A Complete Characterization of a Notion of Contraction Based on Information-Value

Horacio Arló-Costa
Carnegie Mellon University
Baker Hall 161B
Pittsburgh, PA 15213
hcosta@andrew.cmu.edu

Isaac Levi
Columbia University
1150 Amsterdam Avenue
New York, NY10027
levi@columbia.edu

Abstract

We present a decision-theoretically motivated notion of contraction which, we claim, encodes the principles of minimal change and entrenchment. Contraction is seen as an operation whose goal is to minimize losses of informational value. The operation is also compatible with the principle that in contracting A one should preserve the sentences better entrenched than A (when the belief set contains A). Even when the principle of minimal change and the latter motivation for entrenchment figure prominently among the basic intuitions in the works of, among others, Quine (Quine 1951), Levi (Levi 1980) (Levi 1991), Harman (Harman 1999), and Gardenfors (Gärdenfors 1988), recent formal accounts of belief change (like AGM - see (Gärdenfors 1988)) have abandoned both principles (see (Rott 2000)). We argue for the principles and we show how to construct a contraction operation which obeys both. An axiom system is proposed. We also prove that the decision-theoretic notion of contraction can be completely characterized in terms of the given axioms. Proving this type of completeness result is a well-known open problem in the field, whose solution requires employing both decision-theoretical techniques and logical methods recently used in belief change.

Principles for Belief Change

The following principle has guided research in belief revision at least since it was sketched in (Quine 1951) and (van O. Quine & Ullian 1978) (we use here a recent formulation of it presented by Hans Rott and Maurice Pagnucco in (Rott & Pagnucco 1999).)

The principle of Economy
Keep loss to a minimum in contraction.

Of course one needs to specify what is the index that is minimized in changing beliefs. A usual way of making this explicit is done as follows (see (Gärdenfors & Rott 1988), p. 38):

The principle of Informational Economy
Keep loss of information to a minimum in contraction.

Moreover the aforementioned principle is usually complemented by the following additional principle (see (Gärdenfors & Rott 1988), p 40):

The principle of Entrenchment
When a sentence $[A]$ is given up from a belief set, one should preserve the sentences better entrenched than A.

One of the points in (Rott 2000) is to explain in some detail that although many influential authors have maintained versions of these two principles, almost no concrete theory of belief change has endorsed them formally. Therefore in (Rott 2000) Hans Rott proposed the radical move of abandoning them and constructing a theory of belief change on the basis of alternative principles.

In this article we propose that there is nothing wrong with the principle of Entrenchment and that the Principle of Economy admits the following sound and intuitive instance:

The principle of Economy for Information Value
Keep loss of informational value to a minimum in contraction.

Of course, the explicit introduction of the value of information gives the resulting models a decision theoretical flavor. We argue that the project of constructing a notion of contraction based on the joint assumption of our version of the principle of Economy and the principle of Entrenchment is viable (from now on we will use the term 'Economy' to refer to our version of the principle). Moreover we offer a complete characterization of such a contraction. The notion of contraction that thus arises is quite different formally from some of standards in the field, like AGM (Alchourrón C. Gärdenfors & D. 1985), and others. Nevertheless our decision-theoretical model characterizes the same axioms as the ones proposed in (Rott & Pagnucco 1999) on the basis of rather different considerations. The notion of contraction we defend here is also one of the central notions of contraction studied by one of us (Levi) in (Levi 2004). This paper extends and complements this work by tackling various formal problems not considered in (Levi 2004). For example, we offer here a completeness result for the aforementioned notion of contraction. The result depends centrally on developing a representation of contraction operators in terms of *system of shells of informational value*, a notion not considered in (Levi 2004). Further use of this representation, as well as an extension of it useful for iteration, is offered by one of us (Arló-Costa) in a companion paper (Arló-Costa 2004).

Operators of informational value

We will assume a classical propositional language L as a representational tool. We assume that L contains the classical connectives. The underlying logic will be identified with its Tarskian consequence operator Cn: $2^L \rightarrow 2^L$ which is assumed to obey for all subsets X and Y of L:

(Inclusion) $X \subseteq Cn(X)$

(Monotony) If $X \subseteq Y$, then $Cn(X) \subseteq Cn(Y)$

(Iteration) $Cn(X) = Cn(Cn(X))$

(Superclassicality) If A can be derived from X by classical logic, then $A \in Cn(X)$

We also assume that Cn obeys the deduction theorem and that Cn is compact. A *theory* is any set K such that $K = Cn(K)$. Theories can be used advantageously in order to represent *commitments to full belief* of agents. So, our investigation here will focus on the logic of theory change.

We understand the problem of how to contract by removing A from a theory K to be a decision problem where one is called upon to choose a contraction removing A from K *from among all the contraction strategies removing A from K available in the context.*

Let K be a theory (representing the current commitments for full belief) and let LK be a *minimal theory* such that $LK \subseteq K$. The *basic partition* Π is a set of expansions of LK, not necessarily all of them and not necessarily all (or some of) the maximal and consistent ones. Even when the set of expansions of LK constituting Π need not contain the set of maximal and consistent extensions of LK, we will require that these expansions have to be sufficiently informative in order to permit contractions of every non-tautological sentence A in K (which does not happen to be entailed by all partition cells in Π). Syntactical constraints on admissibility of partitions reflecting this requirement will be provided below. The *ultimate partition* is the subset Π_K of partition cells of Π whose intersection is exactly K. In addition Π - Π_K is the *dual ultimate partition* Δ.

A *potential contraction* is obtainable by taking any subset of Δ, forming its intersection and then taking the intersection of K and the result. A *maxichoice contraction* of K relative to Δ is the intersection of K with a single element of Δ. A *maxichoice contraction of K removing A* relative to Δ is the intersection of K with a single element of Δ that entails $\neg A$. A *saturatable contraction* of K removing A relative to Δ is the intersection of a maxichoice contraction of K removing A relative to Δ with the intersection of a set of elements of Δ none of which entail $\neg A$.

A partition Π is *decisive* with respect to a sentence A, if and only if every cell of Π either entails A or its negation. And a partition Π is decisive with respect to a set of sentences if it is decisive with respect to all the sentences in the set. A partition cell is decisive with respect to a sentence as long as it entails the sentence or its negation. Let At be a set of atoms of L such that an atom t is in At if and only if there is at least a cell of Π which is decisive with respect to t. We will assume here as a constraint in the representation that a partition Π is admissible if and only if (1) it is decisive with respect to all the atoms in At[1] and (2) for every atom t in K there should be at least one partition cell in Δ entailing $\neg t$.

Any sentence $A \in K$ which is also entailed by all cells in Δ will be called a T-sentence. The idea is that it will be treated as a tautology in the sense that there are no potential contractions giving up A definable with respect to Π. It is clear that any admissible partition is such that for any sentence A that is not a T-sentence if $A \in K$ then there is a cell in Δ entailing $\neg A$. In other words, admissible partitions always permit to define contractions of non T-sentences of K.

Of course, for any admissible partition there is a largest language $\mathbf{L} \subseteq$ L such that Π is decisive with respect to all the atoms of $\mathbf{L}$, namely the language that has exactly the atoms in At. Consider now such language $\mathbf{L}$. Given Π consider $\Pi_{\mathbf{L}} = \{P \cap \mathbf{L}$: P is a partition cell of $\Pi\}$. Notice that $\Pi_{\mathbf{L}}$ is such that all its partition cells are maximal and consistent theories expressible in $\mathbf{L}$. Of course none of the partition cells in the original partition Π need to be a maximal and consistent theory in the original language L. Nevertheless one can work with $\Pi_{\mathbf{L}}$ without loss of generality and this is the strategy we will adopt here.

DEFINITION 0.1 *Let $S(K, A)$ be the family of A-saturatable sets of K. I.e. if K is a theory, $X \in S(K, A)$ if and only if $X \subseteq K$, X is closed, and $Cn(X \cup \{\neg A\})$ is an element of the partition Δ.*

Now, since we are working in all cases with the partition $\Delta_{\mathbf{L}}$ rather than Δ, the second clause in the previous definition is equivalent to say that $Cn(X \cup \{\neg A\})$ is a maximal and consistent theory in $\mathbf{L}$. Relativities to $\mathbf{L}$ will be tacitly assumed from now on.

Call $\mathcal{M}$ the set of all maximal and consistent sets (of $\mathbf{L}$) in the basic partition. This notation will be useful later. $\Phi = \{X : X = \cap Y$, with $Y \in 2^{\Delta} \cup [K]\}$. Let $[T]$ the set of maximals extending theory T. Then we can now introduce a *measure of informational value* $V : \Phi \rightarrow$ [0,1]. V is not just any value function. As the terminology indicates V is supposed to deliver a measure of the value of *information*. As such we assume that it inherits some basic properties of classical measures of information which are probability-based. A classical manner of utilizing probability in order to measure the content of information is to utilize the measure Cont(.) = 1 - Prob(.) - see for example (Levi 1980) for an account of how this measure can be used in order to construct a decision-theoretically motivated theory of *expansion*.

There are two basic properties that probability-based measures of information satisfy. First they *respect entailment* in the following sense:

(Weak Monotony) For any two sets X, that are elements of Φ, such that $X \subset Y$, $V(X) \leq V(Y)$.

[1] In other words, if a partition cell is decisive with respect to an atom, then all partition cells must be decisive with respect to this atom as well.

In order to introduce the second structural property respected by measures of information in general we need some additional notation. Let X and Y be two saturatable contractions in $S(K,A)$. Any saturatable contraction S in $S(K,A)$ has the canonical form $K \cap T_A \cap m_{\neg A}$, where T_A is an intersection of A-maximals and where $m_{\neg A}$ is a $\neg A$-maximal. Following notation we introduced above, let $[T]$ the set of maximals extending theory T. For any saturatable contraction S in $S(K,A)$ let then $S_{/K}$ be $T_A \cap m_{\neg A}$. So, if we have a saturatable contraction S that has been formed by intersecting K with a set of maximals $[S_{/K}]$ which is distinct from the set of maximals $[X_{/K}]$ and $[Y_{/K}]$ used in the construction of saturatable contractions X and Y (in the sense that the the intersection of $[S_{/K}]$ with $[X]$ and $[Y]$ is empty) then we call S an *alternative* saturatable contraction with respect to X and Y. With the help of these preliminaries we can now introduce our second postulate:

(Extended Weak Monotonicity) Let X, Y be saturatable contractions in a family $T \subseteq S(K,A)$. Then if S is an alternative saturatable contraction with respect to X and Y, and if $V(X) \leq V(Y)$, then $V(X \cap S) \leq V(Y \cap S)$.

The idea of this postulate is that intersecting saturatable contractions X and Y by the same alternative saturatable contraction Z cannot reverse the order of the informational values of X and Z. Most familiar measures of information which are probability-based obey this principle (including, of course, Cont). The trouble of measures like Cont is that the Cont-value of the intersection of two optimal saturatable contractions need not and, in general, will not carry maximum Cont-value. Nevertheless we think that it is reasonable to follow here the standard 'skeptical' approach (utilized in AGM and in other accounts of contraction and revision) according to which when several contractions removing A from K are optimal one should move to a position of suspense. But, notice that the two constraints that we have adopted so far permit the following scenario. Suppose that S is the set of optimal saturatable contractions removing A from K. Nothing so far prohibits to have an optimality set S where the intersection of S is not optimal. If this situation were possible then the aforementioned option of moving to a position of suspense would lack a justification in terms of optimality. And we think that it is constitutive of the notion of informational value we are presenting here that positions of suspense in these situations should be recommended as optimal. One way of guaranteeing the optimality of the intersection of any finite set of optimal contractions removing A from K, is to require that:

(Weak Intersection Equality) For every subset T of $S(K,A)$ each element of which is of equal informational value and for every $X \in T$, $V(\cap T) = V(X)$.

Given a set of optimal saturatable contractions removing A from K relative to Δ, the previous principle guarantees that its intersection is also an optimal saturatable contraction. Consider now the following important property entailed by these requirements.

(Weak Min) If a finite $T \subset S(K,A)$, $V(\cap T) = min(V(X) : X \in T)$.

Observation 0.1 *Weak monotony, extended weak monotony and weak intersection equality imply Weak Min.*

Proof 0.1 *Focus first on the set $S(K,A)$. List all the saturatable contractions in $S(K,A)$ with S_i and $1 \leq i \leq k$. If $T \subseteq S(K,A)$ consider:*

(P) $V(\cap T) = min(V(S_i) : S_i \in T)$.

(P) holds trivially for T of cardinality 1 and we should show that if it holds for all non-empty subsets of T if cardinality $n < k$ then it holds for T of cardinality $n+1$. Then (P) holds for all non-empty subsets of T of cardinality k, i.e. for all non-empty subsets of $S(K,A)$.

So, assume (P) holds for all subsets of T of cardinality $n < k$. Consider then a particular such T and $\cap T \cap S_j$ with $S_j \notin T$. Let M be the set of maximal and consistent theories entailing $\neg A$ used in order to construct saturatable contractions in T. We will consider the most general case where $S_j = K \cap M_A \cap m_{\neg A}$, were M_A is an intersection of maximal and consistent theories entailing A and $m_{\neg A}$ is a maximal and consistent set of sentences entailing $\neg A$ such that $m_{\neg A} \notin M$. We will also assume that there is a non-empty subset of elements of M_A not used in order to construct saturatable contractions in T. Call this subset S_A. Notice that $X = K \cap S_A \cap m_{\neg A}$ is a contraction in $S(K,A)$ which cannot be in T. Moreover $\cap T \cap S_j = \cap T \cap X$. Notice that X is an alternative contraction with respect to all contractions in T, a fact that will be useful below. So, Weak Min holds for $\cap T$. Let Y be a member of T such that $V(\cap T) = V(Y)$. We need then to show that $V(\cap T \cap S_j) = min(X,Y)$.

Consider first the case $V(X) \leq V(\cap T) = V(Y)$, where Y is a member of T. Let Z be a saturatable contraction entailing $\neg A$ that that does not belong to T and is distinct from X and where $V(Z) = V(X)$. We also want Z to be alternative with respect to X and the all saturatable contractions in T. There need not be such a Z in $S(K,A)$. If that were the case we can always embed hypothetically $\mathcal{M}$ into a larger set $\mathcal{M}'$ in such a way that a new Z with the desired properties is added and in such a way that the original structure of values remains unaltered by the addition of Z. In order to do so consider the (logically finite) underlying language L and its expansion $L' = L \cup t$, where t is a fresh atom not occurring in L. For every theory S in L, where $V(S) = x$, and for every S' over L', such that $S' \cap L = S$, $V(S') = x$. In addition construct an embedding partition Π' containing all the t-maximals of L'. Now let Z be the $\neg t$-counterpart of X. Add to Π' the corresponding $\neg t$-counterparts of each maximal and consistent extension of X – in such a way that each $\neg t$-maximal is added to the cell containing its t-counterpart. It is obvious that Z exists and that $V(X) = V(Z)$.

If $T'_{L'}$ are the saturatable contractions expressible in L' over Π', it should be clear as well that if we manage to prove that $V(\cap T'_{L'}) = min(V(R) : R \in T'_{L'})$ then the result also shows that $V(\cap T') = min(V(S_i) : S_i \in T')$ – the reason being that $V(\cap T'_{L'} \cap L) = V(\cap T')$. It is important to realize for what follows that $Z_{L'}$ is alternative both with respect to

$X_{L'}$ and with respect to $Y_{L'}$. In order not to inflate terminology we will drop from now on the sub-index L'.

Notice first that weak monotony guarantees that $V(X \cap [\cap T]) \leq V(X) \leq V(Y)$. The main supposition in this first case together with extended weak monotony give us in addition: $V(X \cap Z) \leq V(\cap T \cap Z)$.

Now, extended weak monotony also gives us as well that $V(\cap T \cap Z) = V(\cap T \cap X)$. Therefore weak intersection equality guarantees that $V(X \cap Z) = V(X)$. It is also easy to see that $V(X \cap Y) = V(X)$, which is enough to establish that $V(X) = V(X \cap Y) = min(V(W) : W \in \cap[T \cup \{X\}])$. This completes the proof of the first case.

For the second case consider $V(\cap T) = V(Y) < V(X)$. Replace Y in T with X. Since the result contains only n contractions, the min-rule applies. Moreover, the result must carry informational value no less than the original, it carries informational value at least as great as Y. In effect, case 2 has been converted into the first case•

The three postulates that we just introduced are the *core postulates* of the notion of informational value used in contraction. Such notion is called by Levi in (Levi 2004) *damped* informational value (as opposed to the notion of undamped informational value characterized by the first two postulates - which is central in decision-theoretical characterizations of expansion). We will assume as well here the following stronger property.

(Strong Intersection Equality) For every subset T of Φ each element of which is of equal informational value and for every $X \in T$, $V(\cap T) = V(X)$.

Strong intersection equality combined with weak positive monotonicity and extended weak positive monotonicity imply the following:

(Min) If X and Y are potential contractions from K in Φ, $V(X \cap Y) = min(V(X), V(Y))$.

It is obvious that there are other forms of contraction parametrically obtainable by relaxing some of the principles that entail Min (in particular Strong Intersection Equality). Nevertheless, the form of contraction we are studying here is salient, we would like to argue, given its compatibility with the Principle of Entrenchment.

Shells of informational value

The assumption of the core postulates and the stronger Min condition allows us to construct the following notion of *rank*.

DEFINITION 0.2 *Let $I = range(V)$ be a set of indices. For $x \in I$ let R^x be the non-empty set X of maximals in Δ such that for every $Y \subseteq X$, $V((\cap Y) \cap K) = x$.*

Intuitively R^x groups the maximals such that the intersection of each of them with K has value x. By Min the intersection of any subset of them with K, has also value x. We can extend here the notion of rank, by adjudicating ranks to propositions $P \subseteq 2^{\Delta}$.[2]

$\rho^+(P)$ = max(y: $R^y \cap P \neq \emptyset$)

So, for $[A] \subseteq 2^{\Delta}$, $\rho([A]) = y$, where R^y is the set of maximals of largest rank intersecting [A]. Of course, we have then that $\rho^+(\{w\}) = y$ when $w \in R^y$ and for every $Y \subseteq R^y$, $\rho^+(Y) = y$.

We can now introduce the notion of *m-shell of informational value*. The idea of a m-shell is to group together all the ranks R^x where x is greater or equal than the index m.

DEFINITION 0.3 *The* x-shell of informational value $S^x = \cup_{i \in I}^{i \geq x} R^i$. *The* system of shells of informational value *(SS)* $\mathcal{S}$ *is defined as:* $\mathcal{S} = \{S^x : \cup S^x = \Delta\}$

It should be obvious that shells of a SS are nested. Notice in addition that for any maximal $w \in \Delta$ we do not necessarily have $V(w) = \rho^+(w)$. For, by definition, $\rho^+(w)$ = V($K \cap w$). The only constraint imposed by WM in this case is that $\rho^+(w) \leq V(\{w\})$. The equality $V(w) = \rho(w)$ is only guaranteed when the intersection of any two arbitrary definable theories carries the minimal value of the two. But requiring this goes beyond our strongest assumption (Min). So every maximal and consistent set in a SS has a *value-level* which might not coincide with its rank.

A SS for a value function V determines a grading on Δ. So, none of the maximals in Π_K appear in the SS. But of course there are some constraints relating the value of K and the value of the sets in the SS. One important constraint (given by WM) is that $V(K) \geq i$, where S^i is the innermost shell of $\mathcal{S}$. Therefore $V(K)$ is greater than the value of any rank in $\mathcal{S}$.

With the help of the previous definitions we can now characterize our operator of informational value as an operation defined in systems of shells of informational value. We only need an additional definition. Let a sentence A be *rejected in K* if and only if $\neg A \in K$.

DEFINITION 0.4 *Let A be a sentence rejected in K. Then S_A is the union of $[K]$ with the set $X \in \mathcal{S}$ such that $X \cap [A] \neq \emptyset$ and for any other $Y \in \mathcal{S}$, such that $Y \cap [A] \neq \emptyset$, $X \subseteq Y$.*

S_A just picks the union of $[K]$ with the innermost shell in the SS $\mathcal{S}$ for V containing A-maximals. Now we can define some salient operators of informational value.

DEFINITION 0.5 *$\div$ is an operator of informational value for a closed set K if and only if there is a selection function γ such that for all A in L: (i) if $A \in K$, then $K \div A$ = $\cap\gamma(S(K, A))$, where $\gamma(S(K, A)) = \{X \in S(K, A)$: $V(Y) \leq V(X)$ for all $Y \in S(K, A)\}$ and (ii) $K \div A = Cn(K)$ otherwise.*

When the value function V is constrained by WM, the resulting operator is called a *basic* operator of informational

[2] We will represent by $[A]$ the proposition expressed by the sentence A, which is encoded by the set of maximal and consistent sets of sentences extending A. Variables P, Q, etc. are used to denote propositions.

value. When it obeys all core postulates the resulting operator is called a *core* operator of informational value. Finally when V is constrained by all cores postulates plus Min, the resulting operator is called the *standard operator of informational value*. From now on we will mainly work with standard operators of informational value and we will use the notation '$\div$' to refer to them. Specific references and clarifications will be made otherwise.[3]

Observation 0.2 $[K \div \neg A] = S_A$

Proof 0.2 *Consider a maximal and consistent theory w, such that $w \in S_A \cap [A]$. Let C(A) = $S_A - (S_A \cap [A])$. Notice that $M_w = \cap C(A) \cap w$ is a saturatable contraction removing $\neg A$ from K. M_w is one of the contractions of maximal value in $S(K, \neg A)$). In general these contractions have the form $K \cap z \cap (\cap S)$ where z is a maximal in $S_A \cap [A]$ and $S \subseteq C(A)$. In fact, it is easy to see that if either of these conditions fails the resulting saturatable contraction cannot carry maximal value. Say that we consider $(K \cap z \cap (\cap S)) = C \in S(K, \neg A))$, where $z \in [A]$ but $z \notin S_A \cap [A]$. Then C can be represented as $(K \cap z) \cap (K \cap (\cap S))$ where $V(K \cap z) < V(K \cap (\cap S))$. The Min rule requires therefore that $V(C) < V(M_w)$. The reasoning is similar when $S \not\subseteq C(A)$ (in this case one has to focus on the intersection of S with the lowest rank intersecting S). It is then clear that the result of intersecting all the maximal contractions in $S(K, \neg A))$ (i.e. $K \div \neg A$) can be represented by just taking the intersection of all the saturatable contractions of the form M_w, with $w \in S_A \cap [A]$ – the value of this intersection is also maximal by WIE. But then it is obvious that $[K \div \neg A] = S_A$, as desired*•

Given a value function V defined on Φ it is possible to define the following useful relation:

DEFINITION 0.6 $P \leq_V Q$ *if and only if* $V(P) < V(Q)$

In particular given a theory of reference K and a value function this relation orders all the potential contractions for the theory K. Moreover, it is immediate how to retrieve a relation $\leq_V$ from the system of shells for V and K. This can be done as follows:

Observation 0.3 *If P, Q are potential contractions of K then $P \leq_V Q$ if and only if there is S^x and S^y, such that $S^x \subseteq S^y$, R^x is the minimum rank intersecting $[P]$ and R^y is the minimum rank intersecting $[Q]$.*

This property flows from Min. Notice that if $P \leq_V Q$ this is so independently of the ranks of $[P]$ and $[Q]$ in the SS for V and K. Propositions in 2^Δ are ordered by $\leq_V$ in virtue of an index different than its rank. In fact, if $P \subseteq 2^\Delta$ we can define the following index of informational value ρ^-:

$\rho^-(P)$ = min(y: $R^y \cap P \neq \emptyset$)

Notice that for any $P \subseteq 2^\Delta$ we have that $V((\cap P) \cap K)$ = $i(P)$. Nevertheless $V((\cap P)$ need not coincide with $i(P)$ – the theory $\cap P$ could have some value lower than $\rho^-(P)$. The index ρ^- has some obvious properties. For example: $\rho^-(P \cup Q)$ = min($\rho^-(P)$, $\rho^-(Q)$). And the index of informational value can be combined with ranks to give a simple definition of contraction. For any A rejected in K:

Corollary 0.1 $[K \div \neg A] = \cup\{P \subseteq 2^\Delta$*:* $\rho^-(P) = \rho^+([A])\} \cup [K]$.

In words, in order to construct $[K \div \neg A]$ we take the union of $[K]$ with all the propositions in 2^Δ such that their index of informational value equals the rank of A. It is quite obvious that S_A is one of these propositions. We can now go back to some additional properties of $\leq_V$:

Observation 0.4 *(d1) Either* $[K \div A] \leq_V [K \div B]$*, or* $[K \div B] \leq_V [K \div A]$

Which, in turn, means that we can easily establish a pretty strong property of informational value contractions, namely that: (d1) Either $[K \div A] \subseteq [K \div B]$, or $[K \div B] \subseteq [K \div A]$. This property will be used later on in the proof of our main result.

Shells of informational value are structures which, at first sight at least, might be easy to conflate with Spohn's *ranking systems* (Spohn 1988), (Spohn 2002). A ranking function κ is a function from $\mathcal{M}$ to the set of extended non-negative integers $\mathcal{N}^+ = \mathcal{N} \cup \{\infty\}$, such that $\kappa(w)$ = 0, for some $w \in \mathcal{M}$. For each proposition $P \subseteq \mathcal{M}$ the *rank* $\kappa(P)$ of P is defined by $\kappa(P)$ = min $\{\kappa(w)$: $w \in P\}$ and $\kappa(\emptyset) = \{\infty\}$.

Ranks in Spohn's system are best interpreted as *grades of disbelief*. $\kappa(P)$ = 0 says that P is not disbelieved at all. It does not say that P is believed; this is rather expressed by $\kappa(P^c) > 0$, i.e., that non-P is disbelieved (to some degree). The set $C_\kappa = \{w$: $\kappa(w) = 0\}$ is called the *core* of κ and C_κ is the strongest proposition believed (to be true) in κ. So, if A^c is believed to be true in κ, one way of representing the contraction of A^c from C_κ is to take the union of C_κ with the set of least disbelieved A points, i.e. $\{w$: $\kappa(w) = \kappa(A)\}$. This is a simple way of defining an AGM contraction in this setting.

Ranking systems are different, both formally and conceptually from shell systems. Notice first that since we are working with finite partitions we can define shells also with range over $\mathcal{N}$, but in our case the domain is restricted to $\Phi = \{X : X = \cap Y$, with $Y \in 2^\Delta\}$. Moreover in the case of rankings one proceeds by assigning first natural numbers to points (maximal and consistent theories in this case) and then ranks are assigned to propositions in an unproblematic manner. In our case an assignment of values to maximal and consistent theories does not fully determine the ranks of contractions for a theory of reference K. In fact, notice that we can also define both $\kappa^-(P)$ = min $\{V(w)$: $w \in P\}$ and $\kappa^+(P)$ = max $\{V(w)$: $w \in P\}$. The second notion is not usually defined in Spohn's systems. But even if we were to use it notice that, given any proposition P in 2^Δ, nothing guarantees that $\kappa^+(P) = \rho^+(P)$ or that $\kappa^-(P) = \rho^-(P)$. As we explained before, the maximal and consistent theories w in Δ receive a rank $\rho_-(\{w\}) = \rho_+(\{w\}) = x$, for R^x such that $w \in R^x$. But this rank need not coincide with w's

[3] It is important to keep in mind here that the core postulates characterize the notion of damped informational value central to the notion of contraction, while the first two are enough to characterize the notion of informational value important in expansion.

value-level (measured by $\kappa_+(\{w\})$ or $\kappa_-(\{w\})$).[4]

In our framework the value-level of propositions is, of course, quite useful. It puts a constraint on permissible rankings ρ and it is crucial for determining iterated contractions (and revisions). But the value-level of points (maximal and consistent sets) does not fully determine ranks in our sense (ρ) and the 'upper' and 'lower' point-ranks κ do not play a significant role in our proposal. Moreover even if we were to restrict our attention exclusively to ranks in our sense to the detriment of value levels we would need to use *both* the 'upper' and 'lower' ranks ρ^+ and ρ^-. So ranking systems and systems of shells are quite different. Both induce an indexed grading, but they induce gradings over different domains and the algorithm for assigned grades to propositions is different in each account. Spohn's ranking functions assign ranks to propositions identical to the degree of disbelief of its least disbelieved points, while in our account the rank of a proposition P (relative to K and Δ) is identical to the degree of informational value carried by the P-maxichoice contractions of K of maximal value. Notice that this notion of 'upper' rank has no operative counterpart in Spohn's system.[5]

All these formal differences flow from the central fact that the intended interpretation of grades in each account (Spohn's and ours) is fundamentally different. Spohn's account is a purely doxastic account where ranks can be (roughly) interpreted as the orders of magnitude of infinitesimal probabilities. As we explain above Spohn's main goal is to develop a non-probabilistic articulation of *degrees of disbelief*. In our account the grades are induced by a probability-based function measuring the *value* of information.

Mild contractions

Here we will proceed axiomatically. The axioms used here are well known in the literature and their names are also more or less standard (see, for example, (Hansson 1999)).

($\div$ 1) $K \div A = Cn(K \div A)$ [closure]

($\div$ 2) $K \div A \subseteq K$ [inclusion]

($\div$ 3) If $A \notin K$ or $A \in Cn(\emptyset)$, then $K \subseteq K \div A$ [vacuity]

($\div$ 4) If $A \notin Cn(\emptyset)$, then $A \notin K \div A$ [success]

[4]The 'upper' rank κ^+ and the 'lower' rank κ^+ can be used in order to determine an epistemic ordering solely on the basis of point-value utility. The procedure, suggested by John Collins in (Collins 2002), consists (roughly) in stipulating that proposition P is preferred to proposition Q if and only if the upper rank of P is greater than the upper rank of Q and the lower rank of P is no worse than the lower rank of Q. Or, alternatively that the lower rank of P is better than the lower rank of Q and the upper rank of P is no worse than the upper rank of Q. The procedure allows for incomparability of preference. This view of preference, nevertheless, does not satisfy postulates that are typical of probability-based notions of utility, like Weak Monotony, and therefore is quite different from the one presented here.

[5]Such notion can, of course, be defined for Spohn's ranking functions as well. According to Spohn's official interpretation the 'upper' rank of a proposition would be determined by the degree of disbelief assigned to its most disbelieved points.

($\div$ 6) If $Cn(A) = Cn(B)$, then $K \div A = K \div B$ [extensionality]

($\div$ 7) If $A \notin Cn(\emptyset)$, then $K \div A \subseteq K \div (A \wedge B)$ [antitony]

($\div$ 8) If $A \notin K \div (A \wedge B)$, then $K \div (A \wedge B) \subseteq K \div A$ [conjunctive inclusion]

All the conditions, except antitony, are AGM properties. On the other hand there is a notorious postulate, AGM's axiom of recovery, which is not in the previous list and that is not derivable from the previous list:

($\div$ 5) $K \subseteq Cn((K \div A) \cup \{A\})$ [recovery]

Antitony is perhaps the most controversial postulate from the list. For example Hansson reports in (Hansson 1999) that antitony does not hold '[...] for any sensible notion of contraction'; while Rott and Pagnuco report in page 513 of (Rott & Pagnucco 1999) that '[...] intuitively antitony makes quite a bit of sense'.

The axiomatic base given here is exactly the one proposed in (Rott & Pagnucco 1999) to characterize *severe withdrawals*. Here we will show that this axiomatic base is indirectly supported by the intuitiveness of the postulates of Economy and Entrenchment. This gives, in turn, indirect support to Antitony.

A representation result for mild contractions

We will focus first on a soundness result. It is interesting to remark that this result can be partitioned in two lemmas.

Lemma 0.1 *Any basic operator of informational value obeys postulates ($\div$ 1)-($\div$ 4), ($\div$ 6) and ($\div$ 8).*

A proof of this fact was given in (Hansson & Olsson 1995). The proof of ($\div$ 8) requires some work if only weak monotony is assumed (as a matter of fact this proof is one of the substantial arguments presented in (Hansson & Olsson 1995)). Things are simpler when we have an operator of informational value which obeys the Weak Min. The use of shells of informational value permits a more direct proof - we will offer below a simpler proof of ($\div$ 8) to illustrate this.

Lemma 0.2 *Any standard operator of informational value satisfies all the postulates of mild contractions.*

Proof 0.3 *Let's first focus on Conjunctive Inclusion. Assume that $A \notin K \div (A \wedge B)$. We need to show that $K \div (A \wedge B) \subseteq K \div A$. We have that $[K \div A] = S_{\neg A}$, i.e. we know that $[K \div A]$ is identical to the smallest m-shell of informational value intersecting $[\neg A]$. And by the same token we have that $[K \div (A \wedge B)] = S_{\neg(A \wedge B)}$. Given the assumption of the proof we have that $S_{(\neg A \vee \neg B)} \cap [\neg A] \neq \emptyset$. This is enough to guarantee that $S_{(\neg A)} \subseteq S_{(\neg A \vee \neg B)}$. Therefore $[K \div A] \subseteq [K \div (A \wedge B)]$, and $K \div (A \wedge B) \subseteq K \div A$, as desired.*

Considering Antitony, we have to show that when $A \notin Cn(\emptyset)$, $K \div A \subseteq K \div (A \wedge B)$. So, assume that $A \notin Cn(\emptyset)$. We need to show that $K \div A \subseteq K \div (A \wedge B)$. Or, in terms of collections of maximals, we need to show that $[K \div (A \wedge B)] \subseteq [K \div A]$. So, we need to show that $S_{\neg(A \wedge B)} \subseteq S_{\neg A}$, or, *equivalently that $S_{(\neg A \vee \neg B)} \subseteq S_{\neg A}$. By the assumptions we know that $(A \wedge B) \notin Cn(\emptyset)$. So, since $[\neg A] \subseteq [\neg A \vee \neg B]$, we have that $S_{\neg(A \wedge B)} \subseteq S_{\neg A}$, as desired. Of course, as we showed in a previous observation, the appeal to the identity $[K \div \neg A] = S_A$ in the proof presupposes that our operator of informational value obeys not only all the core postulates but also Min, i.e. that it is a standard operator* $\bullet$

Aside from soundness we can also establish the following completeness result:

Theorem 0.1 *If '$\div$' is a mild contraction function obeying the correspondent postulates, then '$\div$' can be represented as an operator of informational value.*

Proof 0.4 *We need to show that starting with an operator $\div$ obeying the postulates of mild contractions we can explicitly construct a system of shells of informational value. We have to show as well that the operator $\div'$ obtained from the defined system of shells of informational value by requiring $[K \div' A] = S_{\neg A}$, where $S_{\neg A}$ is the smallest m-shell of informational value intersecting $\neg A$, is identical to $\div$.*

Let's first focus on how to construct a Grove system in terms of the operation $\div$. The method for constructing a Grove system from a contraction operation is well-known (Grove 1988). So, we will only outline here the main steps of the proof, skipping unnecessary details. The proof sketched here follows a suitable modification of Grove's original proof as presented in (Rott & Pagnucco 1999). The central idea is quite simple: a Grove system of spheres **S** *centered on $[X]$, is determined by identifying a sphere in* **S** *with the collection $[X \div A]$ for some $A \in L$. More precisely:*

(d) $X_A = [X \div A]$

In addition define the system of spheres **S** *as follows;*

S *= $\{X_A$: $A \in L\} \cup \mathcal{M}$, when $K \neq L$, and* **S** *= $\{X_A$: $A \in L\} \cup \mathcal{M} \cup \emptyset$ otherwise.*

The gist of this first part of the proof consists on showing that the system **S** *is indeed a Grove system of spheres centered on $[X]$. Since this is important for the rest of the result we are showing, we will remind the reader immediately of the definition of a Grove system of spheres centered on $[X]$.*

Let **S** *be a collection of subsets of the set of all L-maximals of $\mathcal{M}$.* **S** *is a system of spheres, centered on X =[K]$\subseteq \mathcal{M}$ and satisfying:*

(1) **S** *is totally ordered by $\subseteq$.*

(2) *X is the $\subseteq$-minimum of* **S**.

(3) *$\mathcal{M}$ is the $\subseteq$-maximum of* **S**.

(4) *If $A \in L$ and $\emptyset \neq [A] \in 2^{\mathcal{M}}$, then there is a smallest sphere S_A in* **S** *intersecting the set $[A]$.*

Condition (1) is directly satisfied in virtue of (d) above and the fact that the following property can be deduced from the axioms of mild contractions:

(d1) Either $K \div A \subseteq K \div B$ or $K \div B \subseteq K \div A$.

Conditions ($\div$ 2) and ($\div$ 3) guarantee that $K \div \mathbf{true} = K$. This and (d) are enough to show that $X = [K]$ is a sphere. That this is the innermost sphere follows immediately from ($\div$ 2) and (d). This takes care of condition (2). Condition (3) is automatically satisfied by the given definition of **S**.

Condition (4) is slightly harder. Let A be such that $\emptyset \neq [A] \in 2^{\mathcal{M}}$. Now we need to show that there is a sphere $U \in$ **S***, such that $U \cap [A] \neq \emptyset$ and for every other $V \in$* **S***, such that $V \cap [A] \neq \emptyset$, we have $U \subseteq V$. The basic idea of the proof, which we skip here, is to show that $[K \div \neg A] = X_{\neg A}$ satisfies this constraint.*

The proof of condition (4) establishes that (when A is not a tautology) $S_{\neg A} = [K \div A] = X_A$, where $S_{\neg A}$ denotes the smallest sphere intersecting $[\neg A]$ - we used basically the same notation $S_.$ for shells above, indicating the smallest m-shell intersecting $[\neg A]$. This fact will be useful below.

Now we should focus on the second step of the proof, namely the construction of a system of shells of informational value for the constructed system of spheres. Here is the recipe in order to do so.

First construct a system of ranks out of the given system of spheres as follows: index first spheres with natural numbers starting with 0 assigned to the innermost sphere in such a way that S_i denotes the set of maximals in the sphere indexed by i. This is done via an indexing function mapping propositions to natural numbers. Then define a function δ from the range of the indexing function to propositions, such that $\delta(k) = S_{k+1}$ - S_k. As a second step assign an arbitrary V-value x to the innermost sphere of **S***, $[K]$ as long as x is greater than k, where k is the index of the outermost sphere S_k.*

As a third step we need to give a value to each maximal not included in $[K]$. In order to do so assign a uniform value $y > x$ to each maximal in $\delta(0)$ and, in general, for every $\kappa(i + 1)$, for $i \geq 0$, assign a uniform value $z > z'$ to the maximals in $\kappa(i + 1)$, where z' is the uniform value of maximals in $\kappa(i)$.

As a fourth step we need to define ranks and m-shells of informational value. In order to do so we need to re-index the ranks we just defined from the Grove system. Assign to $\delta(0)$ a positive index $x' < x$ and, in general, for every $\delta(i + 1)$, with $i \geq 0$, assign to $\delta(i + 1)$, a positive index z, where $z < z'$ and z' is the value of $\delta(i)$. So, for an arbitrary $\delta(i)$, we have a positive number m assigned to it such that $m < x$. Create then the ranks of informational value R_i^m and the corresponding shells S_i^m as follows: R_i^m = $\delta(i)$ and $S_i^m = S_i$. The last definition allows us to complete the definition of the V-measure, by requiring that (i) for every $[Y] \subseteq R_i^x$, $V((\cap Y) \cap K) = x$; (ii) that for every $[Y] \subseteq S_i^x$, such that R_i^x is the outermost rank such that $R_i^x \cap [Y] \neq \emptyset$, $V((\cap Y) \cap K) = x$; and (iii) for any theory T incompatible with K, let $V(T) = min(V(m) : m is a maximal and consistent theory extending K)$.

It is obvious that V, as defined, is a function. It is also clear it satisfies WM. We need to check that for any two maxichoice contractions of K, X, Y, such that $[Y] \subset [X]$, then $V(X) \leq V(Y)$. Let $R^{[X]} = R_i^m$ be the outermost rank intersecting $[X] - [K]$. Then it is clear that the outermost *rank intersecting $[Y] - [K]$, $R^{[Y]} = R_{i'}^{m'}$ is such that $i' \leq i$ and $m' \geq m$. Since $[Y] - [K]$ is a subset of $R_{i'}^{m'}$ $[Y]$ receives value m', and since $[X] - [K]$ is a subset of R_i^m $[X]$ receives value m. And since $m' \geq m$, weak monotony is satisfied.*

Consider now $Y = \cap X$, where X is a family of maxi-choice contractions of K, and its associated set $[Y]$. Consider again the outermost rank intersecting $[Y] - [K]$, $R^{[Y]} = R_i^m$. Then according to the proposed explicit definition of V, $V(Y) = m$. Take now an arbitrary $Z \in X$ such that $[Z] - [K]$ does not intersect $R^{[Y]}$. Then, by construction, $V(Y) < V(Z)$. And if $[Z]-[K]$ intersects $R^{[Y]}$, $V(Z) = m$. So, clearly we do have that $V(Y) = min\{V(Z) : Z \in X\}$.

Finally we need to check that an operation of contraction $\div'$ defined from the explicitly constructed system of shells of informational value, coincides with the operator $\div$ characterized by the postulates of mild contractions. We will define:

$[K \div' \neg A] = S_A$

where S_A is the smallest m-shell of informational value intersecting $[A]$ union $[K]$. The non-trivial case to consider is when A is not a tautology. We will work under this assumption. Assume first that $B \in K \div \neg A$. This entails that $[K \div \neg A] \subseteq [B]$. The proof of condition (4) for Grove systems above tell us that $[K \div \neg A]$ is a sphere, and since in our construction all the resulting m-shells are the spheres that we constructed in the first step of the proof (conveniently indexed and where $[K]$ has been substracted), $[K \div \neg A] - [K]$ is also an m-shell of informational value. Moreover, the proof of (4) also tells us that the smallest sphere overlapping $[A]$ is identical to $[K \div \neg A]$. Therefore the smallest m-shell of informational value overlapping $[A]$ is identical with $[K \div \neg A] - [K]$. Therefore we have that $[K \div' \neg A] \subseteq [B]$, and $B \in K \div' \neg A$, as desired. Proving that $K \div' \neg A \subseteq K \div \neg A$ only requires reversing the strategy used for the RTL inclusion•

The Principle of Entrenchment and informational value

So far nothing has been said about the Principle of Entrenchment invoked in the introduction. Let's first introduce a relation of entrenchment formally. Let $\leq$ be an ordering of the sentences of L. $\leq$ is a relation of entrenchment for a theory K if and only if the following postulates are satisfied

(i) If $A \leq B$ and $B \leq C$, then $A \leq C$ (transitivity)

(ii) If $A \in Cn(\emptyset)$, then $B \leq A$ (dominance)

(iii) $A \leq A \wedge B$ or $B \leq A \wedge B$ (conjunctiveness)

(iv) If $K \neq L$, then $A \leq B$ for every $B \in$ L if and only if $A \notin K$ (minimality)

(v) If $A \leq B$ for every $A \in$ L, then $B \in Cn(\emptyset)$

Now, we remind the reader that our principle of entrenchment said that in giving up a sentence A from the current view one should preserve the sentences better entrenched than A. This translates formally into:

DEFINITION 0.7 *If $A \in K$ and $A \notin Cn(\emptyset)$, then $K \div A = K \cap \{B : A < B\}$ and K otherwise.*

This simple and elegant manner of characterizing contraction in terms of entrenchment was first proposed by Rott in (Rott 1991). AGM cannot be characterized in terms of entrenchment in this way. The non-equivalent bridge connecting entrenchment and AGM contractions stipulates that: $K \div A = K \cap \{B : A < A \vee B\}$

Observation 0.5 *(Rott and Pagnuco) If $\leq$ satisfies the postulates (i) to (v) then the function $\div$ obtained from $\leq$ by definition 5.1 is a mild contraction.*

Therefore, via the completeness result offered above, the informational value contractions can be retrieved by using the Principle of Entrenchment. Moreover, we can also appeal to a result recently proved by Rott and Pagnuco (and to our completeness result) in order to retrieve the relevant notion of entrenchment from an informational value contraction.

DEFINITION 0.8 *If $A \leq B$ if and only if $A \notin K \div B$, or $B \in Cn(\emptyset)$.*

Observation 0.6 *(Rott and Pagnuco) If $\div$ is a mild contraction then the relation $\leq$ obtained via definition 5.2 satisfies the postulates (i) to (v) for the entrenchment relation.*

So, the Principle of Entrenchment and the Principle of Economy give exactly the same account of contraction. AGM contractions are also mirrored by a corresponding notion of entrenchment, but this notion does not obey the Principle of Entrenchment. One of the consequences of this divergence is the fact that AGM contractions satisfy the controversial principle of recovery, which is not satisfied by mild contractions (alias severe withdrawals according to Rott and Pagnuco's terminology).

Conclusions

The article focuses on determining the logical commitments entailed by our formulation of the principle of Economy (in terms of informational value) and the principle of Entrenchment. Together they give a holistic justification of the axioms of mild contractions. We pointed out above, nevertheless, that the principle of Economy is more accommodating than the principle of Entrenchment. In fact, minimizing loses of informational value under the constraints imposed by the three core postulates of (damped) informational value is coherent with endorsing our version of Economy. But this does not guarantee the satisfaction of:

($\div$ 7) If $A \notin Cn(\emptyset)$, then $K \div A \subseteq K \div (A \wedge B)$ [antitony]

In order to satisfy Antitony we need to assume in addition the full force of Min. One could justify Min (and therefore Antitony) by pointing out that this is exactly the requirement that is needed in order to have a theory of contraction where both Economy and Entrenchment are satisfied.

Apparent counterexamples against Antitony are easy to concoct, nevertheless, by considering scenarios where the sentences A and B (in the formulation of Antitony) express propositions that in some pre-systematic sense are irrelevant to each other. The serious study of these apparent counterexamples requires nevertheless a minimal articulation of the notion of relevance presupposed by the examples. Even when formalizing a notion of relevance is a complicated problem in itself, there are some proposals in the literature that articulate precisely the degree of relevance that two formulas might have to each other (Chopra & Parikh 2000), (Parikh 1999). It should be pointed out here in passing, nevertheless, that an unmodified version of our model is not insensitive to the problem of relevance. The theory of contraction we are proposing has various contextual parameters that are useful for this problem. In particular the space of options $\mathcal{M}$ is determined by the adoption of a basic partition encoding a set of potential answers to a problem. So our model, like some of the existing syntactic approaches to the problem of relevance focuses on certain relevant (syntactic) partitions of the set of all expansions of the basic theory LK. Unlike most of the existing theories of belief change we restrict the set of potential options and the set of potential contractions to a set relevant to the solution of a cognitive problem from scratch. So, counterexamples in terms of relevance are not crucially threatening to the theory presented here. In any case, it is important to realize that potential counterexamples against Antitony do not seem to threaten the core postulates (i.e. the notion of informational value) but the full force of Min.

There are some obvious extensions of our proposal and some ways of weakening it as well. An obvious extension is the consideration of the infinite case. In this case the density of the real interval $[0, 1]$ is important. Otherwise we can as well define a value function with range over the natural numbers. In addition, it would be interesting to study the shape of sequential change of view in this setting. It is also clear that more expressive languages (taking advantage of the numerical features of the semantics in terms of system of shells) can also be studied.

Concerning weakenings of the proposed theory, when Min fails there are a variety of possible contraction theories obeying the core postulates. Their parametric study can help to identify the constraints on value that completely characterize various *withdrawal operators*. Finally it would be interesting to study the more realistic case which permits value to be indeterminate (we conjecture that this is also a weakening of the theory that leads to the failure of Antitony).

Acknowledgements Previous versions of this paper were presented during the Spring semester 2004 both at the Logic and Games Seminar, CUNY/GC and at the Belief Revision and Decision Group at Columbia University. We are grateful for useful comments received during these meetings. We want to thank also the useful comments provided by two anonymous referees.

References

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

Arló-Costa, H. 2004. Decision-theoretic contraction and sequential change. In Olsson, E., ed., *Knowledge and Inquiry: Essays on the Pragmatism of Isaac Levi*. Cambridge University Press, Cambridge Studies in Probability, Induction and Decision Theory.

Chopra, S., and Parikh, R. 2000. Relevance sensitive belief structures. *Annals of Mathematics and Artificial Intelligence* 259–285.

Collins, J. 2002. Belief revision derived from preference. Talk Philosophical Logic and Applications, CUNY Graduate Center.

Gärdenfors, P., and Rott, H. 1988. Belief revision. In D.M. Gabbay, C. H., and Robinson, J., eds., *Handbook of Logic in Artificial Intelligence and Logic Programming, Volume IV*. Oxford University Press.

Gärdenfors, P. 1988. *Knowledge in Flux*. Cambridge, Mass.: MIT Press.

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

Hansson, S. O., and Olsson, E. J. 1995. Levi contractions and agm contractions: a comparison. *Notre Dame Journal of Formal Logic* 36:103–119.

Hansson, S. O. 1999. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Harman, G. 1999. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Levi, I. 1980. *The Enterprise of Knowledge*. Cambridge, Mass.: MIT Press.

Levi, I. 1991. *The fixation of belief and its undoing*. Cambridge: Cambridge University Press.

Levi, I. 2004. Mild contraction: Evaluating loss of information due to loss of belief. Forthcoming Cambridge University Press.

Parikh, R. 1999. Beliefs, belief revision and splitting languages. In Moss L. Ginzburg, J., and de Rijke M., eds., *Logic, Language and Information, vol. 2*. CSLI, Lecture Notes, No. 96. 266–278.

Quine, W. 1951. Two dogmas of empiricism. *Philosophical Review* 60:20–43.

Rott, H., and Pagnucco, M. 1999. Severe withdrawal (and recovery). *Journal of Philosophical Logic* 28:501–547.

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

Rott, H. 2000. Two dogmas of belief revision. *The Journal of Philosophy* 97:503–22.

Spohn, W. 1988. Ordinal conditional functions. a dynamic theory of epistemic states. In Harper, W., and Skyrms, B., eds., *Causation in Decision, Belief Change and Statistics, vol. II*. Kluwer: Dordrecht. 105–134.

Spohn, W. 2002. Laws, ceteris paribus conditions, and the dynamics of belief. *Forschungsberichte der DFG-Forschergruppe Logik in der Philosophie, No. 84*.

van O. Quine, W., and Ullian, J. S. 1978. *The Web of Belief*. New York: McGraw-Hill.