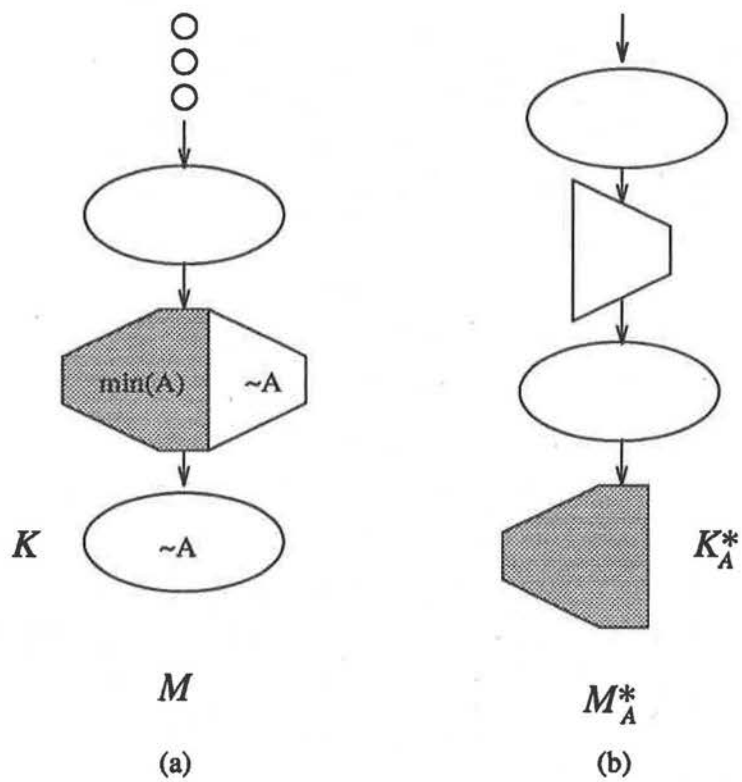


Figure 8: Natural revision with an inconsistent revision.

This is true exactly when there is some B-world in the minimal cluster of M_A^* (the shaded region of the figures); that is, some B-world is contained in $\min(M, A)$. Now, $M_A^* \models B \xrightarrow{\text{KB}} C$ just in case the set $\min(M_A^*, B)$ contains only C-worlds. Since $\neg B \notin K_A^*$, the set $\min(M_A^*, B)$ is a subset of the minimal cluster of M_A^*. Clearly then $\min(M_A^*, B) = \min(M, A \wedge B)$ and $M_A^* \models B \xrightarrow{\text{KB}} C$ iff $M \models A \wedge B \xrightarrow{\text{KB}} C$. Therefore, whenever $K_A^* \not\vdash \neg B$, a conditional $B \xrightarrow{\text{KB}} C$ is in E_A^* iff $A \wedge B \xrightarrow{\text{KB}} C$ is in E. The status of $B \xrightarrow{\text{KB}} C$ in E has no bearing on its acceptance or rejection in E_A^*. This can be explained as follows. $B \xrightarrow{\text{KB}} C$ is in E just in case the minimal B-worlds in M satisfy C. When E is revised by A, the minimal A-worlds become most plausible (in M_A^*) and "carry along" with them certain B-worlds. If these are not the minimal B-worlds in M, the status of $B \xrightarrow{\text{KB}} C$ in E has no influence on the acceptance of $B \xrightarrow{\text{KB}} C$ in E_A^*, since these are no longer the most plausible B-worlds. That honor is now conferred upon those B-worlds in $\min(M, A)$. Notice that this behavior is exactly in accordance with the AGM postulates (R7) and (R8).

The second situation arises when $\neg B \in K_A^*$. This means $M_A^* \models \boxtimes\neg B$, or $A \xrightarrow{\text{KB}} \neg B \in E$, or $M \models A \xrightarrow{\text{KB}} \neg B$. When K (or E) is revised by A, $\neg B$ is in the resulting belief set. This is true exactly when there are no B-worlds in the minimal cluster of M_A^* (the shaded region of the figures); that is, no B-world is contained in $\min(M, A)$. Now, $M_A^* \models B \xrightarrow{\text{KB}} C$ just in case the set $\min(M_A^*, B)$ contains only C-worlds. Since $\neg B \in K_A^*$, the set $\min(M_A^*, B)$ is not contained in the minimal cluster of M_A^*. However, all worlds outside the minimal cluster stand in exactly the same relation as they do in M. Therefore $\min(M_A^*, B) = \min(M, B)$ and it follows that $M_A^* \models B \xrightarrow{\text{KB}} C$ iff $M \models B \xrightarrow{\text{KB}} C$. For conditionals $B \xrightarrow{\text{KB}} C$ whose antecedents are not made plausible by the acceptance of A (i.e., $K_A^* \not\vdash \neg B$), $B \xrightarrow{\text{KB}} C$ is in E_A^* iff $B \xrightarrow{\text{KB}} C$ is in E. Since nothing forces the conditional to be abandoned when A is accepted, it is retained.

We can summarize these considerations in the following theorem and equivalent corollaries.

Theorem 13 *Let M be a revision model, let $*$ be the natural revision operator and let $A, B, C \in \mathbf{L}_{CPL}$.*

a) If $M_A^* \models \boxtimes\neg B$ then $M_A^* \models B \xrightarrow{\text{KB}} C$ iff $M \models B \xrightarrow{\text{KB}} C$.

b) If $M_A^* \not\models \boxtimes\neg B$ then $M_A^* \models B \xrightarrow{\text{KB}} C$ iff $M \models A \wedge B \xrightarrow{\text{KB}} C$.

Corollary 14 *Let M be a revision model with associated belief set K, and let $*$ be the natural* *revision function.*

a) If $K_A^* \vdash \neg B$ *then* $C \in (K_A^*)_B^*$ *iff* $C \in K_B^*$.

b) If $K_A^* \nvdash \neg B$ *then* $C \in (K_A^*)_B^*$ *iff* $C \in K_{A \wedge B}^*$.

Corollary 15 *Let* M *be a revision model with associated extended belief set* E*, and let* $*$ *be the natural revision function.*

a) If $A \xrightarrow{\text{KB}} \neg B \in E$ *then* $B \xrightarrow{\text{KB}} C \in E_A^*$ *iff* $B \xrightarrow{\text{KB}} C \in E$ *iff* $A \xrightarrow{\text{KB}} (B \xrightarrow{\text{KB}} C) \in E$.

b) If $A \xrightarrow{\text{KB}} \neg B \notin E$ *then* $B \xrightarrow{\text{KB}} C \in E_A^*$ *iff* $A \wedge B \xrightarrow{\text{KB}} C \in E$ *iff* $A \xrightarrow{\text{KB}} (B \xrightarrow{\text{KB}} C) \in E$.

These results precisely characterize when conditionals will be preserved in a revised extended belief set. Furthermore, each of these results show that the sentences accepted in the new revision model or belief state can be determined by appeal to the original model or belief state. Theorem 13 shows that the conditional belief set captured by M_A^* can be determined by the conditional beliefs of M. Furthermore, it demonstrates that natural revision preserves as much conditional information in the revised belief set as is consistent with the AGM postulates. It is clear that if B is consistent with the revised belief set K_A^*, then the truth of $B \xrightarrow{\text{KB}} C$ in the new extended set E_A^* cannot be influenced by its truth in E. It's acceptance or rejection is dictated by postulate (R4). This is reflected in clause b) of the theorem. However, as indicated by clause a), the remaining set of conditionals (or negated conditionals) in E_A^* coincides precisely with the conditional information in the original extended set E. Except for those that must be given up because of (R4), every conditional sentence is retained.

Corollary 14 shows that the sequence of two revisions applied to K can be reduced to a single revision, requiring no iterated revision, and that the test to establish which condition holds also requires no iterated revision. Similarly, Corollary 15 shows that the revised extended belief set E_A^* and the nested conditionals in E can be captured by the simple, unnested conditionals in E. These properties will play a vital role in our characterization of revision sequences.

4.3 Revision Sequences

The objective belief set K_A^* formed by revising K with A relies only on one application of $*$ to the original belief set K. Somewhat surprisingly, the simple conditionals in E_A^* can also be discovered by referring only to applications of $*$ to K. As the results above indicate, to ask if $B \xrightarrow{\text{KB}} C \in E^*_A$, one first asks if $\neg B$ is in K^*_A. If this is true, then asking if $B \xrightarrow{\text{KB}} C \in E^*_A$ is equivalent to asking if $C \in K^*_B$. If this is false, then asking if $B \xrightarrow{\text{KB}} C \in E^*_A$ is equivalent to asking if $C \in K^*_{A \wedge B}$. A "hypothetical" revision of K by B or by $A \wedge B$ is sufficient to determine the status of $B \xrightarrow{\text{KB}} C$ in E^*_A (or equivalently, to determine the status of $A \xrightarrow{\text{KB}} (B \xrightarrow{\text{KB}} C)$ in E).

Suppose we have a revision sequence, $A_1, \ldots A_n$, to be applied to E, where $K \subseteq E$. We can clearly use a single revision to verify whether some simple conditional $B \xrightarrow{\text{KB}} C$ is in $E^*_{A_1}$, or if $M^*_{A_1} \models B \xrightarrow{\text{KB}} C$. However, if we replace B by A_2 we see that these conditionals determine precisely the beliefs obtained when $E^*_{A_1}$ is revised by A_2, that is, $(E^*_{A_1})^*_{A_2}$. If single revisions applied to K can establish the content of the belief set obtained by an iterated revision of two levels, there seems no reason single (uniterated) revisions of K cannot be used to decide the outcome of arbitrary revision sequences. In this section we will examine the properties of such sequences and show how they may be reduced to single uniterated revisions, or unnested conditionals.

4.3.1 Order Dependence

The simplest true "sequence" of revisions consists of two elements A and B. An important property of the natural revision function applied to revision sequences is its *order dependence*. In general, the sets $(K^*_A)^*_B$ and $(K^*_B)^*_A$ will differ. To see this, consider a revision model M such that $M \not\models A \xrightarrow{\text{KB}} \neg B$, but $M \models B \xrightarrow{\text{KB}} \neg A$. A model of this type is illustrated in Figure 9 for a belief set K containing $\neg A$, but neither B nor $\neg B$. When M is revised by A, the dark shaded region of A-worlds become most plausible, and when M^*_A is then revised with B, the subregion of $\min(M, A)$ containing B-worlds becomes most plausible. By the results above, since $\neg B \notin K^*_A$ we have $(K^*_A)^*_B = K^*_{A \wedge B}$.

In contrast, when M is revised with B initially, the light shaded region $\min(M, B)$ becomes most plausible, and when M^*_B is revised with A, the dark region $\min(M, A)$ becomes most plausible, "leaving behind" the set $\min(M, B)$. Again, this is supported by our results of the previous section: since $\neg A \in K^*_B$ we have $(K^*_B)^*_A = K^*_A$. In this case, of course, $K^*_A \neq K^*_{A \wedge B}$, so $(K^*_B)^*_A \neq (K^*_A)^*_B$. For this reason, it is important to keep in mind that we are dealing with

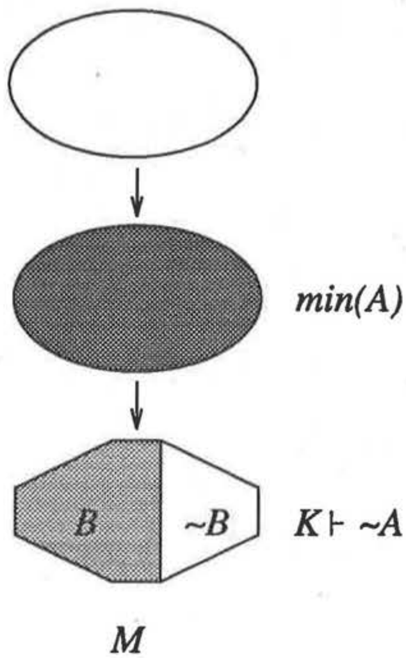


Figure 9: The order dependence of natural revision.

revision *sequences* rather than simple sets of updates. Revision of a belief or extended belief set by some set of new facts will not be order insensitive. In the example above, applying A before B is the same as applying $A \wedge B$ (to K at least), while applying B before A is the same as simply revising by A.

This difference exists for two reasons: first, B is incompatible with A in the sense that $B \xrightarrow{\text{KB}} \neg A$ holds; second, A is less plausible than B. Because A is less plausible than B, revision by A causes more damage to K than revision by B. A is less expected, or conflicts with K to a greater degree. The belief set K^*_A can be thought of as a "radical shift" in belief from K (we draw a very loose analogy to Kuhn's notion of *paradigm shift*). When update B is encountered, it must reconciled with the radically different set K^*_A. If it is consistent with K^*_A then $(K^*_A)^*_B = K^*_{A \wedge B}$. In contrast, if K is revised with B first the results can be thought of as arising from a "routine" revision (routine in comparison to A) followed by the more "radical" revision A. Even though B has been incorporated in K^*_B, the radical shift to $(K^*_B)^*_A$ offers no protection for B. A radical shift has little respect for most routine information in a belief set, and B is as vulnerable as any other fact in K^*_B.

Given this interpretation, it is easy to ascertain just when the order of two revisions is irrelevant, that is, when $(K^*_B)^*_A = (K^*_A)^*_B$.[12]

Definition 10 Let M be a revision model for belief and extended belief sets K and E. Updates A and B are *mutually compatible* (with respect to M or E) iff $M \not\models A \xrightarrow{\text{KB}} \neg B$ and $M \not\models B \xrightarrow{\text{KB}} \neg A$.

Two revisions are mutually compatible just when each is a "routine" revision, relative to the other. This is equivalent to saying A and B are equally plausible: $A \leq_{PM} B$ and $B \leq_{PM} A$.

Proposition 16 *If A and B are mutually compatible with respect to K-revision model M, then $(K^*_B)^*_A = (K^*_A)^*_B$.*

Now, if A and B are incompatible (if one is more plausible than the other) the order of revision is critical. But there are circumstances where the order may still be reversed.

[12]The order of revision is "irrelevant" only with respect to the objective set K. Rarely will the revised models or extended sets be order insensitive. While $(K^*_B)^*_A = (K^*_A)^*_B$ may hold, it will not generally be the case that $(M^*_B)^*_A = (M^*_A)^*_B$. We will discuss this below for arbitrary sequences.

Proposition 17 *Let M be a K-revision model such that $M \models A \xrightarrow{\text{KB}} \neg B$. Then $(K^*_B)^*_A = (K^*_A)^*_B$ iff $M \models B \xrightarrow{\text{KB}} A$.*

Since $(K^*_A)^*_B = K^*_B$ whenever $A \xrightarrow{\text{KB}} \neg B$, it is easy to see that this holds. We simply observe that if $B \xrightarrow{\text{KB}} A$ then $K^*_B = K^*_{A\wedge B}$ and $(K^*_B)^*_A = K^*_{A\wedge B}$. Of course, the only situation left is that where A and B are "super-incompatible". In no instance will the order be irreversible.

Proposition 18 *Let M be a K-revision model such that $M \models A \xrightarrow{\text{KB}} \neg B$ and $M \models B \xrightarrow{\text{KB}} \neg A$. Then $(K^*_A)^*_B = K^*_B$, and $(K^*_B)^*_A = K^*_A$, and $K^*_A \neq K^*_B$.*

4.3.2 Reduction to Single Revisions

In this section, we examine the possibility of simplifying the revision process. Given a revision sequence $A_1, \ldots A_n$, we would like to determine the resulting belief set $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n}$ without having to perform each these n distinct revisions of different belief sets. In fact, we will show that any sequence of revisions can be "reduced" to a single revision. To be more precise we define a *characterizing* sentence for a revision sequence.

Definition 11 Let $A_1, \ldots A_n$ be a revision sequence. We say this sequence is *characterized* by the sentence α iff $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n} = K^*_\alpha$.

Here we show that every revision sequence has such a characterizing sentence, and that this sentence can itself be determined by the simple unnested conditionals contained in the belief set (or, using the Ramsey test, by "hypothetical" unnested revisions of K).

While mutual compatibility is sufficient to ensure that revision ordering can be reversed, we are typically more concerned with processing updates in the order they are received. When A is processed before B, we have seen that $(K^*_A)^*_B = K^*_{A\wedge B}$ whenever $A \xrightarrow{\text{KB}} \neg B$ is false. The mutual compatibility of A and B is not important when revisions are processed in order. Rather the *forward compatibility* of B with A determines the content of $(K^*_A)^*_B$, and how it may be achieved with a single revision. If B is forward compatible with A, that is, if $\neg B \notin K^*_A$, then $(K^*_A)^*_B$ reduces to $K^*_{A\wedge B}$. If B is incompatible, then $(K^*_A)^*_B$ reduces to K^*_B. This can be extended in the obvious fashion to arbitrary sequences of revisions.[13]

[13]We concentrate on revision of models and belief sets, taking for granted the straightforward connections to extended sets and nested conditionals.

Definition 12 Let M be a K-revision model determining natural revision function $*$. The revision sequence $A_1, \ldots A_n$ is *forward compatible* with respect to $*$ (or model M) iff $\neg A_{i+1} \notin ((K^*_{A_1})^*_{A_2} \cdots)^*_{A_i}$ for each $1 \leq i < n$.

This can be restated as

$$M \not\models A_1 \xrightarrow{\text{KB}} (A_2 \xrightarrow{\text{KB}} \cdots (A_i \xrightarrow{\text{KB}} \neg A_{i+1}))$$

for each $i < n$. Loosely, we say that the sequence is forward compatible (or simply compatible) for K, when $*$, M or E is understood. Clearly, we have the following:

Proposition 19 *If a revision sequence $A_1, \ldots A_n$ is forward compatible, so is each subsequence $A_1, \ldots A_i$ for $i \leq n$.*

An obvious inductive argument, generalizing the case of the two-element sequence, gives us the following:

Theorem 20 *If $A_1, \ldots A_n$ is forward compatible for K, then $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n} = K^*_{A_1 \wedge \cdots A_n}$.*

Corollary 21 *$A_1, \ldots A_n$ is forward compatible for K iff $\neg A_{i+1} \notin K^*_{A_1 \wedge \cdots A_i}$ for each $i < n$.*

Thus (by Corollary 21) testing for compatibility can be reduced to the application of single revisions to the belief set K, or testing simple conditionals $A_1 \wedge \cdots A_i \xrightarrow{\text{KB}} A_{i+1}$. Iterated revision or nested conditionals are not required to test for compatibility, nor (by Theorem 20) are they needed to compute the result of such a revision sequence applied to K. Computing compatible revision sequences is a straightforward extension of the case of two compatible revisions, and is reducible to a single revision, the conjunction of the elements.

In the *incompatible* instance, the two-element sequence was again reducible to a single revision: $(K^*_A)^*_B = K^*_B$ when $A \xrightarrow{\text{KB}} \neg B$. Accounting for an incompatible revision after a longer sequence of compatible revisions, however, is not so straightforward. Suppose we have a revision sequence, $A_1, \ldots A_{n+1}$, where $A_1, \ldots A_n$ is compatible but the longer sequence is not. The analogy to the two-element case breaks down here, for in general $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_{n+1}} \neq K^*_{A_{n+1}}$. Unfortunately, earlier revisions leave a residual trace on the structure M, as shown in Figure 10. While $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n} = K^*_{A_1 \wedge \cdots A_n}$, most certainly $((M^*_{A_1})^*_{A_2} \cdots)^*_{A_n}$ is different

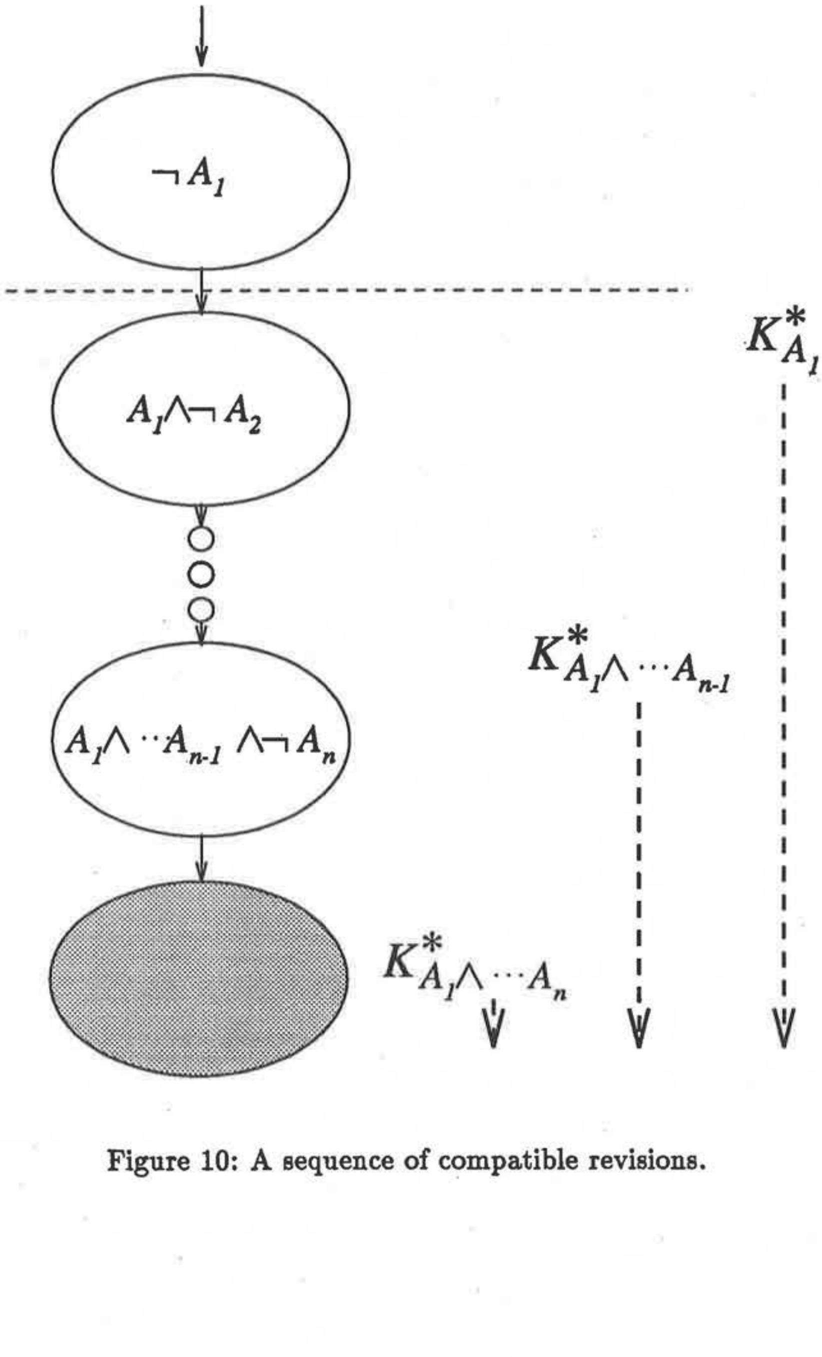


Figure 10: A sequence of compatible revisions.

from $M^*_{A_1\wedge\cdots A_n}$. The history of the belief set K, the process by which it was formed, plays a vital role in future revisions.

In this model, the original structure of the K-revision model lies primarily above the dashed line. Revision by A_1 moved the minimal A_1-worlds to the cluster just below the dashed line. Since A_2 is compatible with A_1, revision by A_2 removed the A_2-worlds from this new cluster (leaving behind $A_1 \wedge \neg A_2$-worlds) and moved them to the second cluster below the line. This process was repeated up to A_n, resulting in a final minimal cluster of $A_1 \wedge \cdots A_n$-worlds (shaded in Figure 10). The same minimal cluster would have been formed had K simply been revised by $A_1 \wedge \cdots A_n$, but the *sequence* of revisions has a drastically different effect on the structure of M, leaving a number of intermediate clusters in its wake. Simply revising by the conjunction $A_1 \wedge \cdots A_n$ would have caused only the shaded cluster to form below the dashed line.

Referring still to Figure 10, if a subsequent revision A_{n+1} is not compatible with the sequence $A_1, \ldots A_n$, then $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_n} \vdash \neg A_{n+1}$. That is, no A_{n+1}-worlds can be found in the shaded minimal cluster. When $((M^*_{A_1})^*_{A_2}\cdots)^*_{A_n}$ is revised with A_{n+1}, it need not be the case that $K^*_{A_{n+1}}$ results. If the minimal A_{n+1}-worlds are found in some cluster below the dashed line, that is, if $K^*_{A_1} \not\vdash \neg A_{n+1}$, then $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_{n+1}}$ will not usually equal $K^*_{A_{n+1}}$.

So exactly where will minimal A_{n+1}-worlds be found in $((M^*_{A_1})^*_{A_2}\cdots)^*_{A_n}$ and what sentences will be in $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_{n+1}}$? Suppose the sequence $A_1,\ldots A_k$, A_{n+1} is incompatible. This means there can be no A_{n+1}-worlds in the cluster formed when $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_{k-1}}$ is revised by A_k, that is, the cluster of $A_1 \wedge \cdots A_k$-worlds. Of course, this implies that there can be no A_{n+1}-worlds is any lower clusters formed by the subsequent revisions A_{k-1} through A_n, since each of these is compatible and will only "select" worlds from this set of $\neg A_{n+1}$-worlds. Conversely, if $A_1,\ldots A_k$, A_{n+1} is a compatible sequence, there must be some A_{n+1}-worlds among the the cluster of $A_1 \wedge \cdots A_k$-worlds representing $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_k}$.

It now becomes clear that the minimal set of A_{n+1}-worlds must be located in the cluster of worlds "labeled" $A_1 \wedge \cdots A_k \wedge \neg A_{k+1}$, where $k \leq n$ is maximal among the set of i such that $A_1,\ldots A_i$, A_{n+1} is a compatible sequence of revisions. Since $A_1,\ldots A_k$ is compatible, $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_k} = K^*_{A_1\wedge\cdots A_k}$, and the minimal A_{n+1}-worlds in this set are captured by $(((K^*_{A_1})^*_{A_2}\cdots)^*_{A_k})^*_{A_{n+1}}$. But this is equivalent to $K^*_{A_1\wedge\cdots A_k\wedge A_{n+1}}$ since A_{n+1} is compatible with the rest of the sequence.

Definition 13 Let $A_1, \ldots A_n$ be a revision sequence. Update A_k is *incompatible* in this sequence iff $\neg A_k \notin ((K^*_{A_1})^*_{A_2} \cdots)^*_{A_{k-1}}$. A sequence is incompatible iff it contains at least one incompatible member.

Proposition 22 *Sequence $A_1, \ldots A_n$ is incompatible iff it is not forward compatible.*

If the belief set $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_{k-1}}$ is characterized by a single revision α but A_k is incompatible, then the sentence representing $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_k}$ is clearly not equivalent to $\alpha \wedge A_k$. The situation we have have described above, where only the last revision in a sequence is incompatible, is easily characterized.

Theorem 23 *Let $A_1, \ldots A_{n+1}$ be an incompatible sequence such that $A_1, \ldots A_n$ is compatible. Let k be the maximal element of*

$$\{i \leq n : \neg A_{n+1} \notin ((K^*_{A_1})^*_{A_2} \cdots)^*_{A_i}\}$$

Then $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_{n+1}} = K^*_{A_1 \wedge \cdots A_k \wedge A_{n+1}}$.

Thus a sequence with one incompatible revision as its last element is reducible to a single revision. Notice that when the set of revisions compatible with A_{n+1} is empty, when this maximal element k does not exist, we have $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_{n+1}} = K^*_{A_{n+1}}$. This is directly analogous to the two-element case, since there is no subsequence compatible with A_{n+1}. In Figure 10 this occurs exactly when there are no A_{n+1}-worlds below the dashed line; that is, when $A_1 \xrightarrow{\text{KB}} \neg A_{n+1}$.

It should be quite clear that subsequent compatible revisions, A_{n+2} and so on, should be treated as previously and simply "conjoined" to $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_{n+1}}$.

Proposition 24 *Let $A_1, \ldots A_n$ be a revision sequence with one incompatible element A_k, and let $j < k$ be the maximal compatible revision for A_k (as in Theorem 23). Then $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n} = K^*_{A_1 \wedge \cdots A_j \wedge A_k \wedge \cdots A_n}$.*

The final piece in the puzzle is the process by which subsequent incompatible revisions are achieved. Consider a revision sequence $A_1, \ldots A_n$ where A_k is incompatible and has as its *most recent compatible revision* A_j (the element defined in Theorem 23). This situation is illustrated in Figure 11. Now suppose update A_{n+1} is incompatible, so that no A_{n+1}-worlds are located in the minimal cluster. This occurs when $K^*_{A_1\wedge\cdots A_j\wedge A_k\wedge\cdots A_n} \vdash \neg A_{n+1}$. Again, to find the minimal A_{n+1}-worlds in this structure we must look for the most recent compatible revision in the sequence $A_1, \ldots A_n$. If A_i is this update then $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_{n+1}}$ is identical to $(((K^*_{A_1})^*_{A_2}\cdots)^*_{A_i})^*_{A_{n+1}}$. Furthermore, since $A_1, \ldots A_n$ has only one incompatible revision, by Theorem 20 and Proposition 24, we have $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_i} = K^*_\alpha$ for some sentence α. Thus $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_{n+1}} = K^*_{\alpha\wedge A_i}$. It is interesting, however, to examine the various situations that arise with respect to the occurrence of this most recent compatible revision A_i.

First, consider $i \geq k$. In this case, A_{n+1} is compatible with the previous incompatible revision A_k. The set of minimal A_{n+1}-worlds lies below the third dashed line, among those worlds representing $K^*_{A_1\wedge\cdots A_j\wedge A_k}$. In this circumstance we have $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_{n+1}} = K^*_{A_1\wedge\cdots A_j\wedge A_k\wedge\cdots A_i\wedge A_{n+1}}$.

Second, consider $j < i < k$. Clearly, A_{n+1} is incompatible with the incompatible revision A_k, but is compatible with the sequence $A_1, \ldots A_j$. This case is rather interesting for we cannot simply "backtrack" within our representative revision for $A_1, \ldots A_n$. Because $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_n} = K^*_{A_1\wedge\cdots A_j\wedge A_k\wedge\cdots A_n}$, one might think we could simply "back up" to the most recent compatible revision A_j in this representation and arrive at $K^*_{A_1\wedge\cdots A_j}$. However, this ignores the consistent revisions between A_j and A_k (between the second and third dashed lines) that were "left behind" when the incompatible revision A_k was incorporated. The minimal A_{n+1}-worlds lie in this region and this must be taken into account. Indeed, since $A_1, \ldots A_i$ is compatible we have $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_{n+1}} = K^*_{A_1\wedge\cdots A_i\wedge A_{n+1}}$.

Finally, consider $i \leq j$. Since $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_{n+1}} = K^*_{A_1\wedge\cdots A_i}$, this case is much like the one just mentioned. It is distinguished by the fact that the single revision $A_1 \wedge \cdots A_i$ that represents it is a proper subsequence of the update $A_1 \wedge \cdots A_j \wedge A_k \wedge \cdots A_n$ representing $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_n}$.

These considerations can be generalized to accommodate any number of incompatible revisions. Furthermore, they provide a constructive means (described inductively) of reducing any sequence of revisions of belief set K to a single revision of K, and demonstrate (through compatibility testing) how that revision can itself be determined using only single

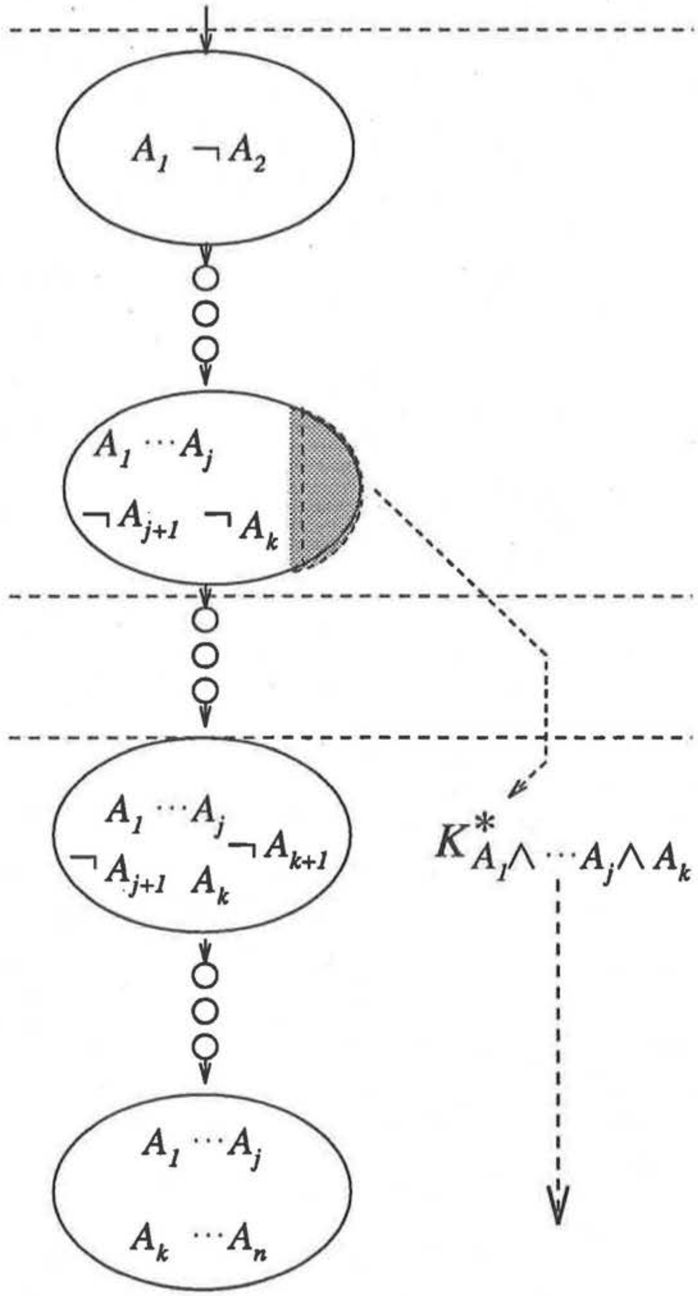


Figure 11: A sequence with an incompatible revision.

(non-iterated) revisions of K. Finally we shall see that, although the inductive description indicates a dependence of the characterization of $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_n}$ on the characterization of $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_i}$, for each $i < n$, it is only necessary to keep track of those sentences that characterize incompatible revisions.

Definition 14 Let $A_1,\ldots A_n$ be a revision sequence with c incompatible updates. We use the strictly increasing, injective function $\sigma : \{1,\cdots,c\} \mapsto \{1,\cdots,n\}$ to denote these incompatible elements: $A_{\sigma(1)}, \ldots A_{\sigma(c)}$. For each $1 < j \leq n$, the *maximal consistent incompatible revision* for A_j is A_k, where

$$k = \max\{\sigma(i) : \sigma(i) < j \text{ and } \neg A_j \notin ((K^*_{A_1})^*_{A_2}\cdots)^*_{A_{\sigma(i)}}\}$$

For each $1 < j \leq n$, the *most recent compatible revision* for A_j is A_k, where

$$k = \max\{i : i < j \text{ and } \neg A_j \notin ((K^*_{A_1})^*_{A_2}\cdots)^*_{A_i}\}$$

If either of these sets is empty, we take the most recent or maximal incompatible revision for A_j to be $\top$.

Lemma 25 *Let $A_1,\ldots A_n$ be a revision sequence such that each proper subsequence $A_1,\ldots A_i$ is characterized by some sentence $s(A_i)$. Then $A_1,\ldots A_n$ is characterized by $s(A_k) \wedge A_n$, where A_k is the most recent compatible revision for A_n. In other words, $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_n} = K^*_{s(A_k)\wedge A_n}$.*

This leads to the main result of this section.

Theorem 26 *For any revision sequence $A_1,\ldots A_n$, there is some subset of these updates $S \subseteq \{A_1,\cdots,A_n\}$ such that $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_n} = K^*_A$ and $A = \wedge S$.*

Corollary 27 *For any revision sequence $A_1,\ldots A_n$, there is some subset of these updates $S \subseteq \{A_1,\cdots,A_n\}$ such that $((E^*_{A_1})^*_{A_2}\cdots)^*_{A_{n-1}} \models A_n \xrightarrow{\text{KB}} B$ iff $E \models A \xrightarrow{\text{KB}} B$, and $A = \wedge S$.*

This result is given its constructive character by Theorem 20, but it seems to suggest that one must keep track of a characterizing sentence $s(A_i)$ for each revision A_i. In fact, the critical sentences are only those corresponding to incompatible revisions, $s(A_{\sigma(i)})$. Every other characterizing sentence $s(A_i)$ is simply the conjunction of subsequent revisions to the *most recent* incompatible revision.

Theorem 28 *Let $A_1, \ldots A_n$ be a revision sequence with c incompatible updates represented by σ. For each $1 \leq k \leq n$, if A_k is a compatible revision, then $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_k} = K^*_{s(A_k)}$, where: a) $s(A_k) = s(A_{\sigma(i)}) \wedge A_{\sigma(i)+1} \wedge \cdots A_k$; b) $\sigma(i) < k < \sigma(i+1)$, or $\sigma(i) < k$ if $i = c$; and c) $s(A_{\sigma(i)})$ characterizes subsequence $A_1, \ldots A_{\sigma(i)}$.*

We can provide a similar characterization of incompatible revisions, but these must be in terms of the *maximal consistent incompatible revision* rather than the most recent incompatible revision.

Proposition 29 *Let $A_1, \ldots A_n$ be a revision sequence with c incompatible updates represented by σ. Let $A_{\sigma(i)}$ be the maximal consistent incompatible revision for A_k. If A_j is the most recent compatible revision for A_k, then $\sigma(i) \leq j < \sigma(i+1)$.*

Theorem 30 *Let $A_1, \ldots A_n$ be a revision sequence with c incompatible updates represented by σ. For each $1 \leq k \leq n$, if A_k is an incompatible revision, then $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_k} = K^*_{s(A_k)}$, where: a) $s(A_k) = s(A_{\sigma(i)}) \wedge A_{\sigma(i)+1} \wedge \cdots A_j \wedge A_k$; b) $A_{\sigma(i)}$ is the maximal incompatible revision for A_k; c) A_j is the most recent compatible revision for A_k; and d) $s(A_{\sigma(i)})$ characterizes subsequence $A_1, \ldots A_{\sigma(i)}$.*

Taken together, these theorems show that one may implement a procedure that tests for membership of B in a multiply-revised belief set $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n}$ using only an "oracle" that answers requests of the form "Is $\beta \in K^*_\alpha$?" for $\alpha, \beta \in \mathbf{L}_{CPL}$. Furthermore, the characterizing sentences $s(A_j)$ that need to be recorded are only those of the form $s(A_{\sigma(i)})$ where $A_{\sigma(i)}$ is some incompatible revision. The core of this algorithm is provided in Figure 12. It takes as input a revision sequence $A_1, \ldots A_n$ and computes the characterizing sentence A such that $K^*_A = ((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n}$. The algorithm is incremental in the sense that a subsequent revision A_{n+1} requires only one further execution of the main "for loop". To ask whether $B \in ((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n}$, one simply computes the required characterizing sentence A and asks if $B \in K^*_A$. In the next section, we describe how this algorithm may be realized using theorem-

```
Input:     revision sequence A_1,...A_n; revision function * applicable to belief set K
Output:    characterizing update A — K*_A = ((K*_{A_1})*_{A_2} ···)*_{A_n}
Variables: L — list of updates
           IL — list of incompatible revisions: elements have form
           ⟨ind, S⟩, where A_ind is incompatible and subsequence
               A_1,...A_ind is characterized by sentence S
           S — characterizing sentence A built up here

Initialize A ← ⊤; L ← ⊤; IL ← ⟨0, ⊤⟩
for i = 1 to n      ;;; one loop for each update A_i
    L ← L + A_i
    if ¬A_i ∉ K*_A then       ;;; A_i is compatible
        A ← A ∧ A_i
    else     ;;; A_i is incompatible
        for j = length(IL) to 1
            ;;; find maximal incompatible revision consistent with A_i
            ;;; Note: non-sequential search methods are possible
            S ← IL(j).S
            idx ← IL(j).ind
            if ¬A_i ∉ K*_S then      ;;; found maximal incompatible revision
                if x = length(IL) then       ;;; top is max location of most recent compatible revision
                    top ← i − 1
                else top ← IL(j).ind
                end if
                break
            end if
        end for
        for k = idx + 1 to top − 1       ;;; find most recent compatible revision
            if ¬A_i ∈ K*_{S∧L(k)} then
                break
            else S ← S ∧ L(k)
            end if
        end for
        A ← S
        IL ← IL + ⟨i, S⟩
    end if     ;;; end of incompatible revision
end for
```

Figure 12: Algorithm to compute characterizing update A for revision sequence $A_1, \ldots A_n$.

proving in CO* for incompletely specified revision functions, and how a revision function may be "completed" so that the "oracle" will always provide an answer.

4.4 Information Preservation

One desideratum of any model of the revision process is the minimization of information loss. In the AGM framework for single revisions, postulate (R4) ensures that no beliefs from set K are retracted if none need to be. An arbitrarily long sequence of *consistent* revisions will simply make a belief set larger, resulting (in the limit) in a *complete* belief set. Eventually, we will believe either A or $\neg A$ for each objective sentence A. Once an inconsistent revision is processed, the AGM model asserts that, among those beliefs that could be given up to accommodate A, only those that are least entrenched are retracted. Unfortunately, no constraints are levied on the relative degrees of entrenchment of the members of the new belief set K^*_A. Much of the information associated with K, including relative degrees of entrenchment, is lost.

The natural revision model is unique in the sense that this type of information is preserved. When a consistent revision A is processed, the minimal cluster of a revision model M is divided, resulting in two smaller clusters. A subsequent consistent revision breaks the minimal cluster of M^*_A, and so on. Eventually (supposing some finite language) after a suitable sequence of consistent revisions, we would end up with a minimal cluster containing one possible world, representing a complete belief set.

The process is not altogether different for inconsistent revisions. Rather than dividing the minimal cluster of K-worlds, the cluster containing $\min(M, A)$ is divided. However, any information implicit in previous revisions that caused the formation of various clusters in M is preserved. All other clusters remain undisturbed. In no case will a revision cause any cluster to "grow" or "lose information." Typically, a revision will cause the number of clusters to increase by one by splitting $\min(M, A)$ from a cluster, thus "shrinking" a cluster and ensuring information gain. Only in certain cases, will no new clusters be formed, for example, when we revise K by A where $A \in K$.

Proposition 31 *Let revision model $M^*_A = \langle W, R, \varphi \rangle$ be the natural revision of M by A. If $C \subseteq W$ is a cluster in M^*_A, then $C \subseteq C'$ for some cluster C' in M.*

If $\min(M, A) \subset C'$ *for some cluster* C' *in* M*, then the set of clusters in* M_A^* *consists of the set of clusters in* M *distinct from* C' *together with* $\min(M, A)$ *and* $C' - \min(M, A)$*. If* $\min(M, A) = C'$ *for some cluster* C' *in* M*, the the set of clusters in* M_A^* *is identical to that in* M.

Thus, a revision sequence $A_1, \ldots A_n$ causes a non-decreasing change in "information" in a belief state. No belief set farther along in the revision sequence can be smaller than an earlier belief set.

Corollary 32 *Let* M *be a* K*-revision model and* $A_1, \ldots A_n$ *a revision sequence. If* $i \leq j$ *then* $((K_{A_1}^*)_{A_2}^* \cdots)_{A_j}^* \not\subset ((K_{A_1}^*)_{A_2}^* \cdots)_{A_i}^*$.

This suggests that, as we process a revision sequence, our revision model becomes more and more *informationally complete* in the following sense.

Definition 15 A revision model M is *informationally complete* iff each cluster in M is a singleton set.

Proposition 33 *Let* M *be an informationally complete* K*-revision model. Then* K_A^* *is a complete theory for any* $A \in \mathrm{L}_{CPL}$.

Corollary 34 *If* M *is an informationally complete* K*-revision model then* K *is a complete theory.*

Proposition 31 ensures furthermore that M_A^* will be informationally complete whenever M is. It also shows that if we restrict our attention to languages with a finite number of propositional variables, we can eventually attain informational completeness.[14]

Proposition 35 *Let* M *be a revision model and* $A_1, \ldots A_n$ *a revision sequence such that every satisfiable sentence* α *is represented in this sequence; that is,* $A_i \vdash \alpha$ *for some* A_i*. Then* $((M_{A_1}^*)_{A_2}^* \cdots)_{A_n}^*$ *is informationally complete.*

[14]We assume that models contain no "duplicate" worlds having the same associated valuation. If so, our informational completeness in the following result is the type that ensures that K_A^* is complete for all A.

An example of such a revision sequence, for a language with p propositional variables, would be the set of the 2^p complete (truth-functionally distinct) conjunctions of literals, each capturing a possible world or valuation, and each causing the corresponding world to be broken off into its own cluster. From that point on all revisions will result in complete belief sets since the set of most plausible A-worlds will have one element for each revision A.

While Proposition 31 ensures that we will constantly approach informational completeness, it also demonstrates that natural revision preserves information by Corollary 32, leading to some interesting recovery properties.

Theorem 36 *Let M be a K-revision model and $A_1, \ldots A_n$ a revision sequence such that $A_2, \ldots A_n$ is a compatible sequence and $M^*_{A_1} \models A_2 \xrightarrow{\text{KB}} \neg A_1$. Then $(((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n})^*_{A_1} = K^*_{A_1}$.*

5 Reasoning with Revision Sequences

In Section 4.1 we provided new truth conditions for the connective $\xrightarrow{\text{KB}}$ based explicitly on the Ramsey test: $A \xrightarrow{\text{KB}} B$ is evaluated by asking whether M^*_A satisfies $\boxtimes B$. While this apparently required the development of a new logical calculus (these truth conditions no longer equivalent to the original modally-specified definition) Theorem 26 and Corollary 27 demonstrated that nested conditionals and iterated revision can be reduced to simple conditionals. By Proposition 11, the two forms of truth conditions are identical for simple conditionals. Therefore, we can reason about natural revision and revision sequences using CO* simply by reducing nested conditionals to their equivalent simple counterparts using the algorithm presented in the last section. A query about the truth of a nested conditional is equivalent to a query about its reduction.

The algorithm for reducing nested conditionals requires some method of establishing the truth of $B \in K^*_A$ for $A, B \in \mathbf{L}_{CPL}$. Typically, as is the case in most reasoning tasks, our premises do not provide us with complete knowledge, and we can only hope to derive as much as possible, leaving certain gaps in our knowledge. It is not reasonable to expect a completely specified revision function $*$, or equivalently, a complete set of conditionals containing one of $A \xrightarrow{\text{KB}} B$ or $\neg(A \xrightarrow{\text{KB}} B)$, among our set of premises. Naturally, the algorithm can then easily be modified to ask whether $A \xrightarrow{\text{KB}} B$ or $\neg(A \xrightarrow{\text{KB}} B)$ is *provable* from a given premise set. If either is the case, the algorithm can proceed, having an answer to the query $B \in K^*_A$. If neither is derivable from the premises, then the algorithm must halt unsuccessfully, or proceed as if either could be the case.[15] As should be expected when reasoning with incomplete information, the answer "unknown" must be returned for certain queries.

While the revision function $*$ is usually specified only partially by means of conditional premises (as well as direct statements of plausibility and entrenchment), there may be circumstances when the revision function is completely known. The problem changes from that of reasoning with incomplete information to that of specifying complete information in a reasonable manner. We cannot expect one to specify a complete conditional theory explicitly containing $A \xrightarrow{\text{KB}} B$ or $\neg(A \xrightarrow{\text{KB}} B)$ for each objective A and B, for we do not want to be forced to reason with infinite sets of premises. Even for finite languages with n atomic variables, where we require only a conditional or its negation for semantically distinct A and B, we are forced to reason with 2^{n+1} premises.[16]

There are cases, though, where a revision function can be captured finitely, and often with a manageable number of sentences. Often when a set of conditional premises is given, we have in mind a certain intended model that can be represented finitely. Pearl's (1990) System Z is an example of this, specifying one type of preferred conditional model in the context of default reasoning. Boutilier (1991) has shown how this model can be compactly represented in CO*. In general, a revision function can be compactly represented if there is a corresponding revision model that is "well-behaved" in the following sense.

Definition 16 A revision model $M = \langle W, R, \varphi \rangle$ is *finitely specifiable* iff $W = \cup_{i \leq n} C_i$, where each C_i is a cluster in M and $C_i = \|S_i\|$ for some sentence $S_i \in \mathbf{L}_{CPL}$.

In other words, M is well-behaved if it consists of a finite number of clusters, each corresponding to some finite classical theory or sentence. It is easy to verify that if M is finitely

[15] It often makes sense to continue, for a subsequent revision may be incompatible with previous revisions, and failing to reduce an earlier revision may have no effect on the effort to reduce an entire sequence.

[16] Alternatively, one could provide a complete set of entrenchment sentences, specifying the relative degrees of entrenchment of each pair of sentences: $A \leq_E B$ and/or $B \leq_E A$ for each $A, B \in \mathbf{L}_{CPL}$. This would allow the derivation of every simple conditional or its negation, and seems to be what Gärdenfors and Makinson (1991; 1990) have in mind in their presentation of *expectation inference*.

specified by sentences $S_1, \ldots S_n$ in the definition above, then these sentences are "mutually exclusive" and "exhaustive"; that is, $\vdash S_i \supset \neg S_j$ if $i \neq j$, and $\vdash S_1 \vee \cdots S_n$. We assume that S_1 characterizes the minimal (most plausible) cluster of M, S_2 the nest most plausible and so on. If M is such a model, it is uniquely characterized by the following CO* theory:

$$\begin{array}{ll} S_i \supset \Box(S_1 \vee \cdots S_i) & \text{for } i \leq n \\ S_i \supset \overleftarrow{\Box}\neg S_i & \text{for } i \leq n \end{array}$$

We denote the conjunction of the sentences in this set by $FSM(S_1, \cdots S_n)$: the sentence finitely specifying a revision model M.

Definition 17 A revision function $*$ is *finitely specifiable* (for belief set K) iff there is some finitely specifiable K-revision model M such that $*^{M} = *$.

Given a finitely specifiable revision function for belief set K (in which case K is also finitely specified by S_1), we can use the premise set $FSM(S_1, \cdots S_n)$ to represent the model M and use it to determine the truth of every simple conditional sentence.

Theorem 37 *Let $S_1, \cdots S_n \in \mathbf{L}_{CPL}$ be such that $\vdash S_i \supset \neg S_j$ if $i \neq j$, and $\vdash S_1 \vee \cdots S_n$. For all $A, B \in \mathbf{L}_{CPL}$ either*

$$\begin{array}{ll} FSM(S_1, \cdots S_n) \vdash_{CO*} A \xrightarrow{\text{KB}} B; & \textit{or} \\ FSM(S_1, \cdots S_n) \vdash_{CO*} \neg(A \xrightarrow{\text{KB}} B) & \end{array}$$

Fortunately, the ability to finitely specify a revision model is not disturbed by natural revision. If $FSM(S_1, \cdots S_n)$ characterizes a revision model M, then M_A^* is formed by simply "dividing" the minimal cluster consistent with A in two clusters: $S_i \wedge A$ becomes most plausible; and $S_i \wedge \neg A$ replaces cluster S_i.

Theorem 38 *Let revision model M be characterized by $FSM(S_1, \cdots S_n)$ and let S_k be the minimal sentence in this set consistent with A; that is, $S_k \not\vdash \neg A$ and $S_i \vdash \neg A$ if $i < k$. Then M_A^* is characterized by*

$$\begin{array}{ll} FSM(S_k \wedge A, S_1, \cdots S_{k-1}, S_k \wedge \neg A, S_{k+1}, \cdots S_n) & \textit{if } S_k \not\vdash A \\ FSM(S_k, S_1, \cdots S_{k-1}, S_{k+1}, \cdots S_n) & \textit{if } S_k \vdash A \end{array}$$

Thus we can completely specify a revision function with a compact set of premises and use this premise set to reduce nested conditional queries to simple conditionals and then establish the truth of these simple conditionals. Furthermore, we can explicitly revise a model and retain a compact representation. It also becomes clear that testing for the truth of a simple conditional in such a revision model is reducible to a simple propositional reasoning task. Given a revision function $*$ completely specified by the sentence $FSM(S_1, \cdots S_n)$, the following procedure will ascertain the truth of $B \in K_A^*$, or $A \xrightarrow{\text{KB}} B$:

(a) Find the minimal S_k consistent with A: $S_k \nvdash \neg A$ and $S_i \vdash \neg A$ if $i < k$.

(b) If $S_k \wedge A \vdash B$ then $B \in K_A^*$.

(c) If $S_k \wedge A \nvdash B$ then $B \notin K_A^*$.

Relative to propositional satisfiability tests, this "algorithm" is relatively efficient, for the minimal A-consistent cluster or sentence can be found using a linear search technique. Thus we need only perform $O(n)$ satisfiability tests to determine the truth of $A \xrightarrow{\text{KB}} B$. The complexity of these tests will depend on the size of the theories S_i and their structure. For instance, if the S_i are Horn theories, this test will be linear in the size of the theory. This is reminiscent of the algorithm used by Pearl (1990) for default reasoning. Pearl's System Z computes a preferred model of a set of default rules in a relatively efficient manner. The same assumptions can also be applied here, given that the formal properties of our subjunctive conditional and the conditionals used in default reasoning are identical (see (Boutilier 1992e)). From a set of simple conditional sentences, the characterizing theory $FSM(S_1, \cdots S_n)$ can be easily computed (Boutilier 1991) for this preferred model. Using the S_i, the truth of further conditionals $A \xrightarrow{\text{KB}} B$ can then be computed as described.

6 Concluding Remarks

We have presented a model for iterated revision that captures sequences of propositional revisions. The hallmark of natural revision is the preservation of subjective information, such as conditional beliefs and entrenchment, across such a sequence. Because this information is retained across belief sets, simply knowing how a revision function behaves on single propositional revision is sufficient to characterize the results of any sequence of revisions. If we adopt the Ramsey test for acceptance of conditionals, this demonstrates that right-nested conditionals are equivalent to simple unnested conditionals, and that this reduction can be performed using only knowledge of accepted unnested conditionals.

The natural revision model has other compelling properties as well. Because it preserves conditional information, the sequence of belief sets corresponding to a revision sequence can never decrease in propositional information content. It preserves nice properties of the revision models and functions being revised as well, for example, informational completeness and finite specifiability. In cases where a revision function is described simply (e.g., by a set of propositional sentences), the characterization of natural revision is also easily computed.

There are several models of revision and opinions regarding nested conditionals that bear some resemblance to ours. Spohn's (1987) use of ordinal conditional functions to represent belief states is much like the possible worlds model we have used (along with other such as Grove (1988) and Katsuno and Mendelzon (1990)), except that "clusters" are given explicit ordinal rankings. Spohn's notion of conditionalization on such models assigns to a proposition A a new "degree of belief" or ordinal ranking (such rankings should be thought of as entrenchment of beliefs). This is the analogue of revising by A and is achieved by "shifting" the rank of all A-worlds so that the minimal A-worlds have an ordinal ranking that is lower (or more plausible) than that of any $\neg A$-world by the specified degree of belief. While the explicit representation of degrees of entrenchment provides a superficial difference, the key distinction between Spohn's approach and ours is that revising by A requires the shift of only the *minimal* A-worlds during natural revision. Spohn's approach is more reminiscent or probabilistic conditionalization, and if we applied our truth conditions for conditionals to ordinal conditional function update, we would see a much larger change in conditional beliefs in general.

It is also interesting to note that a nested conditional $A \xrightarrow{\mathrm{KB}} (B \xrightarrow{\mathrm{KB}} C)$ is often equivalent to $A \wedge B \xrightarrow{\mathrm{KB}} C$ (whenever A and B are compatible). It has been suggested by a number of people that nested conditionals should be reduced to unnested conditionals with all antecedents conjoined to form a single antecedent, Adams (1975) and Levi (1988) among them. This reduction is sanctioned by the natural revision model for nested conditionals with compatible antecedents. But as described in Section 4, this is not the case when incompatible antecedents are present. For instance, $A \xrightarrow{\text{KB}} (B \xrightarrow{\text{KB}} C)$ reduces simply to $B \xrightarrow{\text{KB}} C$ when A and B are incompatible. As suggested by Levi, this reduction may seem inappropriate in normal linguistic usage, for the nested conditional seems to imply that A should continue to hold when $B \xrightarrow{\text{KB}} C$ is evaluated with the Ramsey test, thus suggesting the reduction to $A \wedge B \xrightarrow{\text{KB}} C$. Indeed, the natural revision model cannot account for this circumstance when incompatible antecedents are involved. In contrast, our model provides non-trivial acceptance conditions for nested conditionals such as $A \xrightarrow{\text{KB}} (\neg A \xrightarrow{\text{KB}} B)$. In this sense, our model is more appropriate for conditionals that model a sequence of revisions or changes in belief state. It is this model that may be more appropriate in AI applications such as diagnosis, where tests provide incremental (perhaps "contradictory") changes in belief, and planning, where information-gathering actions are used to change an agent's state of knowledge of the world. Indeed, the *update* models of belief change (Winslett 1988; Katsuno and Mendelzon 1991), more suitable for modeling changes due to the actions of an agent, may also be integrated into our model of revision sequences.

We should also note that Levi (1988) is rather critical of the enterprise of determining truth or acceptance conditions for nested conditionals, or even allowing conditionals to be part of a belief set. He offers the opinion that an element of a belief set or *corpus* ought to be practicable as a standard of serious possibility. Since conditionals do not perform this function, they are not accorded the status of beliefs. While certainly their role differs from that of garden-variety propositional beliefs, we adopt the stance that they perform an indispensable function in the process of deliberation. Conditionals suggest hypothetical possibilities to an agent, and aid an agent in changing its mind. It's hard to imagine how these should be represented if not as beliefs.

There are a number of interesting avenues that remain to be explored. This model is restricted to right-nested conditionals, or propositional revisions. In general, we want to allow revision of a knowledge base with conditional information, or statements of entrenchment as well. A fully general model of this type is currently being developed in (Boutilier 1992c). However, the extension of the *natural* revision model, with its very constrained behavior, to the fully general case has yet to be explored. Such an extension will allow one to assert as premises, for the purposes of natural revision, nested conditional sentences that "indirectly" constrain a revision model in the same way that simple conditionals directly constrain a model. The application of natural revision to default reasoning should also prove interesting, as have the connections of AGM revision to default reasoning (Gärdenfors and Makinson 1991; Boutilier 1992e). The paranoid revision model, also of intrinsic interest as the maximal logic of AGM revision, seems strongly related to epistemic logics used in knowledge representation, and to "minimal knowledge" characterizations of Reiter's default logic (Truszczyński 1991). This may provide an interesting interpretation of default logic in terms of belief revision.

We are currently developing a model of revision and conditionalization for single revisions that adds to the basic revision model probabilistic degrees of belief (Boutilier 1992f). If this model can be grafted onto the natural revision model, "conditional objects" that make statements of conditional probability can be nested in a meaningful way, and given a natural semantics. We are also exploring the application of these ideas to the processes of *J-conditionalization* and *L-conditionalization*. These models were proposed by Goldszmidt and Pearl (1992) to capture changes in belief that have degrees of certainty attached, these degrees corresponding to Spohn's κ-rankings (or degrees of entrenchment or possibility). Naturally, a sequence of changes (for example, as arising from several diagnostic tests suggesting how strongly one should believe in the failure of a given component) need not be mutually consistent. The natural revision model can be used to show how to resolve these conflicts rather than trying to, say, add the conjunction of all changes. Both extensions of natural revision (either probabilistic or possibilistic) offer ways of attaching quantitative strength measurements to our conditionals and revising these conditionals.

Finally, this model reflects the bias of the AGM model to accepting without question the most recent update. The *primacy of the most recent information* is clearly not a principle that should be accepted in all circumstances. Sometimes things we learn are so radically incompatible with our knowledge that we reject them out of hand, and do not attempt to reconcile them with our current beliefs. Generalizing the AGM and natural revision models in this way is a difficult task, but one that certainly deserves inquiry.

Acknowledgements

Thanks to Mirek Truszczyński for discussing these ideas with me, and for suggesting the two-cluster (paranoid) revision models explored in Section 3.

References

Adams, E. W. 1975. *The Logic of Conditionals.* D.Reidel, Dordrecht.

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

Boutilier, C. 1991. Inaccessible worlds and irrelevance: Preliminary report. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence*, pages 413–418, Sydney.

Boutilier, C. 1992a. Conditional logics for default reasoning and belief revision. Technical Report KRR-TR-92-1, University of Toronto, Toronto. Ph.D. thesis.

Boutilier, C. 1992b. Epistemic entrenchment in autoepistemic logic. *Fundamenta Informaticae.* To appear.

Boutilier, C. 1992c. A general model for nested conditionals. Technical report, University of British Columbia, Vancouver. (Forthcoming).

Boutilier, C. 1992d. A logic for revision and subjunctive queries. In *Proceedings of the Tenth National Conference on Artificial Intelligence*, pages 609–615, San Jose.

Boutilier, C. 1992e. Normative, subjunctive and autoepistemic defaults: Adopting the Ramsey test. In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning*, Boston. To appear.

Boutilier, C. 1992f. The probability of a possibility: Adding degrees of belief to conditionals. Technical report, University of British Columbia, Vancouver. (Forthcoming).

Boutilier, C. 1992g. Unifying default reasoning and belief revision. In *Fourth International Workshop on Nonmonotonic Reasoning*, pages 52–64, Plymouth, VT.

Fuhrmann, A. 1989. Reflective modalities and theory change. *Synthese*, 81(1):115–134.

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*. MIT Press, Cambridge.

Gärdenfors, P. and Makinson, D. 1991. Nonmonotonic inference based on expectations. To appear.

Ginsberg, M. L. 1986. Counterfactuals. *Artificial Intelligence*, 30(1):35–79.

Goldszmidt, M. and Pearl, J. 1992. Reasoning with qualitative probabilities can be tractable. In *Proceedings of the Eighth Conference on Uncertainty in AI*, Stanford. To appear.

Grove, A. 1988. Two modellings for theory change. *Journal of Philosophical Logic*, 17:157–170.

Haddawy, P. and Hanks, S. 1992. Representations for decision-theoretic planning: Utility functions for deadline goals. In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning*, Boston. To appear.

Halpern, J. Y. and Moses, Y. O. 1985. A guide to the modal logics of knowledge and belief: A preliminary draft. In *Proceedings of the Ninth International Joint Conference on Artificial Intelligence*, pages 480–490, Los Angeles.

Katsuno, H. and Mendelzon, A. O. 1990. Propositional knowledge base revision and minimal change. Technical Report KRR-TR-90-3, University of Toronto, Toronto.

Katsuno, H. and Mendelzon, A. O. 1991. On the difference between updating a knowledge database and revising it. In *Proceedings of the Second International Conference on Principles of Knowledge Representation and Reasoning*, pages 387–394, Cambridge.

Kraus, S., Lehmann, D., and Magidor, M. 1990. Nonmonotonic reasoning, preferential models and cumulative logics. *Artificial Intelligence*, 44:167–207.

Levesque, H. J. 1990. All I know: A study in autoepistemic logic. *Artificial Intelligence*, 42:263–309.

Levi, I. 1988. Iteration of conditionals and the ramsey test. *Synthese*, 76(1):49–81.

Lewis, D. 1973. *Counterfactuals.* Blackwell, Oxford.

Makinson, D. and Gärdenfors, P. 1990. Relations between the logic of theory change and nonmonotonic logic. In Fuhrmann, A. and Morreau, M., editors, *The Logic of Theory Change*, pages 185–205. Springer-Verlag, Berlin.

McArthur, G. 1988. Reasoning about knowledge and belief: A survey. *Computational Intelligence*, 4(3):223–243.

Moore, R. C. 1985. Semantical considerations for nonmonotonic logic. *Artificial Intelligence*, 25:75–94.

Nebel, B. 1989. A knowledge level analysis of belief revision. In *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning*, pages 301–311, Toronto.

Pearl, J. 1990. System Z: A natural ordering of defaults with tractable applications to default reasoning. In Vardi, M., editor, *Proceedings of Theoretical Aspects of Reasoning about Knowledge*, pages 121–135. Morgan Kaufmann, San Mateo.

Reiter, R. 1980. A logic for default reasoning. *Artificial Intelligence*, 13:81–132.

Rott, H. 1989. Conditionals and theory change: Revisions, expansions, and additions. *Synthese*, 81(1):91–113.

Schwarts, G. F. and Truszczyński, M. 1992. Modal logic S4F and the minimal knowledge paradigm. In *Proceedings of the Fourth Conference on Theoretical Aspects of Reasoning about Knowledge*, pages 184–198, Pacific Grove.

Segerberg, K. 1970. Modal logics with linear alternative relations. *Theoria*, 36:310–322.

Segerberg, K. 1971. *An Essay in Classical Modal Logic.* Department of Philosophy, University of Uppsala, Uppsala.

Spohn, W. 1987. Ordinal conditional functions: A dynamic theory of epistemic states. In Harper, W. and Skyrms, B., editors, *Causation in Decision, Belief Change and Statistics*, volume 2, pages 105–134. D. Reidel, Dordrecht.

Stalnaker, R. C. 1968. A theory of conditionals. In Harper, W., Stalnaker, R., and Pearce, G., editors, *Ifs*, pages 41–55. D. Reidel, Dordrecht. 1981.

Stalnaker, R. C. 1984. *Inquiry*. MIT Press, Cambridge.

Truszczyński, M. 1991. Modal interpretation of default logic. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence*, pages 393–398, Sydney.

Winslett, M. 1988. Reasoning about action using a possible models approach. In *Proceedings of the Seventh National Conference on Artificial Intelligence*, pages 89–93, St. Paul.

A Proofs of Main Results

The truth of many of the propositions in the paper is rather obvious and their proofs are excluded. Certain results are described in the body of the paper and have their proofs sketched there. These proofs are also excluded.

Proposition 9 *Let $M = \langle W, R, \varphi\rangle$ be a paranoid revision model for K, where $K \vdash \neg A$. If $*$ is the paranoid revision function then $M^*_A = \langle W, R', \varphi\rangle$ where $vR'w$ iff $M \models_w A$ or $M \models_v \neg A$. Consequently, $K^*_A = Cn(\{A\})$.*

Proof Let C_1 be the minimal cluster of M and C_2 be the nonminimal cluster. Since $K \vdash \neg A$, all A-worlds in M lie in C_2 (we assume A is satisfiable). Now M^*_A has as a minimal cluster the set of A-worlds minimal in M. Since all A-worlds are located in C_2, the set $\|A\|$ in M forms the minimal cluster of M^*_A. Hence, any world can see w in M^*_A if $M \models_w A$, and v can see any world in M^*_A if $M \models_v \neg A$. Since M^*_A is a CO*-model, all logically possible A-worlds are in its minimal cluster, so $K^*_A = Cn(A)$. ■

Theorem 10 *Let M be a paranoid revision model for K, let $A_1, \ldots A_n$ be a revision sequence, and let $*$ be the paranoid revision function. If $K \not\vdash \neg(A_1 \wedge \cdots A_n)$ then $((M^*_{A_1})^*_{A_2} \cdots)^*_{A_n}$ is* *the (unique) paranoid revision model for*

$$K' = ((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n} = Cn(K \cup \{A_1, \cdots, A_i\})$$

If $K \vdash \neg(A_1 \wedge \cdots A_n)$ *then* $((M^*_{A_1})^*_{A_2} \cdots)^*_{A_n}$ *is the (unique) paranoid revision model for*

$$K' = ((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n} = Cn(\{A_i, A_{i+1}, \cdots, A_n\})$$

where $1 \leq i \leq n$ *is such that* $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_{i-1}} \vdash \neg A_i$ *and* $A_i \not\vdash \neg(A_{i+1} \wedge \cdots A_n)$.

Proof If $K \not\vdash \neg(A_1 \wedge \cdots A_n)$ then this part of the theorem holds immediately given Proposition 8. So suppose $K \vdash \neg(A_1 \wedge \cdots A_n)$. Then there must be some "first" inconsistent revision A_j in the sequence; that is, $K \not\vdash \neg(A_1 \wedge \cdots A_{j-1})$ but $K \vdash \neg(A_1 \wedge \cdots A_j)$. By Propositions 8 and 9 we have $K^*_{A_1 \wedge \cdots A_{j-1}} \vdash \neg A_j$. Since such an inconsistent revision exists, let A_i be the maximal element in this sequence with this property. We then have $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_{i-1}} \vdash \neg A_i$, but also $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_i} \not\vdash \neg(A_{i+1} \wedge \cdots A_n)$. By Proposition 9, $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_i} = Cn(A_i)$. Since A_i is the maximal inconsistent revision, Proposition 8 ensures that $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n} = Cn(\{A_i, A_{i+1}, \cdots, A_n\})$. ■

Theorem 13 *Let* M *be a revision model, let* $*$ *be the natural revision operator and let* $A, B, C \in \mathbf{L}_{CPL}$.
a) If $M^*_A \models \boxtimes\neg B$ *then* $M^*_A \models B \xrightarrow{\text{KB}} C$ *iff* $M \models B \xrightarrow{\text{KB}} C$.
b) If $M^*_A \not\models \boxtimes\neg B$ *then* $M^*_A \models B \xrightarrow{\text{KB}} C$ *iff* $M \models A \wedge B \xrightarrow{\text{KB}} C$.

Proof The proof of this theorem is sketched, for the most part, in the text preceding its statement in the body of the paper. ■

Theorem 23 *Let* $A_1, \ldots A_{n+1}$ *be an incompatible sequence such that* $A_1, \ldots A_n$ *is compatible. Let* k *be the maximal element of*

$$\{i \leq n : \neg A_{n+1} \notin ((K^*_{A_1})^*_{A_2} \cdots)^*_{A_i}\}$$

Then $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_{n+1}} = K^*_{A_1 \wedge \cdots A_k \wedge A_{n+1}}$.

Proof Let M be the revision model for K and $*$ the natural revision function. Let A_k be the last element of $A_1, \ldots A_{n+1}$ such that $\neg A_{n+1} \notin ((K^*_{A_1})^*_{A_2} \cdots)^*_{A_k}$.

a) If there is no such element then $K^*_{A_1} \vdash \neg A_{n+1}$. The minimal cluster of $M^*_{A_1}$ is formed by the set $\min(M, A_1)$. Since the relative ordering of all other worlds is unaffected by this revision, $\min(M^*_{A_1}, A_{n+1}) = \min(M, A_{n+1})$. Now, each subsequent revision is compatible, so $\min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_{i-1}}, A_i)$ is contained in the minimal cluster of $((M^*_{A_1})^*_{A_2} \cdots)^*_{A_{i-1}}$ for each $i \leq n$. An obvious inductive argument shows that (since no A_{n+1}-worlds are contained in the first minimal cluster)

$$\min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_n}, A_{n+1}) = \min(M, A_{n+1})$$

Hence $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_{n+1}} = K^*_{A_{n+1}}$.

b) If such a k exists the, by Proposition 19 and Theorem 20,

$$((K^*_{A_1})^*_{A_2} \cdots)^*_{A_k} = K^*_{A_1 \wedge \cdots A_k}$$

Furthermore, since $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_k} \nvdash \neg A_{n+1}$, the set $\min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_k}, A_{n+1})$ lies within the minimal cluster of $((M^*_{A_1})^*_{A_2} \cdots)^*_{A_k}$. Since the sequence $A_1, \ldots A_n$ is compatible, $\min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_k}, A_{k+1})$ must also lie within the minimal cluster of $((M^*_{A_1})^*_{A_2} \cdots)^*_{A_k}$. However, by the maximality of k, this set must be disjoint from the set $\min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_k}, A_{n+1})$. Thus,

$$\min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_{k+1}}, A_{n+1}) = \min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_k}, A_{n+1})$$

Since all subsequent revisions up to A_{n+1} are compatible, as in case a) above,

$$\min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_n}, A_{n+1}) = \min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_k}, A_{n+1})$$

But A_{n+1} is compatible with $A_1, \ldots A_k$, so

$$((K^*_{A_1})^*_{A_2} \cdots)^*_{A_{n+1}} = K^*_{A_1 \wedge \cdots A_k \wedge A_{n+1}}$$

■

Lemma 25 *Let $A_1, \ldots A_n$ be a revision sequence such that each proper subsequence $A_1, \ldots A_i$ is characterized by some sentence $s(A_i)$. Then $A_1, \ldots A_n$ is characterized by $s(A_k) \wedge A_n$, where A_k is the most recent compatible revision for A_n. In other words, $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n} = K^*_{s(A_k) \wedge A_n}$.*

Proof Let A_k be the most recent compatible revision for A_n. Since A_n is compatible with the subsequence $A_1, \ldots A_k$, we have that $\min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_k}, A_n)$ lies within the minimal cluster of $((M^*_{A_1})^*_{A_2} \cdots)^*_{A_k}$. However, each subsequent revision $A_{k+1}, \ldots A_{n-1}$ is incompatible with A_n so there can be no A_n-worlds in the set $\min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_{i-1}}, A_i)$ for $k < i < n$. Since only the relative status of these worlds is changed by these subsequent revisions, an obvious inductive argument (on the number $n - k$ of subsequent revisions) shows that $\min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_{n-1}}, A_n) = \min(((M^*_{A_1})^*_{A_2} \cdots)^*_{A_k}, A_n)$. Since $A_1, \ldots A_k$ is characterized by $s(A_k)$, the minimal cluster of $((M^*_{A_1})^*_{A_2} \cdots)^*_{A_k}$ is $\min(M, s(A_k))$. Clearly then the minimal cluster of $((M^*_{A_1})^*_{A_2} \cdots)^*_{A_n}$ is simply $\min(M, s(A_k \wedge A_n))$. Therefore $A_1, \ldots A_n$ is characterized by $s(A_k) \wedge A_n$. ■

Theorem 26 *For any revision sequence $A_1, \ldots A_n$, there is some subset of these updates $S \subseteq \{A_1, \cdots, A_n\}$ such that $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_n} = K^*_A$ and $A = \wedge S$.*

Proof This result can be shown using a simple inductive argument on n, the number of updates in the sequence. If $n = 1$ then the therorem obviously is true, for A_1 characterizes itself. Now suppose each subsequence $A_1, \ldots A_i$, $i < n$, is characterized by some sentence $s(A_i) = \wedge S$, where $S \subseteq \{A_1, \cdots, A_i\}$. By Lemma 25, the sequence $A_1, \ldots A_n$ is characterized by the sentence $s(A_k) \wedge A_n$ for some $k < n$, where A_k is the most recent compatible revision for A_n (or it is characterized by A_n if no such k exists). By the inductive hypothesis, $s(A_n) = \wedge S$ where $S \subseteq \{A_1, \cdots, A_k\} \cup \{A_n\} \subseteq \{A_1, \cdots, A_n\}$. (If $k = 0$ we simple observe that $S = \{A_n\} \subseteq \{A_1, \cdots, A_n\}$.) ■

Theorem 28 *Let $A_1, \ldots A_n$ be a revision sequence with c incompatible updates represented by σ. For each $1 \leq k \leq n$, if A_k is a compatible revision, then $((K^*_{A_1})^*_{A_2} \cdots)^*_{A_k} = K^*_{s(A_k)}$, where: a) $s(A_k) = s(A_{\sigma(i)}) \wedge A_{\sigma(i)+1} \wedge \cdots A_k$; b) $\sigma(i) < k < \sigma(i+1)$, or $\sigma(i) < k$ if $i = c$; and c) $s(A_{\sigma(i)})$ characterizes subsequence $A_1, \ldots A_{\sigma(i)}$.*

Proof Let A_k be a compatible revision in the sequence $A_1,\ldots A_n$. If no incompatible revision follows A_k, then $k > \sigma(i)$, where $i = c$. Otherwise $\sigma(i) < k < \sigma(i+1)$ for some $i < c$. In either case $A_{\sigma(i)}$ is the maximal incompatible revision in the subsequence $A_1,\ldots A_k$. Let $s(A_{\sigma(i)})$ characterize $A_1,\ldots A_{\sigma(i)}$. Since each revision $A_{\sigma(i)+1},\ldots A_k$ is compatible, A_{k-1} is the most recent compatible revision for A_k, and by $k-\sigma(i)$ applications of Lemma 25, we have that $A_1,\ldots A_k$ is characterized by $s(A_k) = s(A_{\sigma(i)}) \wedge A_{\sigma(i)+1} \wedge \cdots A_k$. ■

Theorem 30 *Let $A_1,\ldots A_n$ be a revision sequence with c incompatible updates represented by σ. For each $1 \leq k \leq n$, if A_k is an incompatible revision, then $((K^*_{A_1})^*_{A_2}\cdots)^*_{A_k} = K^*_{s(A_k)}$, where: a) $s(A_k) = s(A_{\sigma(i)}) \wedge A_{\sigma(i)+1} \wedge \cdots A_j \wedge A_k$; b) $A_{\sigma(i)}$ is the maximal incompatible revision for A_k; c) A_j is the most recent compatible revision for A_k; and d) $s(A_{\sigma(i)})$ characterizes subsequence $A_1,\ldots A_{\sigma(i)}$.*

Proof Let A_j be the most recent compatible revision for A_k. By Theorem 28 and Proposition 29, $A_1,\ldots A_j$ is characterized by $s(A_j) = s(A_{\sigma(i)}) \wedge A_{\sigma(i)+1} \wedge \cdots A_j$. By Lemma 25, $A_1,\ldots A_k$ is characterized by $s(A_j) \wedge A_k$. ■

Theorem 36 *Let M be a K-revision model and $A_1,\ldots A_n$ a revision sequence such that $A_2,\ldots A_n$ is a compatible sequence and $M^*_{A_1} \models A_2 \xrightarrow{\text{KB}} \neg A_1$. Then $(((K^*_{A_1})^*_{A_2}\cdots)^*_{A_n})^*_{A_1} = K^*_{A_1}$.*

Proof By the definition of M^*_A it is clear that $\min(M, A_1)$ is precisely the minimal cluster of $M^*_{A_1}$. Since $\min(M^*_{A_1}, A_2)$ contains no A_1-worlds, and each subsequent revision A_3, $\ldots A_n$ is compatible, it must be that

$$\min(((M^*_{A_1})^*_{A_2}\cdots)^*_{A_n}, A_1) = \min(M^*_{A_1}, A_1) = \min(M, A_1)$$

Thus $(((K^*_{A_1})^*_{A_2}\cdots)^*_{A_n})^*_{A_1} = K^*_{A_1}$. ■

Theorem 37 *Let $S_1, \cdots S_n \in \mathbf{L}_{CPL}$ be such that $\vdash S_i \supset \neg S_j$ if $i \neq j$, and $\vdash S_1 \vee \cdots S_n$. For all $A, B \in \mathbf{L}_{CPL}$ either*

$$FSM(S_1, \cdots S_n) \vdash_{CO*} A \xrightarrow{\text{KB}} B; \qquad \textit{or}$$
$$FSM(S_1, \cdots S_n) \vdash_{CO*} \neg(A \xrightarrow{\text{KB}} B)$$

Proof This follows immediately since the theory $FSM(S_1, \cdots S_n)$ is "categorical" in the sense that their is only one CO*-structure satisfying it (modulo "duplicate worlds", which can have no influence on the truth of any sentence in the model). ■

Theorem 38 *Let revision model M be characterized by $FSM(S_1, \cdots S_n)$ and let S_k be the minimal sentence in this set consistent with A; that is, $S_k \not\vdash \neg A$ and $S_i \vdash \neg A$ if $i < k$. Then M_A^* is characterized by*

$$\begin{array}{ll} FSM(S_k \wedge A, S_1, \cdots S_{k-1}, S_k \wedge \neg A, S_{k+1}, \cdots S_n) & \textit{if } S_k \not\vdash A \\ FSM(S_k, S_1, \cdots S_{k-1}, S_{k+1}, \cdots S_n) & \textit{if } S_k \vdash A \end{array}$$

Proof Clearly the minimal cluster in M containing A-worlds is that cluster specified by the sentence S_k. Thus the set $\min(M, A)$ consists exactly of those worlds in M satisfying $S_k \wedge A$. In M_A^*, this set forms the minimal cluster and all other clusters remain in the same relative order. However, the cluster that was specified by S_k is now reduced to those worlds satisfying $S_k \wedge \neg A$. If $S_k \vdash A$, this is still true; but the sentence $S_k \wedge A$ is equivalent to S_k, and the cluster in M_A^* satisfying $S_k \wedge \neg A$ is empty. ■