RATIONALLY CHOOSING BELIEFS: SOME OPEN QUESTIONS*

HORACIO ARLÓ-COSTA
CMU

Abstract

Carlos Alchourrón, Peter Gärdenfors and David Makinson published in 1985 a seminal article on belief change in the *Journal of Symbolic Logic* (Alchourrón et al., 1985). Researchers from various disciplines, from computer science to mathematical economics to philosophical logic, have continued the work first presented in this seminal paper during the last two decades. This paper explores some salient foundational trends that interpret the act of changing view as a decision. We will argue that *some* of these foundational trends are already present, although only tacitly, in the original article by the AGM trio. Other accounts decidedly depart from the view of contraction and revision presented in this seminal paper. I shall survey various types of theories that progressively depart form the axiomatic treatment defended by AGM. First, I consider theories where rational agents are considered as maximizers as opposed to optimizers (in the sense of (Sen, 1997a)). Second, I consider which feasible set to use in contraction understood as a cognitive decision. This leads to rethink the very notion of what minimal change in contraction is. I shall conclude with some philosophical reflections concerning the sort of epistemological voluntarism that is tacit in seeing change in view as a rational choice.

KEY WORDS: belief change - contraction- revision - cognitive decision- rational choice.

Resumen

Carlos Alchourrón, Peter Gärdenfors y David Makinson publicaron en 1985 un artículo seminal sobre cambio de creencias en *Journal of Symbolic Logic* (Alchourrón et al., 1985). Investigadores de varias disciplinas, desde la ciencia de la computación hasta la economía matemática y la lógica filosófica, han continuado en las dos últimas décadas esta línea de investigación. Este trabajo explora algunos aspectos fundacionalmente salientes que interpretan el acto de cambio de vista como una decisión. Argumentaremos que algunos de esos aspectos fundacionales ya estaban presentes, aunque solo tácitamente, en el artículo original del trío AGM. Otros abordajes parten decididamente de la contracción y revisión tal como fueran presentadas en el trabajo seminal.
Inspeccionaré varios tipos de teorías que progresivamente parten del tratamiento axiomático defendido por AGM. Primero, considero teorías donde los agentes racionales aparecen como maximizadres opuestos a los optimiadores (en el sentido de (Sen, 1997a)).

* I learned a great deal about belief change from Carlos Alchourrón, who directed my Licentiate Degree at the University of Buenos Aires during the late 1980s. Much of my work since then has been greatly influenced by his ideas, his philosophical style and by the constant pursuit of excellence that characterized his work. This essay is dedicated to his memory.

Segundo, me pregunto cuál conjunto derrotable debe usarse en una contracción entendida como una decisión cognitiva, lo cual lleva a repensar la importante cuestión de en qué consiste la noción de cambio mínimo en la contracción. Concluiré con algunas reflexiones filosóficas acerca de la suerte de voluntarismo epistemológico que está tácito en la concepción del cambio como una opción racional.

PALABRAS CLAVE: cambio de creencias - contracción - revisión - decisión cognitiva - elección racional.

1. Introduction

The view of *partial meet contraction* and *partial meet revision* presented in (Alchourrón et al., 1985) continues to be one of the most salient paradigms for the study of belief change. The views of an agent at a given time are encoded via a closed theory K representing a state of *epistemic equilibrium* for the agent. Equilibrium here means both inductive as well as deductive equilibrium. Then the AGM trio studied three salient cases: the case where a new epistemic input (represented by a sentence) is added to K together with the logical consequences of the addition (expansion); the case when a new input (eventually inconsistent with K) is added but in order to maintain consistency some of the old sentences in K are deleted (revision); and the case when some sentence in K is retracted without adding new information (contraction). In the latter case for the resulting view to be closed under logical consequence some other sentences from K have to be given up.

In order to calculate the contraction of K with a sentence A ($K \div A$) the AGM trio relied on the following principle:

Informational Economy: Minimize information loss in contraction.

A natural consequence of this principle is to focus on the set $K \perp A$ of maximal subsets of K that fail to entail A. The underlying idea is that contraction would require selecting a unique member $\gamma(K \perp A)$ from that set. The function γ used in this selection can be seen as a selection function, which picks out an element $\gamma(K \perp A)$ from $K \perp A$, when $K \perp A$ is non-empty. This type of operation is known in the AGM tradition as a *maxichoice* contraction.

Although maxichoice contraction is recommended by the principle of Informational Economy it is well known that maxichoice contractions are *too large*. In fact, if we see the revision with a sentence A as the result of first contracting with the negation of A and the expanding with A:

Levi Identity: $K*A = (K \div \neg A)+A = Cn((K \div \neg A) \cup \{A\})$[1]

Then it is clear that the revision defined from maxichoice contraction via the Levi Identity would yield maximal and consistent set of sentences, a result that the AGM trio deemed unacceptable. On the other hand taking the intersection of all the members of $K\perp A$ leads to an equally unacceptable solution. Contractions in this case are *too small*.

A solution envisaged in (Alchourrón et al., 1985) was to focus instead on a selection $\gamma(K\perp A)$ and then take the intersection of that selection.[2] The contractions ensuing from this procedure are the so-called partial meet contractions studied in (Alchourrón et al., 1985). It is easy to see that the motivation of this type of contraction cannot be the principle of Informational Economy. A justification of partial meet contraction, if at all possible, requires a more sophisticated defense than the one provided by a simple application of Informational Economy. We will critically review below some of the existing justifications, most of which have relatively recently appeared.

It is instructive to notice that so far we have not put any constraint on the selection function γ. One of the usual articulations of the role of g is to see it as providing a way of selecting the 'best' elements of $(K\perp A)$. In most of the standard presentations γ is seen as a function that guarantees a non-empty selection from the set $(K\perp A)$ when this set is non-empty. The idea of picking out the 'best' elements of $(K\perp A)$ can be made more precise by assuming that there is an ordering $\leq$ of the maximal subsets in $(K\perp A)$. Then we can define a function γ as follows:

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

Notice that if we define the selection function in this way, this definition plus the assumption that the selection is non-empty when the remainder set $K\perp A$ is non-empty, yields that the ordering in question has to have some basic properties. On the one hand it has to be acyclic and on the other hand it has to be complete. This fact it usually not mentioned in the standard presentations of this material. Let us focus for a moment of the assumption of completeness. It is clear that if the underlying ordering fails to be complete then we can have cases where the selection func-

[1] Cn is an operator or logical consequence obeying classical Tarskian postulates.

[2] The appeal to intersections can be justified decision theoretically as the use of a rule for resolving ties. Nevertheless this has not been the usual strategy followed by researchers working in the AGM tradition.

tion fails to guarantee a non-empty selection. Say that the set K⊥A contains only three elements K, K' and K", and that we only have $K \leq K'$ (K" being incomparable both with K' and with K). In this case there is no 'best' element in K⊥A, i.e. there is no element of K⊥A that dominate all its remaining elements –in spite of the fact that K⊥A is non-empty.

If we assume in addition that the relation ≤ is transitive we see that the relation in question has to be a weak order. A contraction function determined from ≤ via the selection function γ is usually called a *transitively relational* partial meet contraction function.

A first step towards seeing the relevance of rational choice for the problem of belief change consists in realizing that the selection function γ can be seen as the type of choice function used in the economic theory of revealed preference. Seeing γ(S) from this point of view helps to realize that the agent we are representing is facing a *decision problem* represented by the set S. The idea is that γ(S) is the choice set for S in the sense that the elements of S are regarded as equally adequate choices for an agent whose values are represented via the selection function γ. Choice sets are then considered as sets of 'best' elements. The equation (γ) above produces a way of articulating this idea by presenting the selection as an act of *optimization*. There is nevertheless a second way in which one can articulate the role of the selection function as instantiating a process of *maximization* (Sen, 1997a):

$$(\mu)\ \mu(K\perp A) = \{K' \in K\perp A\text{: not}(K' < K'') \text{ for all } K'' \in K\perp A\}$$

where '<' is a strict (asymmetric) preference relation. Maximization permits a different insight into the decision process tacit in determining a contraction. Obviously even if we continue to assume that μ(K⊥A) is non-empty when (K⊥A) is non-empty, the relation in question can be very weak. It certainly need not be negatively transitive. It can be, for example, a quasi-transitive relation although it is assumed that it should be non-cyclical.

There are various important motivations for studying the notion of contraction determined via the process of maximization of relations that are, for example, quasi-transitive. One of them is the study of agents whose values are indeterminate. One such agent can be represented as holding various (potentially conflicting) standards of value when facing a contraction. For example, one dimension of value could be simplicity and another coherence. Even if one assumes that each of these standards of value can be represented via a weak ordering, the agreement of all these orderings (the set of pairs holding in all the given standards of value) need not be complete (it would be quasi-transitive).

The study of the notion of liberal contraction obtained via this process of maximization of quasi-orderings has been recently studied in (Arló-Costa, 2006). We will see below that this type of contraction constitutes a first departure from the standard AGM contractions. In order to see this it is useful to familiarize oneself with very important constraints on choice functions. The notation derives form the seminal work of Amartya Sen in this area (Sen, 1997):

(α) For all S, S' $\in$ **S**, if $S \subseteq S'$, then $S \cap \gamma(S') \subseteq \gamma(S)$

(γ) For all $\{S_i: i \in I\} \subseteq$ **S** such that $\cup\{S_i: i \in I\} \in$ S, $\cap\{\gamma(S_i): i \in I\} \subseteq \gamma(\cup\{S_i: i \in I\})$.

(ε) For all S, S' $\in$ **S** such that $S \subseteq S'$, if $\gamma(S') \subseteq \gamma(S)$, then $\gamma(S) \subseteq \gamma(S')$

(β+) For all S, S' $\in$ **S** such that $S \subseteq S'$, if $\gamma(S') \cap S \neq \emptyset$, then $\gamma(S) \subseteq \gamma(S')$.

The first postulate is sometimes as well known as 'Chernoff's axiom' and the last as 'Arrow's axiom', while ε is sometimes known as the 'Superset Axiom'. The set **S** denotes the domain of the choice function. Hans Rott (and before him Sten Lindström) established important connections between modified forms of these postulates and the axioms of AGM contraction (Rott, 1993):

($\div$ 1) $K \div A = Cn(K \div A)$ [*closure*]
($\div$ 2) $K \div A \subseteq K$ [*inclusion*]
($\div$ 3) If $A \notin K$ or $A \in Cn(LK)$, then $K \subseteq K \div A$ [*vacuity*]
($\div$ 4) If $A \notin Cn(LK)$, then $A \notin K \div A$ [*success*]
($\div$ 5) If $Cn(A) = Cn(B)$, then $K \div A = K \div B$ [*extensionality*]
($\div$ 6) If $A \notin K \div (A \wedge B)$, then $K \div (A \wedge B) \subseteq K \div A$ [*conjunctive inclusion*]
($\div$ 7) If $A \notin Cn(\emptyset)$, then $K \div A \subseteq K \div (A \wedge B)$
($\div$ 8) If $A \notin Cn(\emptyset)$, then $Cn(K \div A) \cup \{A\}) = K$ [*recovery*]

For example, it is easy to verify that a notion of partial meet contraction defined via a selection function that satisfies (suitable reformulations) of Sen's α satisfies postulates ($\div$ 1) to ($\div$ 7) and that if β+ is satisfied then ($\div$ 1)- ($\div$ 6) and ($\div$ 8) are satisfied. Moreover a suitable converse of this postulate is also probable. Rott also offers an alternative proof of one of the central results in (Alchourrón et al., 1985), namely that, for logically finite theories, a contraction is a transitively relational partial meet contraction function if and only if it satisfies postulates ($\div$ 1) to ($\div$ 8). He also offers an interesting proof that for logically finite theories a contraction is a negatively transitively relational partial meet contraction function if and only if it satisfies postulates (÷ 1) to (÷ 7), (÷ 8r) and (÷ 8c), where:

(÷ 8r) $K \div (A \wedge B) \subseteq Cn(K \div A \cup K \div B)$
(÷ 8c) If $B \in K \div (A \wedge B)$, then $K \div (A \wedge B) \subseteq K \div A$

These two conditions arise naturally via the connection with the choice constraints γ and ε, although they have not been previously studied in the literature deriving from (Alchourrón et al., 1985).

Rott's model is tantamount to offering a justification in terms of a model of rational choice of the notion of contraction first studied by AGM. The model is nevertheless dependent on two main assumptions: (1) that the method of optimization should be used when facing the type of choices that arise in the resolution of a contraction and (2) that the space of feasible options from which one optimizes is indeed the one given by remainder sets.

2. Maximizing rather than optimizing

Both aforementioned assumptions have been questioned in recent research. We already mentioned that one could maximize rather than optimize. For example Arló-Costa studies in (Arló-Costa, 2006) the *liberal contraction* arising form the maximization of a quasi transitive relation. It is clear that in this case Sen's condition β+ is too strong. For consider the set S = {a, b} and the set S' = {a, b, c} such that c strictly dominates b (and a and b are incomparable as well as a and c). It is clear that this scenario offers an example of a quasi-transitive relation violating condition β+. Moreover it is possible to show (Arló-Costa, 2006, Th, 3.4) that a choice function on a space (X, **S**) where the domain S consists of all finite subsets of X, is quasi transitive μ-rational if and only if it satisfies α, the superset axiom and γ.

This indicates that the notion of contraction arising from the process of maximizing a quasi-transitive preference relation will satisfy axioms (÷ 1)- (÷ 7), (÷ 8r) and (÷ 8c) but not axiom (÷ 8). This constitutes a first departure from the AGM notion of contraction.

The motivation for this first departure is the study of agreements between different standards of value utilized in changing view. One might consider the case when there is indeterminacy in value and the agent facing a change of view has at his disposal various rankings according to a plurality of dimensions. One of them could order options according to simplicity considerations, while another might focus on coherence, for example. Even if one supposes that each of these rankings is complete the agreements across the different rakings (the set of pairs shared by all rankings, if any) need not be complete. We can only suppose that they are quasi-transitive. Then it makes sense to maximize the *categorical preference* (the term is Levi's) encoding the agreements among the different rankings. This is exactly the motivation behind the notion of liberal contraction.

But one might also study the more standard situation when there is no indeterminacy in value (which constitutes a particular case of the general case where there is indeterminacy). This is the case where the main motivation can be to maximize a preference relation of the standard type used in economics. The corresponding notion of preference is usually assumed to be asymmetric and negatively transitive (see (Kreps, 1988)). In this case it is well known that the relevant properties of a choice function are the property α mentioned above and the weaker property β (weaker here means weaker than the property β+):

(β) For all S, S' $\in$ **S** such that S $\subseteq$ S', x, y $\in$ μ(S) and y $\in$ μ(S'), then x $\in$ μ(S').

Amartya Sen explains in (Sen, 1997) that property β is entailed, but does not entail property β+. We can rewrite β in a way that is more amenable for a comparison with β+:

(β') If S $\subseteq$ S' and μ(S')$\cap$$\mu$(S) $\neq\emptyset$, then μ(S) $\subseteq$ μ(S').

It is easy to see that (β') and (β) are equivalent. First assume (β). To prove (β') assume that S $\subseteq$ S' and μ(S')$\cap$$\mu$(S) $\neq\emptyset$. Then there is y $\in$ μ(S')$\cap$$\mu$(S). Assume that x $\in$ μ(S). Since y $\in$ μ(S'), we have that, by (β), x $\in$ μ(S'). Now assume (β'). To prove (β) assume that S $\subseteq$ S', x, y $\in$ μ(S) as well as y $\in$ μ(S'). Therefore we have that μ(S')$\cap$$\mu$(S) $\neq\emptyset$, and since by assumption x $\in$ μ(S), we have that x $\in$ μ(S').

Since (β) is weaker than (β+) presumably we will not have ($\div$ 8) validated (at least the proof offered by Rott in (Rott, 1993) does not seem to work for (β')) although probably there is a weaker property of contraction[3] induced by (β). In any case, it seems that here as well we have a departure from the AGM principles.

[3] Weaker than ($\div$ 8) but apparently different from the already considered postulates ($\div$ 8r) and ($\div$ 8c).

There is a second assumption embedded in Rott's model of AGM contraction that has been recently questioned as well: that choices are made by taking into account the remainder set ($K\bot A$) as the domain of the choice function γ (or μ). Isaac Levi has questioned this second assumption in various recent publications (Levi, 1991, 2004). I shall focus on this second assumption and alternatives to it in the coming section.

1. Feasibility and cognitive choice

In all the previous sections we have assumed that in order to determine the content of a contraction $K \div A$ it is adequate to maximize over a feasible set determined by the contents of the remainder set ($K\bot A$). But the use of the remainder set as a feasible over which decisions are made in contraction is a controversial issue in the foundations of belief change. Isaac Levi has proposed in various writings that one should focus instead on the larger family of *saturatable* contractions removing A.

DEFINITION: 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 a maximal and consistent set.

So, the proposal in many of the previous articulations on the so-called 'Levi contractions' is to widen the scope of the choice function used in order to define contraction. The idea is that these choice functions take saturatable families as arguments. These choice functions should be such that when applied to a family S(K, A), return a non-empty susbset of S(K, A).

DEFINITION: $\div$ is a Levi-contraction of a theory K if and only if there exists a choice function G for K such that for all sentences A: if $A \in K$, then $K \div A = \cap G(S(K, A), \leq)$, where $\leq$ is a weak order, and if $A \notin K$, $K \div A = K$.[4]

A contraction operator of this kind does not obey the controversial postulate of Recovery presented above. Unlike other presentations of contraction this kind of contraction is responsive to the aforementioned criticisms (concerning feasibility), and it is usually complemented by the explicit introduction of a value function V on the set of logically closed sub-

[4] We follow here the presentation of Levi contractions given in (Hansson, 1999). A presentation of Levi contractions closer to the ideas presented by Levi in (Levi, 2004) will be introduced in the next section of this paper.

sets of a theory of reference K. The value function is supposed to obey at least minimal structural postulates like:

(Weak Monotony) For any two sets X, Y in the range of V, such that $X \subseteq Y$, $V(X) \leq V(Y)$.

A more robust notion of contraction can be then introduced by further constraining the choice function G in such a way that it optimizes the underlying value function;

$$G(S(K, A)) = \{X \in S(K, A): V(X) \leq V(Y) \text{ for all } Y \in S(K, A)\}$$

An obvious option here would be to study the general properties of an operation $K \div A = \cap M(S(K, A), <)$ where 'M' stands for an operation of maximization rather than optimization and $<$ is a preference relation. The most salient cases would be when $<$ is a preference relation obeying asymmetry and negative transitivity or the alternative case when the relation is quasi-transitive.

It is clear that the combined result of maximizing over an extended feasible set would entail an even greater departure from the AGM standard. We know that Recovery will not be satisfied and we also know that even if the underlying preference relation is asymmetric and negatively transitive we will not have the full force of the eight postulates of AGM. Nevertheless it would be nice to have a complete study of this weaker operation of contraction. A preliminary study is offered in (Arló Costa, 2006).

In the coming section we consider a concrete decision theoretic account of how to maximize over saturatable contractions which, in turn, are determined relative to a basic partition of events for a posited rational agent. The use of agent-relative partitions constitutes a third departure from the AGM orthodoxy and the appeal to decision theoretic distinctions permits a novel articulation of the notion of cognitive economy on contraction (as opposed to the simple appeal to a principle of Economy recommending minimization of information loss).

3. Cognitive Economy in Contraction

In a recent paper Hans Rott articulated what he thinks are two dogmas of belief revision (Rott, 2000). One of them is the so-called principle of Conservatism or Principle of Economy. The central idea of the principle is that *information loss* has to be minimized in contraction (the principle was introduced above under the name of 'informational economy').

Conservatism played, without doubt, a motivating role in the early stages of research in the AGM tradition. It led to some precise formulations of contraction, like the so-called maxichoice contraction and it paved the way towards the more sophisticated account of contraction defended in (Alchourrón et al., 1985): partial meet contraction.

Maxichoice contractions propose to achieve the contraction of a theory by a sentence A by selecting some maximal subset of K that does not imply A. This account is directly motivated by Conservatism, but it produces an unintuitive account of revision. In fact, if we denote the (maxichoice) contraction of K with $\neg$A, by K / $\neg$A, it seems reasonable to represent the revision of K with A as the logical closure of the set {(K/ $\neg$A) $\cup$ A}. But then all revisions of theories will be represented by maximal and consistent theories, an undesirable result.

AGM departed from the Principle of Conservatism by rejecting the recommendation of maxichoice contractions as mandatory in all cases. As we explained above, the central idea in (Alchourrón et al., 1985) was to: make a selection of the 'best' elements in the set of all maximal non-A-implying subsets of K; and then take the intersection of this selection. This is what is usually called a partial-meet contraction.

It is clear that partial-meet contractions do not follow Conservatism. As Rott and Pagnucco (Rott and Pagnucco, 1999, 503) have recently observed: "The Principle of Economy has been severely compromised in the AGM framework." In a more recent article (Rott, 2000), Rott called the principles of Economy and Entrenchment "dogmas" '... not because almost all researchers kept to these principles (quite the opposite is true) but because so many authoritative voices proclaimed them to be the philosophical or methodological rationale for their theories.'

As a matter of fact, the philosophical motivation for the AGM approach remains unclear. Much of the appeal of the early work by AGM focuses on their use of the axiomatic method and the use of logical techniques to fully represent different axiomatic accounts of contraction. But their work is not grounded on the use of a central epistemological principle. As Rott explains in the previous quotation, Conservatism was invoked as the central principle used in the development of partial meet contraction but nevertheless the principle was not used while defining the notion of contraction in question.

Isaac Levi proposed in a series of articles and books ((Levi, 1991, 2004) and (Arló-Costa and Levi, 2006)) a different principle that we can call the principle of Cognitive Economy:

The principle of Cognitive Economy: Keep loss of informational value to a minimum in contraction.

The crucial added notion to the Principle of Economy is the notion of *information value*. A minimal constraint constitutive of the notion of information value is the following:

(*Weak Monotonicity*) For any two feasible sets X, Y such that $X \subset Y$, $V(X) \leq V(Y)$.

The idea here is that even in the case where a feasible set X carries strictly less information than another feasible set Y, the information value of both sets might be equal. The reason for this is that the extra amount of information in the set Y might not be valuable.

Let L be a classical propositional language containing the classical connectives. The underlying logic will be identified with its Tarskian consequence operator Cn: $2^L \rightarrow 2^L$. We also assume that Cn obeys the deduction theorem and is compact. A *theory* is any set K such that K = Cn(K). Theories can be used advantageously in order to represent the epistemic commitments of rational agents.

According to Levi, in giving an account of belief change it is desirable to focus on changes in theories *relevant* to a given problem or question or cluster of questions. The idea here is that a problem or question in inquiry typically presupposes substantial claims that are intended to be taken for granted throughout the changes in belief that take place. Levi proposes to gather these assumptions in a minimal theory LK included in the current view K.

The potential answers to the problem under consideration are then arranged into a basic partition B where each cell in this partition is an expansion of the theory LK. A necessary constraint on the admissibility of B is that should be formed by expanding LK with sentences that are relevant answers to questions under investigation and that the expansions are restricted to expansions by adding to LK elements of a set of sentences such that LK entails that exactly one of them is true and each element of the set is consistent with LK.[5] The sub-partition of B constituted by the partition cells whose intersection is K will be called U and its complement D. For the most Levi part restricts his discussion to the finite case (B is finite). I shall adopt the same constraint here and I shall

[5] As explained in (Levi, 2004, p. 47) the expansions in question can be obtained via a two-step process. First one can consider the set of maximal and consistent expansions of LK. Then we can partition this set into finitely many cells $c_1, .., c_n$ and we can take the intersection of the maximal and consistent sets in each of these cells. The corresponding set of theories constitutes the cells of the basic partition B.

assume as well that all partition cells are finitely axiomatizable. Of course B, U and D are all relative to K and to LK. I shall omit sub-indexes here for the sake of readability.

Each cell in the basic partition can be represented as the intersection of a family of maximal and consistent sets of the initial language L. I shall adopt the following notation. |A|, for $A \in L$ denotes the set of maximal and consistent extensions M of L such that $A \in M$. For any theory K such that the theory can be represented as the intersection of a set of partition cells C_1,.. C_n, I shall use the notation [K] to denote $\{C_1, \ldots, C_n\}$. Also if A is the finite axiomatization of K, $[A] = \{C_1, \ldots, C_n\}$. In general $[A] = \{C_i \in B: A \in C_i\}$. $L \supseteq \mathbf{L} = \{A \in L: [A] \neq \emptyset \text{ and } [\neg A] \neq \emptyset\}$.

We can immediately define some useful notions. Every *potential contraction removing* $A \in \boldsymbol{L}$ from K is the intersection with K of a non-empty subset *R* of $\neg A$-entailing cells of D and a subset *R** of *A*-entailing cells of D that may or may not be empty. A *maxichoice contraction* of K relative to D is the intersection of K with a single element of D. A *maxichoice contraction of K removing* $A \in \boldsymbol{L}$ relative to D is the intersection of K with a single element of D that entails $\neg$A. A *saturatable contraction* of K *removing* $A \in \boldsymbol{L}$ relative to D is the intersection of a maxichoice contraction of K removing A relative to D with the intersection of a set of elements of B none of which entail A.

DEFINITION. 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 D.

$\Theta = \{X : X = \cap([K], Y) \text{ , with } Y \in 2^D\}$. Now we can introduce a *measure of informational value* V: $\Theta \rightarrow [0,1]$.[6] As the terminology indicates V is supposed to deliver a measure of the value of *information*. As such, Levi assumes 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(.). There are two basic properties that probability-based measures of information satisfy. Weak monotony can then be expressed with respect to the basic partition:

(*Weak Monotonicity*) For any two sets $X, Y \subseteq \Theta$ such that $X \subset Y$, $V(X) \leq V(Y)$.

[6] For most applications in this article I shall endorse the simplifying assumption that the value function takes values in the set of positive integers.

The second important postulate is the following one:

(*Extended Weak Monotonicity*) Let $X;Y \subseteq \Theta$. If $S \subseteq \Theta$ is incompatible with both X and Y, and if $V(X) \leq V(Y)$, then $V(X \cap S) \leq V(Y \cap S)$.

Unfortunately one cannot preserve all the properties of Cont in characterizing a notion of information value useful in contraction. The trouble with Cont is that it cannot rationalize (in terms of optimality) moving to a position of suspense when there is a tie in optimality. In fact, the Cont-value of the intersection of two optimal saturatable contractions need not and, in general, will not carry maximum Cont-value. So Levi proposes to preserve the first two postulates while adding a third that permits rationalizing suspense among optimal options as optimal. In order to present this third postulate we need an additional piece of notation. Notice first that any saturatable contraction in S(K;A) has the canonical form $K \cap T_A \cap m_{\neg A}$, where T_A is an intersection of A-cells in D and where $m_{\neg A}$ is a $\neg$A-cell in D. Then we can say that two saturatable contractions removing A from K are *A-equivalent* if and only if they are constituted as intersections of K with different $\neg$A-cells in D and the same subset T_A of the subset all of whose members entail A. A saturatable contraction S removing A is A-equivalent to an intersection of a set T of saturatable contractions removing A (including S) if S is constituted as the intersection of K, a set T_A of A-entailing cells and a $\neg$A-cell in D, and $[\cap T \cap A] = T_A$.

(*Weak Intersection Equality*) For every subset *T* of potential contractions removing A from K each element of which is of equal informational value and such that all the elements in *T* are A-equivalent and A-equivalent to their intersection, then for every $X \in T$, $V(\cap T) = V(X)$.[7]

These three *core* postulates jointly imply (Arló-Costa and Levi, 2006):

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

[7] This is, I understand, the most recent version of this postulate held by Levi (personal communication). Previous versions did not appeal to constraints in terms of A-equivalence. In particular it is useful to see that the following postulate *does not* follow from the one stated above:

For every subset *T* of S(K,A) each element of which is of equal informational val-

Some of the core conditions (WM especially) induce important constraints on contraction.

DEFINITION. $\div$ 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 $\gamma(S(K;A)) = \{X \in S(K;A): V(Y) \leq V(X) \text{ for all } Y \in S(K;A)\}$ and (ii) $K \div A = Cn(K)$ otherwise.

Observation 1 (Hansson and Olsson, 1995): A *basic* operator of informational value (obeying WM) obeys the following conditions ($\div$ 1) to ($\div$ 6):

($\div$ 1) $K \div A = Cn(K \div A)$ [*closure*]
($\div$ 2) $K \div A \subseteq K$ [*inclusion*]
($\div$ 3) If $A \notin K$ or $A \in Cn(LK)$, then $K \subseteq K \div A$ [*vacuity*]
($\div$ 4) If $A \notin Cn(LK)$, then $A \notin K \div A$ [*success*]
($\div$ 5) If $Cn(A) = Cn(B)$, then $K \div A = K \div B$ [*extensionality*]
($\div$ 6) If $A \notin K \div (A \wedge B)$, then $K \div (A \wedge B) \subseteq K \div A$ [*conjunctive inclusion*]

It is important to see that the core postulates *do not* validate some stronger syntactic postulates of contraction, which have been discussed at length in the literature. I list here two of these postulates:

($\div$ 7) If $A \notin Cn(\emptyset)$, then $K \div A \subseteq K \div (A \wedge B)$ [*antitony*]
($\div$ 8) If $A \notin Cn(\emptyset)$, then $Cn(K \div A) \cup \{A\}) = K$ [*recovery*]

Recovery is part of the AGM theory of contraction but Levi has argued at length against its tenability (Levi, 1991). So, he has proposed in (Levi, 2004) that a condition of adequacy of stronger theories of informational value (obtained by imposing further constraints on V aside from the core postulates) is that they should not lead to the validation of recovery. Of course this constraint is rather weak. There is a relatively large spectrum of permissible stronger theories satisfying this constraint. In (Arló-Costa and Levi, 2004) an argument is presented for selecting exactly one theory among the permissible ones. There is a strengthening of the core postulates leading to a theory of contraction, which can also be independently rationalized in terms of a direct articulation of the notion of entrenchment and its role in contraction. The idea is that when A is given up from a theory K one should preserve all sentences better entrenched than A. A theory of informational value compatible with this simple idea is the one obtained by requiring in addition to the core postulates:

(*Strong Intersection Equality*) If T is a set of maxichoice contractions from K 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 the core postulates entails the following:

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

In (Arló-Costa and Levi, 2006) it is shown that the resulting operator of informational value can be completely characterized syntactically in terms of the postulates ($\div$ 1)- ($\div$ 6), plus the postulate of Antitony presented above. Levi calls the resulting notion *mild contraction* while Rott and Pagnucco call it *severe withdrawal*. Antitony is less known than recovery, but so far it has produced a similar amount of controversy than recovery. Some scholars strongly oppose it. For example, Hansson has argued (Hansson, 1999) that 'Antitony does not hold for any sensible notion of contraction'.

It is interesting that the foundational strategy followed in (Rott and Pagnucco, 1999) is quite different from the one outlined above. Nevertheless both accounts converge on the same set of postulates. And the set of postulates in question diverges in various important ways from the standard theory of contraction defended by AGM. First, as in the case of the theory mentioned above, where contraction is studied in the context of the general theory of rational choice and where the feasible set is extended to a saturatable set, the theory that thus arises does not satisfy recovery. Second, the new theory satisfies the postulate of Antitony, which is not part of the AGM theory.

There is, nevertheless, a clear sense in which the theory of mild contractions does not depart substantially from the general framework for belief change presented by (Alchourrón et al., 1985). In fact, the theory of mild contractions is revision-equivalent to AGM contraction in the sense that the revision operator obtained via the so-called Levi-identity from an operator of mild contraction is an AGM revision operator:

Levi Identity: $K*A = (K\div \neg A)+A$

The operator '*' stands for an AGM operator of contraction. Then we have that if $\div_m$ is an operator of mild contraction and $\div$ is a standard AGM operator:

Equivalence: $(K\div_m \neg A)+A = K*A = (K\div \neg A)+A$

There are, as a matter of fact, various contraction operators not coincident with AGM contraction that are revision equivalent to AGM contraction. We explore this issue briefly in the coming section.

3.1 Systems of shells of information value

A useful ranking-type representation for mild contraction can be obtained as follows in terms of systems of shells of information value. For simplicity we have here a value function with values over the positive integers.

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

Intuitively R^X groups the D-cells 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 $P \subseteq 2^D$: $\rho^+(P) = \max(y: R^y \cap P \neq \emptyset)$

This notion of rank will be useful below. 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 to the index m.

DEFINITION. The m-shell of informational value $S^m = \cup\{R^x: x \in I \text{ and } x \geq m\}$. S is a system of shells of informational value if $S = \{S^X: \cup S^X = D\}$.

It is obvious that shells are nested. So a system of shells for a function V determines (at least) a grading of the cells in D. For any cell $w \in D$ we do not necessarily have $V(w) = \rho^+(w)$. The value-level of a cell in D need not coincide with its rank.

With the help of the previous definitions we can now characterize our standard 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. Let $A \in \mathbf{L}$ be a sentence rejected in K. Then S_A is the union of [K] with the shell $X \in S$ such that $X \cap [A] \neq \emptyset$ and for any other shell $Y \in S$, such that $Y \cap [A] \neq \emptyset$, $X \subseteq Y$.

S_A just picks the union of [K] with the innermost shell in the S for V containing A-cells. Let's call standard any operator of informational value defined via the third definition where the underlying value function obeys the core postulates plus Min. Then we can characterize standard operators in terms of the operation S_A defined in terms of systems of shells:

Observation 2. $[K \div \neg A] = S_A$

Standard operators of informational value can be characterized purely in terms if ranks by defining an additional 'lower' rank:

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

Lower ranks have some properties obeyed by ranking operations in systems like Spohn's. For example: $\rho^{-}(P \cup Q) = \min(\rho^{-}(P), \rho^{-}(Q))$. Now we have the following corollary:

Corollary 1. $[K \div \neg A] = \cup\{P \in 2^{D}: \rho^{-}(P) = \rho^{+}(A)\} \cup [K]$

Ranking systems as defined by Spohn and other scholars should be distinguished from shell systems. A detailed comparison stressing formal differences is presented in (Arló-Costa and Levi, 2004). With the help of the elements just introduced it is easy to show that Antitony holds when the value function is constrained by the core postulates and Min:

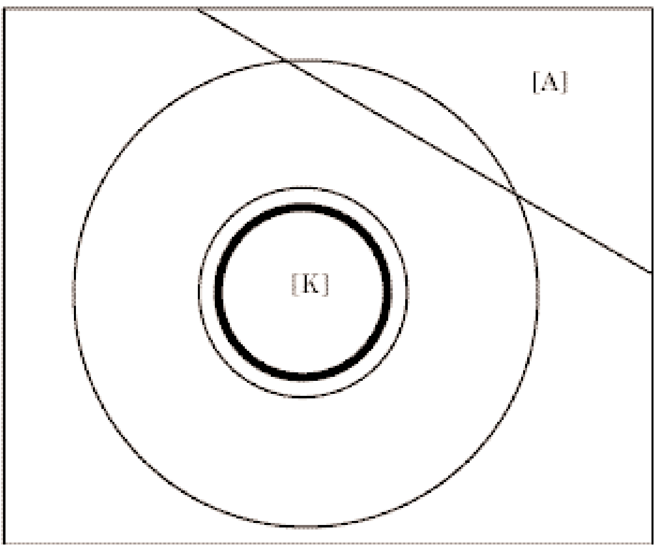

Fig. 1

So, if A is a sentence rejected in K, the outermost circle represented in the picture indicates the innermost shell of information value S_A intersecting [A]. It should be clear that this contraction is uniquely selected if one insists on minimizing loses of information value in contraction. Although the diagram might look familiar to readers familiar with Grove systems of spheres, the construction is here rather different. Grove systems do not have an index of value and under that point of view selecting S_A might be seen as an unnecessary loss of information.

The traditional account of contraction proposed by AGM would propose to form a contraction by intersecting the intersection of S_A and [A] with [K]. This is based on the idea of minimizing pure information loss. But when one focuses on information value the situation changes: the information value carried by $\cap S_A$ coincides with the information value carried by the aforementioned 'partial meet' contraction. Moreover any 'withdrawal' obtained by intersecting the intersection of S_A and [A] not only with [K] but also with other elements of S_A, will carry the same information value than the one carried by $\cap S_A$. All this follows immediately from our definitions of shells of information value and by the utilization of a notion of information value obeying the Min rule.

In (Arló-Costa, 2006a) I considered various operators of information value obeying the core postulates presented above but not necessarily the Min postulate. Then a series of contraction operations arise parametrically. They basically correspond to the many withdrawal operators considered in the literature (see (Rott and Pagnucco, 1999)). All these operators continue to be revision-equivalent with AGM contraction.

Hans Rott and Maurice Pagnucco have provided arguments in defense of what they call severe withdrawals (our mild contractions) without appealing at all to decision theoretic arguments. This convergence is a prima facie reason to think that the postulates of mild contraction offer a solid foundational basis for contraction. Many scholars have argued nevertheless that the multiplicity of notions of contraction that are revision-equivalent (including, of course, salient notions like mild contractions or 'severe withdrawals') speaks against the usual foundational role of contraction as the central basic operation of belief change (together with standard expansion). The argument mentions the apparent fragility and fragmentation of the notion of contraction in opposition to the apparent solidity and invariance of the notion of revision. The target of the argument is the so-called Levi Identity and the associated idea that revision is not a basic operation of belief change, but a notion definable in terms of an underlying notion of contraction. The alternative idea is to take revision as the central operation of belief change. This would allegedly recognize the primary cognitive role of revision. Revisions, the argument goes, are common and intuitive, while contractions have a derivative theoretical role.

It is true that is difficult to find convincing examples of 'pure' contractions. Perhaps some of the most immediate examples come from suppositional reasoning. One might suppositionally contract a piece of information to check what are the consequences of such hypothetical operation. Alchourrón usually mentioned in conversation more interesting examples in jurisprudence, like the act of derogating a law. This seems to be a case of 'pure' contraction where one does not contract in order to give a hearing to a piece of information incompatible with the current view. The sole role of the operation is to expunge some information from a code of law. Notice that not all derogations of laws are motivated by the need of making possible the promulgation of new laws, incompatible with the current code.[8] A simple derogation, that imposes no obligation contrary to that of the existing law, is an example of such type of derogation. To be sure, derogations require some motivation, which could be provided by the fact that new information undermines the reasons for having a particular law. But this information might be perfectly compatible (logically) with the current view. So, there are examples of 'pure' contractions and therefore there is a need for understanding the notion of contraction independently of the associated notion of revision.

4. Voluntarism and rationality

The accounts of contraction just summarized present a view of belief change where an agent *decides* what to believe next. Obviously this kind of account has some severe limitations. It does not apply, for example, to perceptual changes in view where there is no conscious rational election of what to believe next. Levi and others (see the chapters devoted to expan-

ue and for every $X \in T$, $V(\cap T) = V(X)$.

[8] To be sure there are examples in jurisprudence where the derogation of a law is parasitic of the promulgation of a new (eventually) incompatible law. For example the expression *Lex posterior derogat priori*, i.e. a subsequent law imports the abolition of a previous one, seems to indicate this type of change. But these Latin formulas not always indicate the need for a revision, strictly speaking. For example the expression: *Generi derogatur per speciem*, means that a particular law which is a derogation of a general one must always produce its derogatory effect, it being immaterial whether it was issued before the general law or after it. But here the more specific law need not contradict the more general law. Much of these formulas used in jurisprudence appeal to the use of more expressive languages than the propositional one in terms of which most contemporary belief revision theory is usually formulated. In the case of *Generi derogatur per speciem* one would need at least a first order language in order to reflect

sion in (Levi, 1991)) have analyzed change originated by perceptual input in terms of *routine expansions*, where agents expand the current view in accordance with a pre-compiled program. Deliberate contraction of the kind studied here has a role in cases of this sort only when inconsistency is injected into the view of an agent via routine expansions. In cases of this sort deliberate contraction can be used in order to restore consistency.

Deliberate contraction of the sort considered above deals with cases of belief change where it makes sense to make *doxastic decisions*. One of these cases could be scientific change. But even in this case the applicability of the theory we are presenting here has some limitations. If we focus on the notion of mild contraction presented in the previous section its main motivating idea is to describe the norms that guide belief change for agents who share the notion of value articulated by the core postulates and the Min rule. Agents who change view in a deliberate manner but that do not share this notion of value might violate the postulates of mild contraction without lapsing into irrationality.

The notion of liberal contraction is less constrained decision-theoretically. In this case no specific constraint is imposed on each of the different dimensions of value assessment that an agent might have aside from the fact that each of them yields a weak ordering of options.

4.1 Some open questions

The review presented above discussed two different techniques for constructing contractions according to the prescriptions of the theory of rational choice. Therefore they constitute two manners of articulating a form of doxastic voluntarism where one decides what to believe next.

The first manner of proceeding appeals to choice functions and techniques used by economists in the area of revealed preference. The second offers a more direct account where decision theory is used in order to clarify what is the index that is minimized while giving up information.

Even when Rott has defended the theory of mild contractions in (Rott and Pagnucco, 1999) (severe withdrawals according to the terminology used by Rott and Pagnucco) he has not given a decision theoretic rationale for it. Surprisingly Rott has used decision theoretical techniques for providing a justification of the traditional account of AGM contraction. A full decision theoretic justification of mild contractions is offered in (Levi, 2004) and (Arló-Costa and Levi, 2006), but this model does not appeal to choice functions. Levi has argued that an account of this sort (an account of mild contractions in terms of choice functions) cannot be provided (Levi, 2004, p. 161). It is nevertheless unclear whether this is the case. The use of choice functions appeals basically to a translation procedure where the main constraints on contraction are derived form basic constraints on choice functions. The account presented by Rott in (Rott, 1993) derives the constraints on belief change from well-known constraints on choice functions like α, β+, or γ. The account presented in (Arló-Costa, 2006b) also utilizes well-known constraints on selection functions that have an independent meaning in economic theory.

So, one of the ideas behind the assertion that a theory of mild contractions cannot be justified in terms of a theory of choice functions might be that some of the principles of the theory of mild contraction might not have known counterparts in the theory of social choice. As we have seen the only new postulate in the theory of mild contractions aside from the usual AGM ones (minus Recovery) is the postulate of Antitony. The constraint on choice functions corresponding to Antitony is easy to state (when selections are performed over remainder sets):

(A) If $A \notin Cn(\emptyset)$, $\gamma(K \perp (A \wedge B)) \subseteq \gamma(K \perp A)$

Nevertheless, there does not seem to be an established principle constraining choice functions (in the theory of social choice) and corresponding to this condition. More importantly we argued above that remainder sets are not the adequate feasible set for a choice function. But this is easy to fix by making selections over saturatable sets. This has the immediate pay off of eliminating Recovery as one of the basic contraction postulates. This also puts additional constraints on the possible identification of principles of rational choice corresponding to the following modified version of A:

(A') If $A \notin Cn(\emptyset)$, $\gamma(S(K,(A \wedge B))) \subseteq \gamma(S(K, A))$

As we explained above, the postulates of mild contraction can be obtained from the standard AGM postulates by eliminating Recovery and adding Antitony. So, in order to obtain constraints on choice functions that validate the postulates of mild contractions it would be enough to add (A') above to conditions on choice functions operating over saturatable sets rather than remainder sets. Even if (A') lacks a well established counterpart on the theory of choice functions this situation is in a way expected. In the same manner that the translation procedures helped to individualize new principles constraining belief contraction (corresponding to well known principles used in the theory of social choice), it is expected that new postulates on choice functions can be identified by fixing a notion of contraction and focusing on the corresponding constraints it entails on choice functions. Nevertheless the determination of the exact constraints on choice functions needed to validate the theory of severe withdrawal (alias mild contraction) remains as an open problem. We only sketched here the general form of a theory that can be utilized to determine such constraints.

References

Alchourrón, C., Gärdenfors, P., and Makinson, D., "On the Logic of Theory Change: Partial Meet Contraction and Revision Functions", *Journal of Symbolic Logic* 50 (1985): 510–530.

Arló-Costa, H., "Decision-Theoretic Contraction and Sequential Change", in E. Olsson (ed.), *Knowledge and Inquiry: Essays on the Pragmatism of Isaac Levi*, Cambridge University Press, Cambridge (2006a).

——— "Rationality and Value: The Epistemological Role of Indeterminate and Agent-dependent Values", *Philosophical Studies*, 128 (2006b): 7-48.

Arló-Costa, H. and Levi, I., Contraction: "On the Decision-Theoretical Origins of Minimal Change and Entrenchment", *Synthese*, vol. 152, n 1 (2006): 129-154.

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

Hansson, S.O. and Olsson, E.J., "Levi Contractions and AGM Contractions: a Comparison", *Notre Dame Journal of Formal Logic* 36 (1995): 103-119.

Kreps, D.M. *Notes on the Theory of Choice*, Westview Underground Classics in Economics (1988).

Levi, I. *Mild Contraction: Evaluating loss of Information due to loss of belief*, Oxford University Press (2004).

——— *The fixation of belief and its undoing*, Cambridge: Cambridge University Press (1991).

Rott, H., *Change, Choice and Inference: A study of belief revision and nonmonotonic reasoning*, Oxford Science Publications (2001).

——— "Two Dogmas of Belief Revision", *The Journal of Philosophy*, 97 (2000): 503-22.

——— "Belief Contraction in the context of the general theory of rational choice", *Journal of Symbolic Logic*, 58 (1993): 1426-1450.

Rott, H. and Pagnucco, M., "Severe Withdrawal (and Recovery)", *Journal of Philosophical Logic 28* (1999): 501-547.

Sen, A., *Choice, Welfare and Measurement*, Harvard University Press; New Edition (1997).

——— "Maximization and the act of choice", *Econometrica*, 65 (1997a): 745-779.