Reset reproduction of preprint distributed at the *Third International Workshop on Nonmonotonic Reasoning*, South Lake Tahoe, California, June 1990. Reprinted July 1994.

Rational Belief Revision*

(Preliminary Report)

Jon Doyle
Massachusetts Institute of Technology
Laboratory for Computer Science
545 Technology Square
Cambridge, Massachusetts 02139, USA

Abstract

Theories of rational belief revision recently proposed by Gärdenfors and Nebel illuminate many important issues but impose unnecessarily strong standards for correct revisions and make strong assumptions about what information is available to guide revisions. We reconstruct these theories according to an economic standard of rationality in which preferences are used to select among alternative possible revisions. By permitting multiple partial specifications of preferences in ways closely related to preference-based nonmonotonic logics, the reconstructed theory employs information closer to that available in practice and offers more flexible ways of selecting revisions. We formally compare this notion of rational belief revision with those of Gärdenfors and Nebel, adapt results about universal default theories to prove that there is no universal method of rational belief revision, and examine formally how different limitations on rationality affect belief revision.

Keywords: Belief revision, reason maintenance, rationality postulates, social choice theory, default reasoning, limited rationality, preference change.

1 Introduction

Philosophers of science and statisticians have long studied the problem of revising beliefs when actions are taken and information is acquired. Many of these studies view knowledge in probabilistic terms, phrasing the problem as one of revision of probability assessments, with Bayes's rule the central method (see, for example, [39, 21]). Recently these studies have been complemented by detailed studies by philosophers (e.g., [36, 28, 42, 20, 14]) and artificial intelligence researchers (e.g., [9, 17, 44, 6, 29, 32, 37]) of belief revision in a nonprobabilistic setting, in which one views the beliefs of an agent as a set of propositions and seeks to describe how a rational agent should change its set of believed propositions. Numerous practical systems for belief revision called reason maintenance or truth maintenance systems have also been developed in artificial intelligence (see [41, 8, 30, 7, 19]), and even some philosophers have implemented systems corresponding to their theories [34].

One of the best developed formal theories of belief revision is that expounded by Gärdenfors [14], based on work with Alchourrón and Makinson [1]. Gärdenfors models belief states with deductively closed sets of propositions (logical theories) and develops, motivates, and studies a small set of axioms that characterize how a rational agent should change its belief states when new beliefs are added, subtracted, or changed. These axioms require, for example, that changes to belief states result in new belief states; that changes are independent of the syntactic details of the specification; and that changes have the intended effects and no others (that is, they make as few consequential changes as possible). The central result of his theory is that these axioms are equivalent to the existence of an ordering of all propositions according to their degree of *epistemic entrenchment* such that belief revisions always retain more entrenched propositions in preference to less entrenched ones. They also imply the existence of a complete preordering of all belief states according to their degree of change.

Gärdenfors's theory of belief revision is not directly applicable to formalizing belief revision as practiced in artificial intelligence because it models states of belief as infinite, deductively closed sets of propositions. Computational approaches require that states of the reasoner be finitely representable as, for instance, finite sets of propositions, and usually view states of belief as the deductive closures of these finite sets. (Alternatively, the belief states may be closures under some weaker logic corresponding to the inferences feasible for the reasoner to compute. See [25, 26].) In Harman's [20] terminology, Gärdenfors's theory is a *coherence* theory as it requires only that belief states be internally coherent. In particular, it allows the agent to alter any belief directly as long as enough of its consequences are also altered to preserve the consistency and closure of the resulting set of beliefs. In contrast, Harman views most

AI approaches as *foundational* theories, which require that belief states be grounded in some set of beliefs that provides foundations or justifications for all other beliefs. A foundational theory restricts revisions to altering directly only those propositions appearing in the finite representation.

Because Gärdenfors's theory is a coherence theory, it cannot capture some revision techniques commonly used in AI. For example, one of the most common ways to revise a finite set of propositions to include a new one that is inconsistent with the existing propositions is first to find some maximal subset of the existing propositions consistent with the new one and then to make a new set consisting of the new proposition plus the selected maximal subset. But this "maxichoice" method has highly undesirable properties in the coherence theory, as it often produces complete sets of beliefs, even if the original set took no position on many questions. To better model AI practice, Nebel [32] adapted Gärdenfors's theory so that finite sets of representing propositions mediate revisions. He then shows that the resulting revisions satisfy most of Gärdenfors's rationality axioms. In fact, Nebel shows how to define certain orders over propositions and sets of propositions, representing notions of *epistemic relevance*, so that selecting revisions to be maximal in these orders yields finite revisions which satisfy *all* the rationality axioms. He also proves the very satisfying result that revision by picking maximal consistent subsets is, in the foundational case, just as rational as more elaborate forms of revision in which the reasoner retains only those beliefs following from all maximal consistent subsets. This makes mechanization of belief revision systems more practical since picking one consistent subset is much simpler than intersecting the consequences of all consistent subsets.

Unfortunately, even Nebel's theory does not completely succeed at formalizing practical notions of belief revision, even though it makes crucial improvements in Gärdenfors's theory. One problem lies in some notions retained from Gärdenfors's theory which undermine the applicability of Nebel's theory almost as much as infinite states do Gärdenfors's. Both theories suffer from unnecessarily strong standards for correct revisions, and from very strong assumptions about what information is available to guide revisions. Specifically, the sense of "rationality" postulated by these theories requires that rational revisions be unique (they do not permit equally acceptable alternative revisions). They also require total orderings of all propositions, even though most domains of knowledge formalized in AI systems are too incomplete and ambiguous to supply such complete orderings.

We show that these two problems are related by reconstructing the theory of belief revision according to an economic standard of rationality. Where the Gärdenfors axioms refer only to logical properties of revised sets of beliefs and are motivated in terms of logical coherence and conservation of beliefs, we make the underlying motivation more explicit and precise by identifying various preferences guiding revisions and viewing rational revision as choosing the new belief state to be of maximal preferability according to these revision preferences. This change has two major effects. First, since there may be several alternatives of maximal preferability, our theory does not make the strong assumption of unique revisions. Second, the enlarged framework is closer to AI practice than its predecessors since it demands only partial information about revision preferences and permits these partial preferences to be combined and used in more flexible ways.

The plan of the paper is as follows. Sections 2 and 3 summarize the formal theories and central results of Gärdenfors and Nebel in turn. Section 4 introduces the formal theory of rational revision in the economic sense and studies its impact on Gärdenfors's axioms. Determinism aside, rational revision turns out to be quite different from Gärdenfors notion, which makes the plausible but sometimes overly special assumption that believing more is always better. Section 5 examines the informational requirements of rational revision and Gärdenfors's and Nebel's theories. We argue that in practice the preferences available to guide belief revision consist of numerous partial and sometimes conflicting preferences, and show how the orderings assumed by Gärdenfors and Nebel are too inflexible for many applications. We then present the formal theory of rational revision guided by multiple partial preferences, which is based on several principles for rationally aggregating partial preferences. This theory is formally similar to Doyle and Wellman's [12] theory of rational default reasoning and to the economic theory of social choice [3]. We prove that no method for belief revision based on partial preferences satisfies all the rationality conditions on preference aggregation. We examine how irrationalities in preference aggregation lead to violations of Gärdenfors's rationality axioms. Section 6 expands the theory further to cover rational revision of preferences. This final expansion, which is related to our earlier theory of constructive attitudes [11], introduces further ways in which rational belief revision can diverge from Gärdenfors's notion.

2 Revising belief states

Gärdenfors's formalization of belief revision may be summarized as follows, using an adaptation of Nebel's notation. We suppose that $\mathcal{L}$ is a propositional language over the standard sentential connectives ($\neg$, $\wedge$, $\vee$, $\rightarrow$, $\leftrightarrow$), denote individual propositions by x, y, and z, and denote sets of propositions by A, B, and C. We write $\vdash$ to mean classical propositional derivability, and write Cn to mean the corresponding closure operator

$$Cn(A) \stackrel{\text{def}}{=} \{x \in \mathcal{L} \mid A \vdash x\}. \qquad (1)$$

The theory also makes sense if $\vdash$ is taken to be weaker or stronger than classical derivability, for example, by corresponding to some set of limited deduction rules or to entailment with respect to some background theory.

Gärdenfors models states of belief by propositional theories; that is, deductively closed sets of propositions. Thus each set $A \subseteq \mathcal{L}$ such that $A = Cn(A)$ corresponds to a belief state. (The inconsistent set $\mathcal{L}$ is itself a belief state, but will not play a significant role in the following.)

2.1 Expansions, contractions, and revisions

Gärdenfors considers three types of operations on belief states. For each belief state A and proposition x we have:

Expansion: Expanding A with x, written $A+x$, means adding x to A and requiring that the result be a (possibly inconsistent) belief state.

Contraction: Contracting A with respect to x, written $A \dot{-} x$, means removing x from A in such a way to result in a belief state.

Revision: Revising A with x, written $A \dot{+} x$, means adding x to A in such a way that the result is a consistent belief set.

Expansion is naturally defined as the closure of the union of the belief set and the new proposition:

$$A + x \stackrel{\text{def}}{=} Cn(A \cup \{x\}) \tag{2}$$

Contraction and revision, on the other hand, have no single natural definitions, only the standard requirement that the change made be as small as possible so as to minimize unnecessary loss of knowledge. (Quine [35] calls this requirement "minimum mutilation;" Harman [20] calls it "conservativity.") This requirement does not define these operations since there are usually several ways to get rid of some belief. In the case of contraction, for example, there is generally no largest belief set $B \subseteq A$ such that $B \nvdash x$.

Prevented from identifying unique natural contraction and revision operations, Gärdenfors formulates and motivates sets of rationality postulates that these operations should satisfy. The axioms for rational contractions are as follows. For each belief set A and propositions x and y, we have:

($\dot{-}$1) $A \dot{-} x$ is a belief set; *(closure)*

($\dot{-}$2) $A \dot{-} x \subseteq A$; *(inclusion)*

($\dot{-}$3) If $x \notin A$, then $A \dot{-} x = A$; *(vacuity)*

($\dot{-}$4) If $\nvdash x$, then $x \notin (A \dot{-} x)$; *(success)*

($\dot{-}$5) If $\vdash x \leftrightarrow y$, then $A \dot{-} x = A \dot{-} y$; *(equivalence)*

($\dot{-}$6) $A \subseteq (A \dot{-} x) + x$; *(recovery)*

($\dot{-}$7) $(A \dot{-} x) \cap (A \dot{-} y) \subseteq A \dot{-} (x \wedge y)$;

($\dot{-}$8) If $x \notin A \dot{-} (x \wedge y)$, then $A \dot{-} (x \wedge y) \subseteq A \dot{-} x$.

The *closure* axiom ($\dot{-}$1) says that the result of contraction is always a belief set. The *inclusion*, *vacuity*, and *success* postulates ($\dot{-}$2)-($\dot{-}$4) state that removing a proposition does not introduce any new propositions, that no change occurs when one tries to remove a proposition that is not already present, and that removed propositions are indeed removed, unless they are tautologies (in which case they are present in every belief set, and so cannot be removed). The *equivalence* axiom ($\dot{-}$5) states that the results of contraction do not depend on the syntactic form of the proposition removed; removing any logically equivalent proposition has the same effect. The *recovery* postulate ($\dot{-}$6) states a version of the conservativity principle by requiring that contraction with respect to a proposition removes nothing except some of the consequences of the proposition; that is, anything removed can be recovered by adding the proposition back in. Axioms ($\dot{-}$7) and ($\dot{-}$8) relate contracting with respect to a conjunction to the contractions with respect to the individual conjuncts; they imply that retracting a conjunction preserves more information than retracting both of its conjuncts simultaneously. If contraction satisfies ($\dot{-}$1)-($\dot{-}$6), then ($\dot{-}$7) and ($\dot{-}$8) taken together are equivalent to the "factoring" condition (3).

$$A \dot{-} (x \wedge y) = \begin{cases} (A \dot{-} x) \cap (A \dot{-} y) & \text{or} \\ A \dot{-} x & \text{or} \\ A \dot{-} y \end{cases} \tag{3}$$

Gärdenfors[14] also develops a parallel set of axioms for revisions and proves that they are logically equivalent to the contraction postulates if the revision $A \dot{+} x$ is defined by means of the *Levi identity* (after [27])

$$A \dot{+} x \stackrel{\text{def}}{=} (A \dot{-} \neg x) + x, \tag{4}$$

so that revision by x is equivalent to contracting by $\neg x$ to remove any inconsistent beliefs and then expanding with x. One can also define contractions in terms of revisions by means of the *Harper identity* (after [22])

$$A \dot{-} x \stackrel{\text{def}}{=} (A \dot{+} \neg x) \cap A, \tag{5}$$

so that the contraction by x is equivalent to taking those beliefs that would be preserved if $\neg x$ were now believed.

2.2 Epistemic entrenchment

Though Gärdenfors first characterizes rational revisions by means of the preceding axioms, he views the behaviors these axioms describe as arising from a more fundamental notion, that of *epistemic entrenchment* [14, 16], which is related to the notion of database priorities in [13]. Epistemic entrenchment is characterized by a complete preorder (a reflexive and transitive relation) over propositions which indicates which propositions are more valuable than others. This ordering, which may vary from belief state to belief state, influences revisions by the requirement that revisions retain more entrenched beliefs in preference to less entrenched ones.

If x and y are propositions, we write $x \leq y$ to mean that y is at least as epistemically entrenched as x. We define the strict part of this order, $x < y$, which means that y is more entrenched than x, by the conjunction of $x \leq y$ and $y \not\leq x$. Gärdenfors characterizes the qualitative structure of this order with the following axioms.

($\leq$1) If $x \leq y$ and $y \leq z$, then $x \leq z$; *(transitivity)*

($\leq$2) If $x \vdash y$, then $x \leq y$; *(dominance)*

($\leq$3) Either $x \leq x \wedge y$ or $y \leq x \wedge y$; *(conjunctiveness)*

($\leq$4) If A is consistent, then $x \leq y$ for all y iff $x \notin A$; [1in] *(minimality)*

($\leq$5) If $x \leq y$ for all x, then $\vdash y$. *(maximality)*

Axiom ($\leq$1) just says that $\leq$ is an ordering relation, while the other axioms all concern how the logic of propositions interacts with the ordering. Postulate ($\leq$2) says that x entails y, then retracting x is a smaller change than retracting y, since the closure requirement on belief states means that y cannot be retracted without giving up x as well. Axiom ($\leq$3) reflects the fact that a conjunction cannot be retracted without giving up at least one of its conjuncts. Taken together axioms ($\leq$1)-($\leq$3) imply that $\leq$ is a complete ordering, that is, that either $x \leq y$ or $y \leq x$. Propositions not in a belief state are minimally entrenched in that state, according to ($\leq$4), and according to ($\leq$5), the only way a proposition can be maximally entrenched is if it is logically valid.

The influence of epistemic entrenchment on belief revisions is characterized by two conditions relating entrenchment orderings and contraction functions. The first condition,

$$x \leq y \text{ iff either } x \notin A \dot{-} (x \wedge y) \text{ or } \vdash x \wedge y, \quad (6)$$

says that in contracting a belief state with respect to a conjunction, we must give up the conjunct of lesser epistemic entrenchment, or both conjuncts if they are equally entrenched. It says, in essence, that $x < y$ is the same as $y \in A \dot{-} (x \wedge y)$. The second condition,

$$y \in A \dot{-} x \text{ iff } y \in A \text{ and either } x < x \vee y \text{ or } \vdash x, \quad (7)$$

explicitly characterizes contraction functions in terms of epistemic entrenchment orderings.

The main result of the theory of epistemic entrenchment is that this notion is essentially equivalent to the previously axiomatized notion of rational revision. Gärdenfors and Makinson [16] prove this equivalence in two parts. They first prove that rational contraction functions may be constructed from orderings of epistemic entrenchment, specifically, that if an ordering $\leq$ satisfies ($\leq$1)-($\leq$5), then the contraction function uniquely determined by the contraction condition (6) satisfies ($\dot{-}$1)-($\dot{-}$8) as well as (7). They then prove that entrenchment orderings may be constructed from rational contraction functions, specifically, that if a contraction function $\dot{-}$ satisfies ($\dot{-}$1)-($\dot{-}$8), then the ordering $\leq$ uniquely determined by (6) satisfies ($\leq$1)-($\leq$5) as well as (7).

2.3 Some important contraction functions

The central construct in many studies of contractions and revisions is the set of all maximal subsets of a belief set A consistent with a proposition x, which we write as $A \downarrow x$ and read as "A less x." Formally, we have

$$A \downarrow x \stackrel{\text{def}}{=} \{B \subseteq A \mid B \nvdash x \wedge (B \subset C \subseteq A) \rightarrow C \vdash x\}. \quad (8)$$

It is easy to see that elements of $A \downarrow x$ are themselves belief sets.

Using this construct, one can define the contraction $A \dot{-} x$ as the set of beliefs in either one, some, or all states in $A \downarrow x$. These definitions correspond to Gärdenfors's notions of maxichoice, partial meet, and full meet contractions.

2.3.1 Maxichoice contraction

Maxichoice contraction is contraction to one maximal consistent subset. If there are no alternatives, then x must be logically valid and impossible to retract, so we may take A itself as the contraction. Formally, we assume the existence of a choice function $\mathcal{C}$ which selects one of the elements of $A \downarrow x$ and define the full meet contraction operation $\stackrel{\text{m}}{-}$ by

$$A \stackrel{\text{m}}{-} x \stackrel{\text{def}}{=} \begin{cases} \mathcal{C}(A \downarrow x) & \text{if } \nvdash x \\ A & \text{otherwise.} \end{cases} \quad (9)$$

Gärdenfors [14, Lemma 4.1] shows that this operation satisfies ($\dot{-}$1)-($\dot{-}$6), but not necessarily ($\dot{-}$7) and ($\dot{-}$8). However, failure to satisfy these last axioms is the least of the problems with maxichoice contraction, for one may prove that using maxichoice contraction to effect the revision $A \dot{+} x$ via (4) makes $A \dot{+} x$ a *complete* belief set as long as $\neg x \in A$ [14, Corollary 4.6]. That is, $A \dot{+} x$ in this case contains either y or $\neg y$ for each proposition y. Since A need not have contained either y or $\neg y$, these new beliefs are clearly gratuitous.

2.3.2 Full meet contraction

Full meet contraction revises beliefs by retaining only those propositions that appear in *every* alternative in $A \downarrow x$. This simply means taking the intersection of these states. We define the full meet contraction operation $\stackrel{\text{f}}{-}$ formally by

$$A \stackrel{\text{f}}{-} x \stackrel{\text{def}}{=} \begin{cases} \bigcap(A \downarrow x) & \text{if } \nvdash x \\ A & \text{otherwise.} \end{cases} \quad (10)$$

Full meet contraction, unlike maxichoice contraction, satisfies all the rationality postulates ($\dot{-}$1)-($\dot{-}$8) [32, Lemma 4].

Unfortunately, full meet contraction gives up too many beliefs to be very useful. In particular, when one uses full meet contraction to effect revisions, one can easily wind up with nothing left except the consequences of the revised belief alone. Specifically, if $\dot{+}$

is a revision operation defined by (4) and (10) and if $\neg x \in A$ and $\nvdash \neg x$, then

$$A \dot{+} x = Cn(\{x\}). \tag{11}$$

(See [2] or [32, Corollary 3].) Taken together, these results about maxichoice and full meet contraction indicate that the notion of rational contraction captured by axioms ($\dot{-}$1)-($\dot{-}$8) and (4) is really very weak.

2.3.3 Partial meet contraction

If the extremes of defining the contraction $A \dot{-} x$ by the propositions in one or all of the states in $A \downarrow x$ are unacceptable, better results might be obtained by using some intermediate subset of $A \downarrow x$. We assume a selection function $\mathcal{S}$ which selects such subsets, and define the operation of *partial meet* contraction $A \stackrel{\text{p}}{-} x$ by

$$A \stackrel{\text{p}}{-} x \stackrel{\text{def}}{=} \begin{cases} \bigcap \mathcal{S}(A \downarrow x) & \text{if } \nvdash x \\ A & \text{otherwise.} \end{cases} \tag{12}$$

Partial meet contraction satisfies the basic rationality postulates ($\dot{-}$1)-($\dot{-}$6), and in fact is equivalent to them in the sense that any operation satisfying these axioms is a partial meet contraction operation (see [14, Theorem 4.13]).

2.3.4 Relational contraction

Some partial meet contractions, namely those derived from orderings of belief states are fully rational, satisfying ($\dot{-}$1)-($\dot{-}$8). We say that the contraction operation is *relational* if there is a binary relation $\sqsubseteq$ over $A \downarrow x$ such that the selected subsets are exactly the $\sqsubseteq$-"maximal" subsets, that is, if for all $B, C \subseteq A$,

$$B \sqsubseteq C \text{ iff } B \in A \downarrow x \text{ and } C \in \mathcal{S}(A \downarrow x). \tag{13}$$

This gives us a way of constructing a selection function, and hence a partial meet contraction relation, from every relation $\sqsubseteq$ over $A \downarrow x$. The most important case is that when $\sqsubseteq$ is a transitive relation. Gärdenfors [14, Theorem 4.16] proves that if $\sqsubseteq$ is transitive, the contraction function defined by (12) and (13) satisfies ($\dot{-}$1)-($\dot{-}$8).

Since rational contraction functions may be constructed from either epistemic entrenchment orderings $\leq$ of propositions or from transitive relations $\sqsubseteq$ over substates of belief states, it is natural to ask if there is some special connection between these orders. Gärdenfors [14] views $\sqsubseteq$ as comparing the epistemic entrenchment of sets of propositions, but makes no explicit connection between the two notions. One obvious difference is that $\leq$ is a total order, while $\sqsubseteq$ may be partial; but this difference need not be important, since Gärdenfors also proves that for each transitive relation $\sqsubseteq$ there is a transitive total order $\sqsubseteq'$ which yields the same contraction function as $\sqsubseteq$ [14, Theorem 4.17]. That is, if a contraction function satisfies ($\dot{-}$1)-($\dot{-}$8), there is a complete preordering of belief states such that contraction is partial meet contraction with respect to this total ordering.

3 Revising belief bases

As noted earlier Gärdenfors's theory is not directly applicable to artificial intelligence since it presumes infinite belief states while AI works with finite representations of belief states. Nebel [32] modified Gärdenfors's theory by modeling finite representations of belief states as finite sets of propositions, called *belief bases*, with each belief base $B \subseteq \mathcal{L}$ representing the belief set $Cn(B)$. (Of course, not all belief sets are finitely representable, and the representation is never unique.)

3.1 Belief base contraction

Nebel defines belief base contraction, which we write as $\ominus$, by

$$B \ominus x \stackrel{\text{def}}{=} \begin{cases} (B \vee \neg x) \wedge \bigvee_{C \in (B \downarrow x)} C & \text{if } \nvdash x \\ B & \text{otherwise,} \end{cases} \tag{14}$$

and belief base revision ($\oplus$) by

$$B \oplus x \stackrel{\text{def}}{=} (B \ominus \neg x) \wedge x. \tag{15}$$

We may view $\ominus$ as an "implementation" of $\dot{-}$ by lifting base revision to the belief state level. That is, we may use the contraction postulates to judge base contraction by identifying A with $Cn(B)$ and identifying $A \dot{-} x$ with $Cn(B \ominus x)$. With these identifications, Nebel [32, Lemma 11] proves that belief base contraction satisfies ($\dot{-}$1)-($\dot{-}$6).

Nebel proceeds to show that belief base contraction is a form of partial meet contraction when lifted to belief states. Let $\mathcal{S}_B$ be a selection function corresponding to belief set B and defined by

$$\begin{aligned} &\mathcal{S}_B(Cn(B) \downarrow x) \stackrel{\text{def}}{=} \\ &\quad \{C \in (Cn(B) \downarrow x) \mid \forall C' \in (Cn(B) \downarrow x) \\ &\qquad\qquad C' \cap B \not\supset C \cap B\}. \end{aligned} \tag{16}$$

Nebel [32, Theorem 14] then proves that contraction of finite premise sets B using $\ominus$ is identical (with respect to $\vdash$) to a partial meet contraction $\stackrel{\text{p}}{-}$ defined by the selection function $\mathcal{S}_B$, that is,

$$Cn(B \ominus x) = Cn(B) \stackrel{\text{p}}{-} x. \tag{17}$$

This guarantees that the lifted version of $\ominus$ satisfies ($\dot{-}$1)-($\dot{-}$6). By defining the order $\sqsubseteq$ by

$$X \sqsubseteq Y \text{ iff } X \cap B \not\supset Y \cap B, \tag{18}$$

Nebel proves that partial meet contraction using $\mathcal{S}_B$ satisfies ($\dot{-}$7) as well [32, Theorem 15]. Since the order $\sqsubseteq$ is not transitive in general, this definition of contraction need not satisfy ($\dot{-}$8).

The definitions of belief base contraction and revision have much to recommend them as characterizations of the effect of the contraction or revision on the *content* of the agent's beliefs. But they are less satisfying as characterizations of how contraction and revision affect the *form* of the agent's representation of its beliefs because iterated contraction of belief bases does not always make much sense. While the operations $\dot{-}$ and $\dot{+}$ take belief states into other belief states of the same complexity, the operations $\ominus$ and $\oplus$ sometimes replace all previous base propositions with a single new proposition that was not an element of of the previous set. Subsequent contraction may then discard this singular residue of the original belief base. Nebel discusses two approaches to ameliorating this difficulty, namely factoring a subset of the original beliefs out of the disjunction resulting from contraction, and representing the new state as a set of belief bases. Unfortunately, each of these leads to belief bases whose sizes grow exponentially as more and more contractions are performed. In the end, Nebel leaves unsolved the practical problem of how to represent general belief base contractions.

3.2 Epistemic relevance

While one may apply the notion of epistemic entrenchment in the foundational view of belief revision, this approach is not always practical (but see [15]). The dominance axiom ($\leq$2), which requires that the logical consequences of a belief be at least as epistemically entrenched as the belief itself, means that to determine whether one proposition is more entrenched than another may require determining whether it entails the other, and since entailment is not decidable, this may not be possible. Nebel [33, pp. 162-166] introduces the notion of *epistemic relevance* as an analogue of epistemic entrenchment more suited to the needs of computational belief revision. The basic idea is to view the syntactical form of the belief base as indicating or determining which propositions are *relevant* to the agent's purposes, and to view the propositions not in the belief base as irrelevant. The guideline for belief revision is then to minimize the loss of epistemically relevant propositions. Nebel proposes the selection function $\mathcal{S}_B$ defined in (16) as a formalization of this notion.

Nebel goes on to enlarge the conception of epistemic relevance from a simple binary distinction to a total preordering $\leq_\rho$ of the propositions in B, where $x \leq_\rho y$ means that y is at least as relevant as x. He shows how such orderings of epistemic relevance can be used to embed $\sqsubseteq$ in a transitive total ordering $\sqsubseteq_\rho$ over subsets of B, defined by

$$X \sqsubseteq_\rho Y \text{ iff } \forall x \in (X - Y)\ \exists y \in (Y - X) \quad x \leq_\rho y. \tag{19}$$

This order ranks X less relevant than Y just in case X's most relevant elements are less relevant than Y's most relevant elements, ignoring the elements X and Y have in common. Since all propositions are totally ordered by $\leq_\rho$, $\sqsubseteq_\rho$ totally orders all subsets of B. Furthermore, since over subsets of B it is clear that $X \sqsubset Y$ just means $X \subset Y$, we see that $X \sqsubset_\rho Y$ holds trivially if $X \sqsubset Y$, so the new order extends the old one. Nebel then defines a new selection function $\mathcal{S}_{B,\leq_\rho}$ in terms of $\sqsubseteq_\rho$ by

$$\begin{aligned} &\mathcal{S}_{B,\leq_\rho}(A \downarrow x) \stackrel{\text{def}}{=} \\ &\quad \{C \in (A \downarrow x) \mid \forall C' \in (A \downarrow x) \\ &\qquad\qquad C' \cap B \sqsubseteq_\rho C \cap B\} \end{aligned} \tag{20}$$

and proves [32, Theorem 16] that partial meet contraction defined by this selection function gives rise to a fully rational contraction function satisfying ($\dot{-}$1)-($\dot{-}$8).

If $\leq_\rho$ is a linear ordering (no ties allowed), then every finite set of beliefs has a most relevant element, and we may rewrite the definition of $\sqsubseteq_\rho$ as

$$X \sqsubseteq_\rho Y \text{ iff } \max(X - Y) \leq_\rho \max(Y - X). \tag{21}$$

In this case, $\sqsubseteq_\rho$ always singles out the greatest element of $(B \downarrow x)$, and the partial meet contraction function defined using $\mathcal{S}_{B,\leq_\rho}$ resembles a maxichoice contraction on the belief base (see [32, Lemma 17]). Nebel concludes from this that maxichoice contraction on belief bases does not have the undesirable completion behavior exhibited by maxichoice contraction of belief states. Moreover, maxichoice base contraction is just as rational as the more complicated base contraction. Since it can be iterated easily and does not introduce complicated disjunctions, it is to be preferred in practice. He points to McAllester's [31] RUP system as an implementation of just this sort of rational belief base revision (ignoring the issue of RUP's deductive incompleteness).

3.3 Epistemic relevance vs. epistemic entrenchment

The orders of epistemic relevance and epistemic entrenchment are superficially different in that epistemic entrenchment orders all propositions in $\mathcal{L}$ while $\leq_\rho$ may order only the propositions in a belief set B. But the more fundamental difference, as Nebel notes, is that epistemic entrenchment is restricted in how it compares logically related propositions, while $\leq_\rho$ is an arbitrary ordering of propositions that may ignore logical dependencies. All restrictions of $\leq$ to B are of course acceptable as definitions of $\leq_\rho$, but no converse implication is true since $\leq_\rho$ need satisfy none of ($\leq$2)-($\leq$5).

There may be more interesting things one can say about the relation between these two concepts. For example, since $\mathcal{S}_{B,\leq_\rho}$ determines a rational contraction function, it also determines an entrenchment order $\leq$. What is the relation between this $\leq$ and $\leq_\rho$? If $\leq_\rho$ is the restriction of some entrenchment ordering $\leq'$ to B, is the determined order $\leq$ identical to $\leq'$ on B?

4 Choosing revisions rationally

Gärdenfors intends his principles of rational belief revision to capture logical constraints on revisions and contractions, as opposed to pragmatic influences. Accordingly, his axioms for contraction functions concern mainly logical relationships among beliefs. But it is not clear that there are any constraints free of pragmatic motivation. In particular, the whole motivation for conserving as many beliefs as possible seems entirely nonlogical. Conservatism has nothing to do with the consistency or completeness of beliefs or the soundness of inferences, which are the only characteristics of concern to logic. As Gärdenfors [15] acknowledges, the motivations for conservatism are instead economic: beliefs are valuable (useful in acting, costly to infer or acquire), so getting rid of beliefs unnecessarily is irrational. This economic motivation for conservatism may not be pragmatic in the sense of being specific to some particular reasoning situation, but it is certainly not logical. Take away these economic motivations for conservatism, however, and the only ones of Gärdenfors's axioms that seem logically motivated are ($\dot{-}1$), that the results of contraction be logically acceptable, and ($\dot{-}3$), that taking away absent propositions has no effect.[1]

If we really desire a theory of rational belief revision, we must expand the notion of rationality from the purely logical sense, in which one is rational if one's beliefs are consistent and one's inferences are sound, to the economic sense, in which one is rational if one has a consistent set of preferences and makes choices that are optimal with respect to these preferences. This means the theory should use preferences about revisions in guiding revisions, and draw on the standard theory of economic rationality to formalize the notions of rational contraction and revision. In particular, the theory should take seriously the variability of the costs and benefits of beliefs. If the theory is to be general and cover all reasoning, it should allow preferences to vary with the reasoning situation and task, and should not presuppose any special measures of costs and benefits.

4.1 Economic rationality

According to the normative theory of economic rationality, rational agents choose maximally preferred alternatives. Preferences may stem from many different motivations, such as computational costs or moral principles, but all the formal theory requires is that a preference ordering $\precsim$ is a complete, reflexive, and transitive relation that represents the agent's judgments of relative overall preferability (or *utility*) of possible alternatives. That is, $\precsim$ satisfies, for all alternatives X, Y, and Z, the axioms

($\precsim$1) Either $X \precsim Y$ or $Y \precsim X$, and *(completeness)*

($\precsim$2) If $X \precsim Y$ and $Y \precsim Z$, then $X \precsim Z$. *(transitivity)*

We use $\prec$ to denote *strict preference*, the asymmetric part of the preference order, and $\sim$ to indicate *indifference*, the reflexive part of the order. Thus $X \prec Y$ iff $X \precsim Y$ and $Y \not\precsim X$, and $X \sim Y$ iff $X \precsim Y$ and $Y \precsim X$. The combined order $\precsim$ is sometimes called *weak preference*. We assume that the agent's preference ordering may change with its state.

4.2 Epistemic preferences

If belief revision is to be rational in the economic sense, then preferences about beliefs or belief states must play a role. Since belief revision involves moving from one belief state to another, the preferences involved directly must be preferences among belief states, not preferences among individual beliefs. Even if preferences among individual beliefs constitute the fundamental information employed, the belief revision process must embody some way of constructing preferences among belief states from preferences among beliefs.

For example, Gärdenfors's theory contains two orders, $\leq$ and $\sqsubseteq$, over beliefs and belief states, respectively. Each order $\leq$ is complete and transitive, and so satisfies the axioms for preferences over propositions. Moreover, belief revision that satisfies ($\dot{-}1$)-($\dot{-}8$) corresponds, as Gärdenfors shows, to partial meet contraction in which the selection function picks out the maxima of a complete preorder $\sqsubseteq$ over belief states. Thus $\sqsubseteq$ can always be chosen to satisfy the axioms for preferences over belief states. The equivalence of the axioms for rational contraction and epistemic entrenchment mean that these orders are derivable from each other, so that the order $\sqsubseteq$ indicates degrees of overall epistemic entrenchment of belief states. Similarly, Nebel's orders of epistemic relevance satisfy the axioms for preferences over propositions and extend, via his constructions, to preference orders over belief states.

4.3 Rationality and skepticism

If we wish contraction based on epistemic entrenchment or epistemic relevance to be rational in the logical sense, then the resulting belief states must be maximal with respect to these orderings of belief states. That is not usually the case, however. First consider epistemic entrenchment. The maximally preferred states in $A \downarrow x$ are just the states in $\mathcal{S}(A \downarrow x)$. But partial meet contraction makes the new belief state be the intersection of these, and this intersection is generally not itself a preferred belief state, that is, generally we have $\bigcap \mathcal{S}(A \downarrow x) \notin \mathcal{S}(A \downarrow x)$. For example, if we judge

[1] One can raise doubts about these two as well. The whole notion of belief state idealizes far beyond what seems reasonable for agents of limited cognitive abilities, whose explicit beliefs (in the sense of Levesque [26]) might not be closed or consistent, calling ($\dot{-}1$) into question. One might doubt ($\dot{-}3$), for while it may make sense in computational systems to consider contractions and revisions with respect to individual propositions, humans cannot be instructed to perform these operations without also informing them about other things, such as the fact of the instruction itself. Indeed, telling someone to not believe something about which they previously had no opinion is often a sure-fire way to lead them to investigate the question, possibly coming to believe the proposition after all. Similarly, one might doubt ($\dot{-}4$). Just because someone is told to abandon something does not mean they will agree to do so, especially if they strongly prefer to believe it to not believing it.

the preferability of belief states along numerous dimensions, each of the maximal belief states will be better along some dimensions and worse along others, and the intersection will be nonoptimal along as many dimensions as there are alternatives. The problem is clearest when examined in terms of propositional entrenchment. Consider the contraction $A \dot{-} (x \wedge y)$, and suppose that $x \leq y$, $\nvdash x, y$ and $x, y \in A$. If $x < y$, then x is given up, while if $y \leq x$, both x and y are given up. There is no way to take the entrenchment-equivalence of x and y as a license to give up just one, chosen indifferently. Instead, revisions based on epistemic entrenchment are fundamentally "skeptical" in the sense that ambiguities are always resolved by refusing to believe any of the possibilities. The same "skepticism" appears when revisions employ Nebel's epistemic relevance orders instead of epistemic entrenchment.

The skepticism of partial meet contraction is similar to the skepticism of certain theories of inheritance and default reasoning, in which the conclusions of a theory are defined to be those which hold in every maximally preferred interpretation (see [23, 40]).[2] The difficulty is that skepticism is not always rational. The agent cannot always rationally choose to remain skeptical about questions concerning actions that are very important to an agent's prosperity. For some issues, it may be better to adopt a stance and risk error than to take no stance at all and risk paralysis, as in the fabulous tale of the man forced, under threat of death, to choose between doors containing either a beautiful woman or a ravenous tiger. Here skepticism guarantees death, while taking a stand offers a chance at life.

We conclude that if belief revision is to be rational, the actual belief states obtained in contraction and revision must be rational choices. Rather than increasing rationality, the skepticism resulting from choosing intersections of preferred belief states usually results in choosing suboptimal belief states.

4.4 Rationality and determinism

The first consequence of abandoning the skeptical approach to belief revision is that there is no longer any epistemological reason from logic or economics to suppose that contraction or revision are deterministic operations, as is presupposed in Gärdenfors's theory by viewing contraction and revision as functions taking belief states into belief states. Rationality does not prevent belief revision from being functional in deterministic agents, since every revision performed may be optimal with respect to the logically possible revisions, even though only one of these is constitutionally possible for the deterministic agent itself. But determinism is in no sense a logical constraint on belief revision.

Accordingly, we expand the formalization of contraction and revision to cover the cases in which contraction and revision are correspondences, that is, set-valued functions taking belief states into sets of belief states. We write $\hat{-}$ to denote rational contraction, and $\hat{+}$ to denote rational revision. For example, if $\nvdash x$ we might define the rational contraction $A \hat{-} x$ to be $\mathcal{S}(A \downarrow x)$ if the selection function $\mathcal{S}$ is based on a preference order. Of course, we can always construct skeptical deterministic contraction and revision functions from nondeterministic ones by intersection, that is, by defining $A \dot{-} x$ to be $\bigcap(A \hat{-} x)$.

4.5 Rational contraction and revision

Even if $\bigcap \mathcal{S}(A \downarrow x) \in \mathcal{S}(A \downarrow x)$, so that the intersection of the preferred alternatives is itself a preferred alternative, partial meet contraction still need not be rational because the range of alternatives is restricted to sets in $A \downarrow x$. That is, if $\sqsubseteq$ represents preferences over belief states, there is no reason *a priori* why its maxima should be found only among the sets in $A \downarrow x$ rather than among all subsets of $A - \{x\}$. Accordingly, we do not assume that contractions are restricted to only the maximal consistent subsets $A \downarrow x$, but allow contraction to select among all belief states generated by subsets of A. More precisely, we define $A \Downarrow x$, read "A without x," by

$$A \Downarrow x \stackrel{\text{def}}{=} \{X \subseteq A \mid X \nvdash x\}. \tag{22}$$

Clearly, $A \downarrow x \subseteq A \Downarrow x$ for all A and x. Usually we will be interested in the closures of these consistent subsets, which we indicate by $A \Downarrow^* x$, defined by

$$A \Downarrow^* x \stackrel{\text{def}}{=} \{Cn(X) \mid X \in A \Downarrow x\}. \tag{23}$$

We then define rational contraction by

$$A \hat{-} x \stackrel{\text{def}}{=} \begin{cases} \{A\} & \text{if } \vdash x \\ \{A\} & \text{if } A \hat{-}' x = \emptyset \\ A \hat{-}' x & \text{otherwise} \end{cases} \tag{24}$$

where

$$A \hat{-}' x \stackrel{\text{def}}{=} \{X \in (A \Downarrow^* x) \mid \forall X' \in (A \Downarrow^* x) \quad X' \precsim X\}.$$

Similarly, we define rational revision by

$$A \hat{+} x \stackrel{\text{def}}{=} \begin{cases} \{\mathcal{L}\} & \text{if } \vdash \neg x \\ \{\mathcal{L}\} & \text{if } A \hat{+}' x = \emptyset \\ A \hat{+}' x & \text{otherwise} \end{cases} \tag{25}$$

where

$$A \hat{+}' x \stackrel{\text{def}}{=} \{X + x \mid X \in (A \Downarrow^* \neg x) \wedge \forall X' \in (A \Downarrow^* \neg x) \quad X' + x \precsim X + x\}.$$

[2] Both belief revision and default reasoning are forms of nonmonotonic reasoning. Belief revision is *temporally* nonmonotonic in that the set of beliefs need not increase monotonically with passing time, while default reasoning is *logically* nonmonotonic, in that the set of conclusions from a set of premises need not contain the conclusion from subsets of premises (see appendix A).

These definitions ensure that some contraction or revision always exist. It is not clear that this assumption is necessary, or that the definitions here are the best possible. How to deal with the case in which no alternatives are maximally preferred is an interesting question for future study. It would appear that the notion of *rational representation* [10, 11], in which consistent subsets of inconsistent preferences are chosen to temporarily represent the inconsistent sets, should play a role here.

The generality exhibited in the definitions of rational contraction and revision goes against the usual presupposition of epistemologists that knowing more is always better. Indeed, one can even prove that knowing more is better in some standard theories (see [18]). But this presupposition (or theorem) is not always justified when beliefs incur costs to the agent. For example, the set $A \downarrow x$ need not be computable, or if computable, may take too long to compute. In such cases, practical belief revision cannot consider all alternatives and contraction may well simply seek to remove the easily identified inconsistencies rather than work to remove all inconsistencies. Similarly, knowing more may make the agent worse off if beliefs produce non-epistemological effects. For example, *1984*'s Winston Smith is much better off not knowing some of the facts about his crazy world, since he will be severely tortured if he reveals this knowledge, and has no hope of keeping up a perfect pretense of ignorance. More mundanely, there are many facts that people find very painful to know (infidelity of a spouse, criminality of an employer) unless they also know enough extra facts to permit effective action (adequate legal grounds for divorce, sufficient evidence for conviction). The general pattern here is that believing something may be worse than not believing it, even though not believing it is worse than believing it for good reason. (See [5, 11] for more examples.)

4.6 Economic vs. Gärdenfors rationality

Are economically rational contractions rational in Gärdenfors's sense? The obvious answer is no: rational contraction is of a different type than Gärdenfors's contraction, a nondeterministic relation rather than a functional relation, and this nondeterminism can cause big differences. For example, it is clear that an analogue of the equivalence axiom ($\dot{-}5$) holds for rational contraction, in that $A \hat{-} x = A \hat{-} y$ if $\vdash x \leftrightarrow y$. But if we look to specific revisions performed, ($\dot{-}5$) need not hold since we may pick a contraction when revising A by x that is different from one we may pick when revising by y (or even when revising by x at some later time).

Nevertheless, we can compare rational contractions with Gärdenfors contractions in a more interesting way by seeing if deterministic contraction operations that are rational in the economic sense need also satisfy Gärdenfors's axioms. Formally, we say that $\dot{-}$ is a *functional* rational contraction if $A \dot{-} x \in A \hat{-} x$ for every A and x. We then have the following result.

Theorem 1 *Functional rational contraction satisfies the Gärdenfors axioms* ($\dot{-}1$), ($\dot{-}2$), *and* ($\dot{-}4$), *and can violate* ($\dot{-}3$), ($\dot{-}6$)-($\dot{-}8$), *and the Levi and Harper identities* (4) *and* (5).

Proof: Suppose that $\dot{-}$ is a functional rational contraction operation and that x, y, and z are logically independent and collectively consistent propositions.

To begin with, $A \dot{-} x$ is clearly a belief set, so ($\dot{-}1$) holds. Similarly, $(A \dot{-} x) \subseteq A$, so ($\dot{-}2$) holds, as does ($\dot{-}4$).

Now suppose that $A = Cn(\{y\})$ and $A \not\vdash x$, so $A \Downarrow^* x = \{Cn(\emptyset), Cn(\{y\})\}$. If $Cn(\{y\}) \prec Cn(\emptyset)$, then $A \dot{-} x = Cn(\emptyset) \neq A$, so ($\dot{-}3$) need not be satisfied. Similarly, if $A = Cn(\{x, y\})$, then $A \Downarrow^* x = \{Cn(\emptyset), Cn(\{y\})\}$. If

$$Cn(\{y\}) \prec Cn(\{x, y\}) \prec Cn(\emptyset) \prec Cn(\{x\}),$$

then $A \dot{-} x = Cn(\emptyset)$ and $(X \dot{-} x) + x = Cn(\{x\})$ and $A \not\subseteq Cn(\{x\})$, so ($\dot{-}6$) need not hold.

Next, neither ($\dot{-}7$) nor ($\dot{-}8$) always hold. Suppose that $A = Cn(\{x, y, z\})$, so that

$$\begin{aligned} A \Downarrow^* x &= \{Cn(\emptyset), Cn(\{y\}), Cn(\{z\}), Cn(\{y, z\})\} \\ A \Downarrow^* y &= \{Cn(\emptyset), Cn(\{x\}), Cn(\{z\}), Cn(\{x, z\})\} \\ A \Downarrow^* (x \wedge y) &= \{Cn(\emptyset), Cn(\{z\})\}. \end{aligned}$$

Suppose further that we have

$$\begin{aligned} &Cn(\{x\}) \prec Cn(\{y\}) \prec Cn(\{z\}) \\ &\qquad \prec Cn(\emptyset) \prec Cn(\{x, z\}) \prec Cn(\{y, z\}) \end{aligned}$$

Then $Cn(\{y, z\})$ is maximal in $A \Downarrow^* x$, $Cn(\{x, z\})$ is maximal in $A \Downarrow^* y$, $Cn(\emptyset)$ is maximal in $A \Downarrow^* (x \wedge y)$, but

$$Cn(\{y, z\}) \cap Cn(\{x, z\}) = Cn(\{z\}) \neq Cn(\emptyset)$$

which violates ($\dot{-}7$). On the other hand, if

$$\begin{aligned} &Cn(\emptyset) \prec Cn(\{x, z\}) \prec Cn(\{y, z\}) \\ &\qquad \prec Cn(\{z\}) \prec Cn(\{x\}) \prec Cn(\{y\}), \end{aligned}$$

then $Cn(\{y\})$ is maximal in $A \Downarrow^* x$, $Cn(\{z\})$ is maximal in $A \Downarrow^* (x \wedge y)$, but $Cn(\{z\}) \not\subseteq Cn(\{y\})$, which violates ($\dot{-}8$).

Finally, neither the Levi nor Harper identities hold. Suppose $A = Cn(\{\neg x, y\})$, so that $A \Downarrow^* \neg x = \{Cn(\emptyset), Cn(\{y\})\}$. If we have

$$Cn(\{y\}) \prec Cn(\emptyset) \prec Cn(\{x\}) \prec Cn(\{x, y\}),$$

then $A \hat{-} \neg x = Cn(\emptyset)$, and

$$(A \hat{-} \neg x) + x = \{Cn(\{x\})\} \neq \{Cn(\{x, y\})\} = A \hat{+} x,$$

which contradicts the Levi identity (4). Similarly, we have

$$(A \hat{+} x) \cap A = Cn(\{y\}) \neq Cn(\emptyset) = A \hat{-} \neg x,$$

which contradicts (with x and $\neg x$ interchanged) the Harper identity (5). $\square$

These complications may be avoided if we require that contractions be chosen from among $A \downarrow x$ rather than $A \Downarrow^* x$, as shown by the following result.

Theorem 2 *If* $(A \mathbin{\hat{-}} x) \subseteq (A \downarrow x)$ *whenever* $\nvdash x$, *then rational functional contraction satisfies the Levi and Harper identities and the Gärdenfors axioms* $(\dot{-}1)$-$(\dot{-}8)$.

Proof: Suppose $(A \mathbin{\hat{-}} x) \subseteq (A \downarrow x)$ whenever $\nvdash x$. Then there is some maxichoice function $\mathcal{C}$ such that $\mathcal{C}(A \downarrow x) \in (A \mathbin{\hat{-}} x)$ for all x such that $\vdash\!\!\!/ x$. Since each particular contraction given by $\hat{-}$ can be viewed as the result of a maxichoice contraction, rational contraction in this case satisfies all the axioms that maxichoice contraction satisfies, namely (4), (5), and $(\dot{-}1)$-$(\dot{-}6)$. In fact, rational contraction satisfies $(\dot{-}7)$ and $(\dot{-}8)$ as well because any rational functional contraction operation is *orderly*, that is, any element of $A \mathbin{\hat{-}} x$ is an element of $A \downarrow x$ that is maximal with respect to a partial ordering of subsets of A. Gärdenfors [14, Lemma 4.3] shows that any orderly maxichoice contraction operation satisfies all the contraction axioms, and to apply this result, we need only consider the partial order $\sqsubseteq$ defined so that $X \sqsubseteq Y$ iff either $X = Y$ or $X \prec Y$. □

Another way of avoiding the complications of rational contraction is to require that believing more is better than believing less. Formally, we say that the preference relation $\precsim$ is *(positively) informationally monotone* if $\mathcal{L} \precsim X$ for all X and $X \precsim Y$ whenever X and Y are both consistent and $Y \vdash X$. We then have the following result.

Theorem 3 *If* $\precsim$ *is informationally monotone, then rational functional contraction satisfies the Levi and Harper identities and the Gärdenfors axioms* $(\dot{-}1)$-$(\dot{-}8)$.

Proof: Obviously, if $\precsim$ is informationally monotone, then $(A \mathbin{\hat{-}} x) \subseteq (A \downarrow x)$ if $\nvdash x$, so theorem 2 applies. □

Many questions remain to be answered. Are the Gärdenfors axioms equivalent to specific conditions on preferences over belief states? If there is an axiomatization of economically rational belief revision, is it equivalent to the existence of an expected utility preordering of states? That is, is there an analogue of Gärdenfors's theorem relating belief revision with epistemic entrenchment?

5 Partial epistemic preferences

Nebel suggests two reasons why epistemic relevance is a more practical basis for belief revision than epistemic entrenchment. In the first place, epistemic relevance orderings need not respect logical dependencies among propositions. That is, we are free to order $x \leq_\rho y$ without regard to whether $x \vdash y$ or $y \vdash x$, in contrast with the dominance condition (≤ 2) on epistemic entrenchment orderings, which can be quite costly (or impossible) to ensure. In the second place, linear epistemic relevance orderings of the propositions in a belief base make belief base contraction easy to implement (as in RUP [31]) by simply dropping the lowest ranked propositions in any conflicting set. Moreover, this form of belief base contraction is fully rational, and corresponds to maxichoice contraction [32, Lemma 17ff]. Epistemic entrenchment is not entirely impractical, as Gärdenfors and Makinson [16] show that epistemic entrenchment orders can be efficiently represented by information linear in the number of atomic propositions. But pure representation is only half the problem, and they leave the problem of logical dependencies unaddressed. Orderings of propositions and belief bases will only be useful in practice if they can be both represented and computed quickly.

While Nebel's epistemic relevance orderings make rational belief revision more practical, this approach is not without significant costs. The main problem is the inflexibility of this means for effecting contractions. Specifically, linear epistemic relevance orderings rank each possible contraction by the most valuable proposition retained, irrespective of what other propositions are retained. This might seem reasonable since $\sqsubseteq_\rho$ is applied to the elements of $B \downarrow x$, in which case choosing the subset with the maximal element is the same as choosing the subset which abandons the least valuable propositions. But suppose $X = \{a_1, \ldots, a_{1000}\}$ and $Y = \{a_{1001}\}$, where $\leq_\rho$ ranks these propositions by the natural indicial order, so that $X \sqsubseteq_\rho Y$. If all of these propositions are of roughly the same value (but each differing slightly from the rest), then the revision

$$\{a_1, \ldots, a_{1001}\} \ominus (a_{1001} \wedge (a_1 \vee \ldots \vee a_{1000})) = Y \quad (26)$$

chooses among the two alternatives X and Y and discards a thousand good propositions in favor of a single proposition that is little better than those discarded. This seems unreasonable compared with, for example, using a weighted comparison or voting scheme in which equally valuable propositions get equal say in the selection. Unfortunately, there is no way to express schemes like majority voting with linear orders over propositions. More generally, achieving any dependence of ordering on the global composition of the alternatives means revising the linear propositional order to fit each set of alternatives.

It would be valuable to have some more flexible way of specifying preferences for guiding contraction and revision. If we look to the usual explanations of why one revision is selected over another (such as those in [36]), we see that many different properties of propositions influence whether one proposition is preferred to another. For example, one belief might be preferred to another because it is more specific, or was adopted more recently, or has longer standing (was adopted less recently), or has higher probability of being true, or comes from a source of higher authority. The key observation, however, is that these criteria for comparing beliefs are often partial, that is, each may be viewed as a preorder (reflexive and transitive relation) $\precsim$ such that both $X \not\precsim Y$ and $Y \not\precsim X$ for some X and Y. For example, there are many different dimensions of specificity, and two beliefs may such that neither is more specific than the other. Similarly, probabilities need not be known for all propositions, and authorities need not address all questions. If we want contraction and revision to be truly flexible, we need some way of combining different partial orderings of propositions into complete global orderings of belief states.

But combining partial orderings into a global ordering can be difficult because the partial criteria may conflict in some cases. To borrow an example from nonmonotonic logic, we might reasonably prefer to believe that Quakers are pacifist, and that Republicans are not pacifists. These preferences can conflict on cases like that of Nixon, and a preferences for more specific rules does not help since "Quaker" and "Republican" are incomparable categories. Indeed, as argued in [12], other preference criteria can conflict as well, including very specific criteria corresponding to individual default rules. Constructing a global ordering thus means resolving the conflicts among the preference criteria being combined.

In addition to flexibility, we seek a revision method which is potentially mechanizable. This means that whatever method is employed for resolving conflicts must also be mechanizable because placing responsibility for resolving potential conflicts on the theorist is infeasible. For large sets of criteria it is difficult to anticipate all of the potential conflicts and all of the varying circumstances that may influence how the conflicts should be resolved. It also seems difficult to anticipate discovery of new criteria. Thus we seek conflict resolution mechanisms based on general, modular rules of combination that apply even as the criteria used evolve.

5.1 Constructing global preference orders

To analyze the problem of modular construction of orderings, we follow the formal approach elaborated in [12] for analyzing the related problem for preference-based nonmonotonic logics. We say that an *aggregation policy* is a function that specifies the global order corresponding to any given set of partial preference orders. Let the set I index the set of partial preference orders that are to be combined, so that if $i \in I$, $\precsim_i$ denotes the preference order corresponding to the ith pattern of inference to be included in the unified logic. The problem is then to aggregate the set of orders $\{\precsim_i | \ i \in I\}$ into a global preference order $\precsim$.

The principled design of an aggregation policy for partial preference criteria begins with a consideration of properties we think a reasonable policy should exhibit. The properties we propose are analogs of Arrow's [3] desiderata for social choice.[3]

[3]Consult sources on social choice theory [3, 38] for somewhat more rigorous versions of these desiderata, though for the case of total orders.

1. *Collective rationality.* The global order $\precsim$ is a function of the individual orders $\precsim_i$, which are unrestricted, possibly partial, preorders.

   Collective rationality simply stipulates that aggregation policies define general methods for combining multiple preference criteria.

2. *Pareto principle (unanimity).* If $X \prec_i Y$ for some $i \in I$ and for no $j \in I$ does $Y \prec_j X$, then $X \prec Y$. In other words, the global order agrees with unanimous strict preferences.

   The Pareto principle is a clearly desirable property of aggregation functions; a violation of unanimous preference would be difficult to justify.

3. *Independence of irrelevant alternatives (IIA).* The relation of X and Y according to the global order depends only on how the individual orders rank those two candidates. That is, considering new alternatives does not alter rankings among the originals.

   IIA simply means that if A is maximal among a set of alternatives, it is maximal in any subset including A.

4. *Nondictatorship (partial criteria).* There is no $i \in I$ such that for every X and Y, $X \precsim Y$ whenever $X \precsim_i Y$, regardless of the $\precsim_j$ for $j \neq i$. That is, there is no "dictator" whose preferences automatically determine the group's, independent of the other individual orderings.

   The nondictatorship condition simply states the problem faced by theorists of belief revision at this time: namely, that all known (and foreseeable) preference criteria to be aggregated are in fact partial, prone to override in the face of enough opposition by other criteria. This condition merely rules out the trivial solution to the aggregation problem; it says we cannot simply assume we possess some universal criterion that we in actuality lack.

5. *Conflict resolution.* If $X \precsim_i Y$ for some i, then $X \precsim Y$ or $Y \precsim X$. That is, if two candidates are comparable in an individual order, then they are comparable in the global order.

   The conflict resolution condition rules out the easy form of skepticism by mandating that the global order commit one way or the other whenever the individual orders express a preference. It requires that two alternatives be given the same rank if there is no reason to prefer one to the other.

Leaving aside the conflict resolution condition for now, the following theorem states that the desirable and apparently reasonable properties enumerated above are not simultaneously satisfiable by any aggregation policy for preferences expressed by total orders.

Theorem 4 (Arrow) *If the domain includes more than two alternatives, no aggregation policy mapping* *sets of total preorders to total global preorders satisfies the collective rationality, Pareto, IIA, and nondictatorship conditions.*

Proof: With the restriction to total orders, this is exactly Arrow's theorem. For a proof of the original result see [4] or [38, Chapter 7]. □

There is no problem finding good aggregation policies for choices among only two alternatives: majority rule works fine, for example. But for the case of belief revision, there are typically several possible alternatives to choose from. This means that the following theorem applies to the typical case of belief revision.

Theorem 5 *No aggregation policy for partial preference preorders satisfies the collective rationality, Pareto, IIA, nondictatorship, and conflict resolution conditions.*

Proof: The only difference between the belief revision problem and the classic social choice setup is that the individual and global orders can be partial whereas individual and social rankings are taken to be total. Partiality is constrained, however, by the conflict resolution condition's restriction that the global order be at least as complete as the constituent orders. Therefore, any set of belief states that is totally ordered by some $\precsim_i$ is also totally ordered by $\precsim$. The impossibility of the special case of aggregation functions mapping sets of total orders to a total order as in Theorem 4, together with the IIA condition, implies impossibility of the generalized problem where the orders may be incomplete.□

Thus if rational belief revision requires a preorder that completely orders the alternative contractions or revisions, we may expect that the only general way of obtaining the order is manual construction, that is, to supply a dictatorial policy. One way to do this is to impose a linear ordering over all the criteria, so that the first criterion always gets its way regardless of what the rest of the criteria say, unless it expresses no preference, in which case the second criterion gets its way, and so on. Choice rules of this form are called *lexicographic* because they resemble the method for ordering words alphabetically: compare the first letters; if tied compare the second, etc. But like Nebel's orderings, these linear orderings are not very flexible, and can be expected to require ongoing manual revision to achieve satisfactory performance. (See [12] for more discussion of ways around theorem 5 in the context of nonmonotonic logic.)

5.2 Revision rationality and aggregation rationality

Some irrationalities in preference aggregation do not affect the formal rationality of belief revision. For example, aggregation policies which violate the Pareto or the nondictatorship principles do not satisfy one or more of the partial preference criteria but may still yield global preference orderings. In such cases, the resulting belief revisions will be rational with respect to the aggregate order but may not be rational with respect to the ignored preference criteria. On the other hand, violations of collective rationality, IIA, or conflict resolution have immediate effects on the nature of rational belief revision. We examine these in turn.

An aggregation policy may violate collective rationality in several ways. First, it may not aggregate every possible partial preference criterion. This need not affect the formal rationality of belief revision. As with violations of the Pareto or nondictatorship conditions, it just means that some preferences will be ignored. Second, the result of aggregation may not be complete. Incompleteness may mean that some preferences are ignored, but in some cases it may mean that there are no rational choices among $A \Downarrow^* x$, so that the resulting contraction is vacuous. For example, suppose that $A \Downarrow^* x = \{X, Y\}$. If the aggregate order does not relate these two alternatives, then neither X nor Y is maximally preferred, so neither is a rational choice, and $A \dot{-} x = A$. Intransitivity of $\precsim$ means that for some X, Y, Z we have $X \precsim Y$ and $Y \precsim Z$ but $X \not\precsim Z$, which again is incompleteness of the aggregate order. Another way of looking at incompleteness and intransitivity is through the equivalence of $X \prec Y$ and $Y \not\precsim X$. This equivalence really only holds in the case of complete orders, but if we apply it to incomplete orders, incompleteness appears as inconsistency of strict preference. For example, $X \not\precsim Y$ and $Y \not\precsim X$ mean $Y \prec X$ and $X \prec Y$ under this translation, and intransitivity means that $X \precsim Y$ and $Y \precsim Z$ but $Z \prec X$.

An aggregation policy that violates IIA can exhibit erratic behavior, specifically, it can lead to different revisions depending on the presence or absence of irrelevant information. For example suppose $A = \{x, y, z\}$. We then have

$$\begin{aligned} A \Downarrow^* x \wedge y &= \{Cn(\emptyset), Cn(\{x\}), Cn(\{y\}), \\ &\qquad Cn(\{x,z\}), Cn(\{y,z\})\} \\ A \Downarrow^* (x \wedge y) \vee z &= \{Cn(\emptyset), Cn(\{x\}), Cn(\{y\})\} \end{aligned}$$

Suppose further that P_1 and P_2 are two preference criteria to be combined:

$$\begin{aligned} P_1 : \quad & Cn(\emptyset) \prec Cn(\{x\}) \sim Cn(\{x,z\}) \\ & \qquad \prec Cn(\{y\}) \sim Cn(\{y,z\}) \\ P_2 : \quad & Cn(\emptyset) \prec Cn(\{y\}) \sim Cn(\{y,z\}) \\ & \qquad \prec Cn(\{x\}) \sim Cn(\{x,z\}) \end{aligned}$$

If the aggregation policy violates IIA, we may have $A \dot{-} (x \wedge y) = \{Cn(\{y\})\}$ but $A \dot{-} (x \wedge y) \vee z = \{Cn(\{x\})\}$, with the overall choice between $Cn(\{x\})$ and $Cn(\{y\})$ depending on what other alternatives are available.

An aggregation policy that violates conflict resolution incompletely orders the alternatives for which preferences exist. This has the same consequence of vacuous contractions as violation of collective rationality that result in incomplete aggregate orders. Incompleteness when no relevant preferences are available is not necessarily bad. It means that the choice may not be rational because there is no information, and we may be justified in assuming that any choice is equally reasonable. But incompleteness when preferences do exist is a different matter, since uniform incompleteness in the face of conflicts is just the skeptical approach criticized earlier.

6 Revising revision preferences

Gärdenfors points out that the epistemic entrenchment ordering used in contraction and revision may depend on the state of belief. For example, if one holds strongly to some belief because an impeccable authority vouched for it, and then learns that the authority has lied about some matter, the original belief may become less entrenched. Nebel acknowledges a similar variability of epistemic relevance. For example, if one finishes one task and moves on to work on another, the beliefs about the particulars of the first task are no longer epistemically relevant and may be discarded. But as this example suggests, it is more accurate to view these relations as depending on the overall mental state of the agent, rather than on just the belief state.

We will view epistemic preferences as functions of the mental state of the agent. Let Σ stand for the set of mental states. We will assume that each $S \in \Sigma$ can be interpreted as containing a set A of beliefs and a set P of preferences.[4] For simplicity, we will assume that states consist of sets of various mental attitudes, so that $A, P \subseteq S$.

Steps of reasoning, such as expansions, contractions, and revisions, take one mental state into another, and so any complete theory of belief revision must address the question of how preferences change in reasoning. For Gärdenfors's theory, this means formulating some more precise conception of the nature or origins of epistemic entrenchment; for Nebel's, epistemic relevance.

6.1 Pure changes in preferences

The axioms for preferences provide the basis for the logic of preferences, by stating that preferences should be transitively closed and consistent. This logic can be viewed as a standard logic, like that of the logic of beliefs. For example, we may view the set of preferences as a set of "propositions" about the preference relation, and introduce a logical relation $\vdash_P$ to encompass the transitivity and completeness axioms, and so produce a closure operator Cn_P which takes a partial set of preferences and takes its transitive closure.

In this setting, revision of preferences may be viewed as entirely analogous to revision of beliefs, guided by principles like the Gärdenfors axioms, but subject to $\vdash_P$ instead of $\vdash$. But a new issue arises if we seek to make preference revision rational by using preferences about preferences. Rational choice means selection guided by preferences, but if preferences are changing, or if there are preferences about preferences, we must be clearer about which preferences are used to make rationality judgments. In particular, the optimal revision according to the preferences held prior to a revision may not be optimal according to the preferences obtaining after the revision. Paraphrasing Jeffrey [24], we may say that a *ratified* revision or contraction is one that is rational according to the preferences that result from the revision or contraction.

6.2 Belief-dependent preferences

This picture leaves preferences completely independent of beliefs. In general, however, we may expect that preferences may be explicitly conditional on assumptions as well as on other preferences. For example, preferences may themselves be assumptions during periods of experimentation in which the agent seeks to find out what it likes, and these tentatively held preferences may be revised upon gaining new information about the effects of different choices. More generally, beliefs and preferences may be conditional on other beliefs and preferences. If $\vdash$ and Cn represent the combined logic of beliefs and preferences, then the basic closure condition on states is simply $A \cup P = Cn(A \cup P)$.

However, we must extend this theory significantly if we wish to describe practical approaches to belief revision, which are based on states consisting of finite representations of beliefs and preferences, such as Nebel's belief bases. In Nebel's theory, belief states are just the deductive closures of belief bases. This follows a long tradition of deductive representation. But as argued elsewhere [11], the representation relation between belief bases and belief states should be viewed instead as rational selection of sets of consequential beliefs and preferences given an initial set of beliefs and preferences. That is, the meaning of finite belief representations will depend on preferences in practical approaches to representation. In this setting, the "logic" of mental states may be nonmonotonic, and even nondeterministic.

Permitting the logic of states to be nonmonotonic has dramatic consequences for the principles of belief revision. Since removing x may enable some default inference previously blocked by the belief x, a rational "contraction" might actually add beliefs as well as remove others. There are many ways of formalizing this, and one possible approach is as follows. Let Cn' be a new closure operator, and define two new operations $A \dagger x$ and $A \dagger^* x$ by

$$A \dagger x \stackrel{\text{def}}{=} \{X \subseteq A \mid x \notin Cn'(X) \land X \subset Y \subseteq A \rightarrow x \in Cn'(Y)\} \tag{27}$$

[4] Thomason [43] raises doubts about whether beliefs and preferences are really determined by mental states, but we will not address this issue here.

and

$$A \dagger^* x \stackrel{\text{def}}{=} \{Cn'(X) \mid X \in (A \dagger x)\} \qquad (28)$$

$A \dagger x$ gives the maximal subsets of A whose Cn'-closures do not contain x, while $A \dagger^* x$ gives the closures of these maximal subsets. Suppose that $\mathcal{S}(A \dagger^* x)$ picks out the states in $A \dagger^* x$ that are maximal with respect to a complete preorder $\sqsubseteq$. We define *nonstandard* (or nonmonotonic) partial meet contraction, written $\stackrel{\text{n}}{-}$, by

$$A \stackrel{\text{n}}{-} x \stackrel{\text{def}}{=} \begin{cases} \bigcap \mathcal{S}(A \dagger^* x) & \text{if } \nvdash x \\ A & \text{otherwise.} \end{cases} \qquad (29)$$

We then have the following result concerning nonstandard partial meet contraction.

Theorem 6 *If* Cn' *is not monotonic and* $\mathrm{Cn}(X) \subseteq \mathrm{Cn}'(X)$*, then* $\stackrel{\text{n}}{-}$ *satisfies* ($\dot{-}$3)-($\dot{-}$5) *and fails to satisfy* ($\dot{-}$1) *and* ($\dot{-}$2)*.*

Proof: Suppose that $Cn(X) \subseteq Cn'(X)$ and that Cn' is not monotonic. We then have:

($\dot{-}$1): Contraction need not be a belief state because the intersection of Cn'-closed sets need not itself be Cn'-closed.

($\dot{-}$2): Removing one belief may make believing others possible if the removed belief had been defeating some assumption.

($\dot{-}$3): Clearly, $A \dagger x = \{A\}$ if $x \notin A$, so $A \stackrel{\text{n}}{-} x = A$.

($\dot{-}$4): Suppose $\nvdash x$. If $A \dagger x = \emptyset$, then $x \notin (A \stackrel{\text{n}}{-} x)$. Otherwise, $x \notin X$ for each $X \in (A \dagger^* x)$, and so x is not in the intersection.

($\dot{-}$5): Since closure under Cn' implies deductive closure, we have $A \dagger^* x = A \dagger^* y$ whenever $\vdash x \leftrightarrow y$, hence the contractions are identical. $\Box$

If Cn' is not required to subsume deductive closure, partial meet contraction may fail to satisfy ($\dot{-}$5), since detecting the equivalence of two beliefs may not be feasible.

The close connection between belief revision and assumption-making means that many of the preferences about what assumptions to make are also preferences about what revisions to make. For example, it seems reasonable that if x is to be believed in preference to $\neg x$ when there is no other reason to believe either, a revision containing x should be preferred to one containing $\neg x$ if both are possible. But there may be cases in which these two types of preferences disagree, since the operations of assuming some belief for the purpose of action and adopting the belief outright have different consequences. This can happen when repeated choices may be made revocably until they must be made irrevocably. For example, in the past most young men believed in monogamy, but many would nevertheless choose to believe they could "sow their wild oats" with women other than their preferred bride up until the time when marriage seemed desirable, and would never marry their lovers except in shotgun marriages.

7 Conclusion

We examined Gärdenfors's theory of rational belief revision and its extension by Nebel. The heart of Gärdenfors's theory is the ordering of beliefs according to epistemic entrenchment and its equivalence with his notion of rational belief revision. Nebel adapted this theory to cover revision of finite bases of belief, and his results show that finite revision can be rational in Gärdenfors's sense, and that in the finite case maxichoice revision is as rational as skeptical revision.

We argued that these theories are inadequate because they demand more information about which revisions are preferable than is typically available in practice, and because their notion of "rationality" has little to do with the economic notion of rational choice among alternatives. We presented a modification of these theories which makes weaker, more realistic assumptions, namely that belief revisions must be guided by partial preferences which may conflict with each other. This theory is closely related to our theory of rational default inference, and supports similar results. We showed how the notion of rationality proposed by Gärdenfors and Nebel is violated by revisions which are rational in the economic sense, and proved that fully rational belief revision is sometimes impossible when rationality is judged by partial preferences.

It is an open problem to find rationality postulates to describe revision that is rational in the economic sense. There may not be any. There certainly will not be axioms that refer only to beliefs, as in Gärdenfors's axioms. There may not be a single ideal theory of rational revision because one may not exist, as suggested by theorem 5. Is there a complete axiomatization if one presupposes a global expected utility preordering of beliefs, or of belief states?

Since completely rational belief revision appears to be impossible, we must settle for revisions of imperfect rationality. Each feasible form of revision will be irrational in one way or another, and their strengths and weaknesses in different domains or applications may be compared.

Acknowledgments

This work draws substantially on previous work conducted in collaboration with Michael Wellman, and I thank him for many valuable suggestions and for reading drafts of this paper. I also thank Peter Gärdenfors, Sten Lindström, Bernhard Nebel, Wlodzimierz Rabinowicz, and Joseph Schatz for valuable discussions. This work was supported by the National Library of Medicine through National Institutes of Health Grant No. R01 LM04493 and by the International Business Machines Corporation through a grant to MIT's Laboratory for Computer Science.

A Temporal and logical nonmonotonicity

The adjective "nonmonotonic" has suffered much careless usage recently in artificial intelligence, and the only thing common to many of its uses is the term "nonmonotonic" itself. In fact, two principal ideas stand out among these uses: namely, that attitudes are gained and lost over time, that reasoning is nonmonotonic—this we call *temporal* nonmonotonicity—and that unsound assumptions can be the deliberate product of sound reasoning, incomplete information, and a "will to believe"—which we call *logical* nonmonotonicity. Indeed, much of the confusion reigning about the subject stems from a confusion between these two sorts of nonmonotonicity, and between logical nonmonotonicity and nonmonotonic logic.

Let us differentiate these uses in precise formal terms. In mathematics, the terms monotonic and nonmonotonic (or monotone and nonmonotone) refer to properties of functions between ordered sets, so to use these terms with precision in describing a reasoning agent, we must identify specific functions with ordered domains and codomains to which we may attribute these properties. When we view states as closed sets of beliefs, we can distinguish two functions between ordered sets. Let $\mathcal{T}$ be the set of temporal instants of a history ordered by increasing time. Consider the state space $\Sigma \subseteq 2^{\mathcal{L}}$ ordered by set inclusion, so that states with additional elements are considered bigger. We may then view each history of the agent as a function

$$S : \mathcal{T} \rightarrow 2^{\mathcal{L}},$$

and the closure operator as a function

$$Cn : 2^{\mathcal{L}} \rightarrow 2^{\mathcal{L}}$$

such that $S(t) = S_t = Cn(S_t)$ for each instant t.

Now S and Cn are functions between ordered sets, and so may be monotonic or not. We say that monotonicity of S with increasing time is *temporal* monotonicity of the agent's attitudes; that is, the agent's belief states exhibit temporal monotonicity if they are *cumulative*, if $S_t \subseteq S_{t'}$ whenever $t \leq t'$. *Logical* monotonicity is the usual property of deductive closure functions Cn; the set of conclusions grows monotonically with increasing sets of axioms, that is, $Cn(X) \subseteq Cn(X')$ whenever $X \subseteq X'$. Thus temporal or logical nonmonotonicity occurs when the agent's characterization employs nonmonotonic functions for S or instead of Cn.

The idea that reasoning may be nonmonotonic is very old, for in almost all familiar situations the attitudes of agents change nonmonotonically over time; that is, the function S is temporally nonmonotonic in ordinary situations. Confusion between logical and temporal nonmonotonicity arises when we view histories as given by transition functions $\tau : \Sigma \rightarrow \Sigma$ such that $\tau(S_t) = S_{t+1}$, for transition functions are usually logically nonmonotonic (producing, for example, rational maxima which may change if new preferences are added), in addition to being used to characterize temporally nonmonotonic histories.

References

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

[2] C. Alchourrón and D. Makinson. The logic of theory change: Contraction functions and their associated revision functions. *Theoria*, 48:14–37, 1982.

[3] K. J. Arrow. *Social Choice and Individual Values.* Yale University Press, second edition, 1963.

[4] K. J. Arrow. Values and collective decision-making. In P. Laslett and W. G. Runciman, editors, *Philosophy, Politics and Society*, Third, pages 215–232. Basil Blackwell Oxford, 1967.

[5] J. A. K. Cave. Learning to agree. *Economics Letters*, 12:147–152, 1983.

[6] M. Dalal. Investigations into a theory of knowledge base revision: Preliminary report. In *Proceedings of the Seventh National Conference on Artificial Intelligence*, volume 2, pages 475–479, San Mateo, CA, August 1988. AAAI, Morgan Kaufmann.

[7] J. de Kleer. An assumption-based TMS. *Artificial Intelligence*, 28:127–162, 1986.

[8] J. Doyle. A truth maintenance system. *Artificial Intelligence*, 12(2):231–272, 1979.

[9] J. Doyle. Some theories of reasoned assumptions: An essay in rational psychology. Technical Report 83-125, Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA, 1983.

[10] J. Doyle. Artificial intelligence and rational self-government. Technical Report CS-88-124, Carnegie-Mellon University Computer Science Department, 1988.

[11] J. Doyle. Constructive belief and rational representation. *Computational Intelligence*, 5(1):1–11, Feb. 1989.

[12] J. Doyle and M. P. Wellman. Impediments to universal preference-based default theories. In R. J. Brachman, H. J. Levesque, and R. Reiter, editors, *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning (KR'89)*, pages 94–102, San Mateo, CA, May 1989. Morgan Kaufmann.

[13] R. Fagin, J. D. Ullman, and M. Y. Vardi. On the semantics of updates in databases. In *Proceed-*

ings of the Second ACM SIGACT-SIGMOD Conference, pages 352–365, 1983.

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

[15] P. Gärdenfors. The dynamics of belief systems: Foundations vs. coherence theories. *Revue Internationale de Philosophie*, 172:24–46, January 1990.

[16] P. Gärdenfors and D. Makinson. Revisions of knowledge systems using epistemic entrenchment. In M. Y. Vardi, editor, *Proceedings of the Second Conference on Theoretical Aspects of Reasoning About Knowledge*, pages 83–95, Los Altos, CA, March 1988. Morgan Kaufmann.

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

[18] I. J. Good. On the principle of total evidence. In *Good Thinking: The Foundations of Probability and Its Applications*. University of Minnesota Press, 1983. Originally appeared in *British Journal of Philosophy of Science*, 17:319-321, 1967.

[19] J. W. Goodwin. *A Theory and System for Non-Monotonic Reasoning*. PhD thesis, Department of Computer and Information Science, Linköping University, Linköping, Sweden, 1987. Linköping Studies in Science and Technology, No. 165.

[20] G. Harman. *Change in View: Principles of Reasoning*. MIT Press, Cambridge, MA, 1986.

[21] W. L. Harper. Rational belief change, Popper functions, and counterfactuals. In W. L. Harper and C. A. Hooker, editors, *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, volume 1, pages 73–115. Reidel, Dordrecht, 1976.

[22] W. L. Harper. Rational conceptual change. In *PSA 1976*, volume 2, pages 462–494, East Lansing, MI, 1976. Philosophy of Science Association.

[23] J. F. Horty, R. H. Thomason, and D. S. Touretzky. A skeptical theory of inheritance in nonmonotonic semantic networks. *Artificial Intelligence*, 42(2-3):311–348, March 1990.

[24] R. C. Jeffrey. *The Logic of Decision*. University of Chicago Press, Chicago, second edition, 1983.

[25] K. Konolige. Belief and incompleteness. In J. R. Hobbs and R. C. Moore, editors, *Formal Theories of the Common-Sense World*, pages 359–403. Ablex, Norwood, 1985.

[26] H. J. Levesque. A logic of implicit and explicit belief. In *Proceedings of the National Conference on Artificial Intelligence*, pages 198–202. American Association for Artificial Intelligence, 1984.

[27] I. Levi. Subjunctives, dispositions and chances. *Synthese*, 34:423–455, 1977.

[28] D. Lewis. *Counterfactuals*. Blackwell, Oxford, 1973.

[29] J. a. P. Martins and S. C. Shapiro. A model for belief revision. *Artificial Intelligence*, 35(1):25–79, May 1988.

[30] D. A. McAllester. A three valued truth maintenance system. AIM 473, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA, 02139, 1978.

[31] D. A. McAllester. Reasoning Utility Package user's manual. AIM 667, Massachusetts Institute of Technology, Artificial Intelligence Laboratory, 545 Technology Square, Cambridge, MA, 02139, 1982.

[32] B. Nebel. A knowledge level analysis of belief revision. In R. J. Brachman, H. J. Levesque, and R. Reiter, editors, *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning (KR'89)*, pages 301–311, San Mateo, CA, May 1989. Morgan Kaufmann.

[33] B. Nebel. *Representation and Reasoning in Hybrid Representation Systems*. Number 422 in Lecture Notes in Artificial Intelligence. Springer-Verlag, Berlin, 1990.

[34] J. L. Pollock. Defeasible reasoning. *Cognitive Science*, 11:481–518, 1987.

[35] W. V. Quine. *Philosophy of Logic*. Prentice-Hall, Englewood Cliffs, NJ, 1970.

[36] W. V. Quine and J. S. Ullian. *The Web of Belief*. Random House, New York, second edition, 1978.

[37] A. S. Rao and N. Y. Foo. Formal theories of belief revision. In R. J. Brachman, H. J. Levesque, and R. Reiter, editors, *Proceedings of the First International Conference on Principles of Knowledge Representation and Reasoning (KR'89)*, pages 369–380, San Mateo, CA, May 1989. Morgan Kaufmann.

[38] F. S. Roberts. *Discrete Mathematical Models*. Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1976.

[39] L. J. Savage. *The Foundations of Statistics*. Dover Publications, New York, second edition, 1972.

[40] Y. Shoham. *Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence*. MIT Press, Cambridge, MA, 1988.

[41] R. M. Stallman and G. J. Sussman. Forward reasoning and dependency-directed backtracking in a system for computer-aided circuit analysis. *Artificial Intelligence*, 9(2):135–196, 1977.

[42] R. C. Stalnaker. *Inquiry*. MIT Press, Cambridge, MA, 1984.

[43] R. H. Thomason. The context-sensitivity of belief and desire. In M. P. Georgeff and A. L. Lansky, editors, *Reasoning about Actions and Plans: Proceedings of the 1986 Workshop*, pages 341–360. Morgan Kaufmann, 1986.

[44] M. Winslett. Is belief revision harder than you thought? In *Proceedings of the Fifth National Conference on Artificial Intelligence*, volume 1, pages 421–427, August 1986.