# Generalized Change and the Meaning of Rationality Postulates

Renata Wassermann (renata@ime.usp.br)
*Department of Computer Science*
*Institute of Mathematics and Statistics*
*University of São Paulo, Brazil*

**Abstract.** The standard theory of belief revision was developed to describe how a rational agent should change his beliefs in the presence of new information. Many interesting tools were created, but the concept of rationality was usually assumed to be related to classical logics.

In this paper, we explore the fact that the logical tools used can be extended to other sorts of logics, as proved in (Hansson and Wassermann, 1999), to describe models that are closer to the rationality of a real agent.

## 1. Introduction

The theory of belief revision deals with the problem of how to accommodate new information received by a rational agent. The new piece of information may be inconsistent with what the agent previously believed and he may have to give up some previous beliefs in order to accept the new one.

The standard model for belief revision became known as *AGM* due to the initials of the authors of the seminal paper (Alchourrón et al., 1985). In the AGM theory, belief states are represented by beliefs sets, i.e., logically closed sets of sentences. In formal terms, a set $K$ of sentences is a belief set if and only if $K = Cn(K)$, where $Cn$ is the (supraclassical) consequence operator that represents the logic.

The original AGM framework is a theory about how highly idealized rational agents should revise their beliefs when receiving new information. The agents are idealized in that they have unlimited memory and ability of inference. They are able to immediately close a set under deductive inference and to keep a belief set that is usually infinite. Beliefs are represented by formulas of a propositional language and the logic underlying the agents' reasoning is supraclassical.

A belief base is a set not necessarily closed under logical consequence (Fuhrmann, 1991; Hansson, 1989; Nebel, 1992). For every belief base $B$, its closure $Cn(B)$ is a belief set that represents the beliefs held by the agent. The elements of $B$ are assumed to be in a sense more basic beliefs, from which the elements of $Cn(B) \setminus B$ are derived. Belief bases are usually assumed to be finite, which makes them more attractive from the computational point of view. They are also more expressive than belief sets, allowing for the distinction between what is usually called explicit and merely derived beliefs. Explicit beliefs are the most basic ones, with independent standing, while beliefs which are merely derived from them can be automatically given up when the basic beliefs are retracted. On the other hand, belief sets have important advantages in terms of simplicity.

In AGM theory (Alchourrón et al., 1985), three forms of belief change are identified: contraction, expansion, and revision. Contraction consists of retracting a specified sentence from the belief set. Expansion consists of adding a specified sentence set-theoretically to the belief set. If the old and the new information are not logically compatible, then the new belief state after expansion will be inconsistent. Revision is consistency-preserving incorporation of new information, i.e. if the input sentence is consistent, then the new belief set will be consistent. If necessary, consistency is obtained by deleting parts of the original belief set. These operations have also been defined for belief bases (Fuhrmann, 1991; Hansson, 1991; Hansson, 1999).

Of the three AGM operations, only expansion is characterized in a unique way. The expansion of a belief set $K$ by $\alpha$ is given by $K + \alpha = Cn(K \cup \{\alpha\})$ and the expansion of a belief base $B$ by $\alpha$ is given simply by $B + \alpha = B \cup \{\alpha\}$. Contraction and revision operations are not directly defined, but constrained by a set of rationality postulates. These postulates express the properties that reasonable operations of contraction or revision should satisfy. There are many different ways in which the operations can be constructed. Usually, it is assumed that revisions ($*$) are defined in terms of contraction (-) and expansion ($+$) by means of the *Levi identity*: $K * \alpha = (K - \neg\alpha) + \alpha$. In this paper, we will concentrate on the most well-known construction, known as *partial meet* contraction.

The underlying logic used for belief revision is usually assumed to be at least Tarskian[1] and supraclassical. Assuming that such a logic is used to model an agent's reasoning amounts to assigning a great deal of logical power to the agent. We are mainly interested in modeling less idealized agents.

In (Hansson and Wassermann, 1999), it was seen that the representation results linking constructions to postulates for operations of rational belief base change also hold for logics which are not Tarskian. In particular, logics that do not satisfy inclusion ($B \subseteq Cn(B)$) can be used.

[1] A Tarskian consequence operator is an operator $Cn$ that satisfies monotony ($A \subseteq B \Rightarrow Cn(A) \subseteq Cn(B)$), inclusion ($A \subseteq Cn(A)$) and idempotency or iteration ($Cn(Cn(A)) = Cn(A)$).

The fact that we can now use a larger class of logics enables us to model agents with different characteristics and still remain close to the AGM paradigm. We are interested in agents which may have:

- Limited memory.
- Limited logical ability.
- Inconsistency tolerance.

In this paper, we explore logics that address each of the three desiderata and see how they fit the generalized version of the representation results. We also discuss what the rationality postulates mean once we change the underlying logic. Finally, we discuss how the different logics can be combined in order to address the three properties above.

In the rest of this paper we consider $L$ to be a propositional language closed under the usual truth-functional connectives and containing a constant $\perp$ denoting falsum. We call *inference operation* any total function taking sets of formulas to sets of formulas, that is, any function from $\mathcal{P}(L)$ to $\mathcal{P}(L)$. We use $C$ to denote inference operators. We use $Cn$ to denote a Tarskian consequence operator. The Greek letters $\alpha, \beta, \gamma \ldots$ denote arbitrary formulas; $p, q, r \ldots$ denote propositional atoms; $A, B \ldots$ denote sets of formulas. For any formula $\alpha$, $Var(\alpha)$ is the set of propositional letters which occur in $\alpha$.

## 2. Generalized Belief Change

In (Hansson and Wassermann, 1999), the representation theorems existing in the literature for operations of belief base change were generalized. These results relate existing constructions to sets of rationality postulates, which give us what kind of behavior to expect from the operations. In the literature, these representation theorems had been proved for classical, or at least Tarskian logics.[2] In (Hansson and Wassermann, 1999), it was shown that they can be extended to a larger class of logics, limited by some very basic properties.

For example, the common construction of partial meet contraction has been axiomatized by Hansson in (Hansson, 1992). The construction is based on the idea of remainder set, the collection of maximal subsets of the belief base which fail to imply a given formula:

[2] One of the few works to consider weaker logics for AGM revision is (Restall and Slaney, 1995). We will deal with their system in Section 3.2. More recently, (Coniglio and Carnielli, 2001) have also addressed the fact that AGM revision can be "transferred" from classical logic to some non-classical logics.

DEFINITION 2.1. *(*Alchourrón and Makinson, 1982*) Let $X$ be a set of formulas and $\alpha$ a formula. The* **remainder set** *$X \perp \alpha$ of $X$ and $\alpha$ is defined as follows. For any set $Y$, $Y \in X \perp \alpha$ if and only if:*

- $Y \subseteq X$
- $Y \nvdash \alpha$
- *For all $Y'$ such that $Y \subset Y' \subseteq X$, $Y' \vdash \alpha$.*

DEFINITION 2.2. *(*Alchourrón et al., 1985*;* Hansson, 1991*) The* **partial meet base contraction operator** *on $B$ based on a selection function $\gamma$ is the operator $\dot{-}_\gamma$ such that for all sentences $\alpha$:*

$$B \dot{-}_\gamma \alpha = \begin{cases} \bigcap \gamma(B \perp \alpha) & \textit{if } B \perp \alpha \neq \emptyset \\ B & \textit{otherwise} \end{cases}$$

Hansson has given the following axiomatic characterization of partial meet base contraction:

THEOREM 2.3. *(*Hansson, 1992*) An operator $\dot{-}$ is an operator of partial meet base contraction on $B$ if and only if:*

- *If $\alpha \notin Cn(\emptyset)$, then $\alpha \notin Cn(B \dot{-} \alpha)$ (success)*
- *$B \dot{-} \alpha \subseteq B$ (inclusion)*
- *If $\beta \in B \setminus (B \dot{-} \alpha)$, then there is some $B'$ such that $B \dot{-} \alpha \subseteq B' \subseteq B$, $\alpha \notin Cn(B')$ and $\alpha \in Cn(B' \cup \{\beta\})$ (relevance)*
- *If for all subsets $B'$ of $B$, $\alpha \in Cn(B')$ if and only if $\beta \in Cn(B')$, then $B \dot{-} \alpha = B \dot{-} \beta$ (uniformity)*

In (Hansson and Wassermann, 1999), it was shown that as long as the logic considered satisfies monotony and compactness, the theorem continues to hold. The notion of remainder set was extended to collect the maximal subsets of the belief base that fail to imply a given formula *according to the inference operator $C$*. In this way, a generalized construction was obtained, based on a logic which is not necessarily Tarskian. The postulates were also generalized by replacing $Cn$ by some other inference operator $C$. The following representation result was obtained:

THEOREM 2.4. *(*Hansson and Wassermann, 1999*) Let $C$ satisfy monotony and compactness. Then $\dot{-}$ is an operator of partial meet contraction on $B$ based on $C$ if and only if for all sentences $\alpha$:*

- *If $\alpha \notin C(\emptyset)$, then $\alpha \notin C(B \dot{-} \alpha)$ (success)*

- $B \dot{-} \alpha \subseteq B$ *(inclusion)*
- *If* $\beta \in B \setminus (B \dot{-} \alpha)$, *then there is some* $B'$ *such that* $B \dot{-} \alpha \subseteq B' \subseteq B$, $\alpha \notin C(B')$ *and* $\alpha \in C(B' \cup \{\beta\})$ *(relevance)*
- *If for all subsets* $B'$ *of* $B$, $\alpha \in C(B')$ *if and only if* $\beta \in C(B')$, *then* $B \dot{-} \alpha = B \dot{-} \beta$ *(uniformity)*

Similar results were obtained for partial meet revision, but in this case, besides compactness and monotony, an extra property is needed:

**($\alpha$-local non-contravention)** if $\neg\alpha \in C(B \cup \{\alpha\})$, then $\neg\alpha \in C(B)$.

In the rest of the paper, we will explore the use of different logics to instantiate the generalized change paradigm.

## 3. Rational Agents

If we think of models of rational agents, using classical logic as the tool for reasoning implies the assumption that agents are perfect reasoners, in the sense that they are logically omniscient.

When we try to model more realistic agents, there are several characteristics of these agents that must be taken into account. For example, most agents have limited memory recall, they cannot remember everything they know at once. Also, most agents would not give up all the information they had only because they happen to hold inconsistent beliefs related to one particular topic. Another important limitation of realistic agents is the cost of computation. Some beliefs may follow classically from the agent's beliefs, but the agent may lack the time or logical ability to derive them.

Using the generalized change paradigm, we can try different logics to model the agents' reasoning and, as long as these logics satisfy the properties seen in Section 2, we can keep the representation results.

It is interesting to note that even if the postulates have the same form, each logic gives rise to a different meaning of the postulates. We will discuss this in the rest of this section.

### 3.1. Limited Memory

The original motivation behind the generalized form of the representation theorems was to design belief change operations which considered only the relevant part of a belief base. A belief base may be too large to be considered as a whole, but it may contain several pieces of information that are completely irrelevant for the change operation considered. As an example, suppose that an agent observes that, contrary to what he believed, it is raining. He will have to revise his beliefs (giving up the belief that it is not raining), but for this revision he does not have to take into account his beliefs about other subjects, such as Mathematics.

In (Hansson and Wassermann, 1999) and (Wassermann, 2001), two different approaches for isolating the relevant part of a belief base were introduced. We will briefly describe both and discuss their use.

In (Hansson and Wassermann, 1999), given two sets of formulas $A$ and $B$, the relevant formulas in $B$ for $A$ are defined to be those that contribute to proving or disproving any formula of $A$. This notion (denoted by $c(A, B)$) is formalized using the notion of *kernel sets* - minimal subsets implying a given sentence:

DEFINITION 3.1. (Hansson, 1994) *Let $C$ be an inference operation on $L$. The kernel operation $\perp\!\!\!\perp_C$ is the operation such that for every set $B$ and every formula $\alpha$, $X$ is an element of $B \perp\!\!\!\perp_C \alpha$ if and only if:*

1. $X \subseteq B$
2. $\alpha \in C(X)$
3. *for all* $Y$, *if* $Y \subset X$ *then* $\alpha \notin C(Y)$

*The elements of $B \perp\!\!\!\perp_C \alpha$ are called $\alpha$-kernels.*

We write $\perp\!\!\!\perp$ for $\perp\!\!\!\perp_{Cn}$, where $Cn$ is a classical consequence operation.

DEFINITION 3.2. (Hansson and Wassermann, 1999) *The $A$-compartment of $B$, where $A$ and $B$ are sets of sentences, is defined as:*
$c(A, B) = \bigcup_{\alpha \in A}(\bigcup((B \perp\!\!\!\perp \alpha) \cup (B \perp\!\!\!\perp \neg\alpha) \setminus (B \perp\!\!\!\perp \bot))$
*We call $c$ the compartmentalization function for the sets $A$ and $B$.*

The compartments defined above are usually overlapping, and should not be seen as a partition of the belief base into different topics or subjects.[3] The *localization* of a given inference operation $C$ to $A$ is given by:

[3] The definition of compartment around a set as the union of the compartments around each of its elements has some limitations. It may be the case that for some formulas $\alpha$, $\beta$ and $\gamma$, $\alpha$ is not relevant for $\beta$ or for $\gamma$, but is relevant for the set $\{\beta, \gamma\}$. An example due to Dubois and cited in (Herzig, 1997): let $\alpha$ be "I take a bath", $\beta$ be "I use a hair-dryer" and $\gamma$ be "I die".

DEFINITION 3.3. *(*Hansson and Wassermann, 1999*) Let $c$ be as in Definition 3.2. Then the $A$-localization of $Cn$ is the inference operation $C_A$ such that: $C_A(B) = Cn(c(A, B))$.*

*A set $B$ is $A$-locally consistent if and only if $\bot \notin C_A(B)$.*

The construction of the compartments required for the local change inference operations is computationally expensive, involving several checks for entailment. Local change operations are computationally as costly as the original versions, although intuitively more appropriate. However, the representation results proven in (Hansson and Wassermann, 1999) do not rely on the way the compartments are defined, but only on properties of the local inference operation obtained, namely, monotony, compactness, and $\alpha$-local non-contravention.

In (Wassermann, 2001), we have shown how to add structure to a belief base and use this structure in order to find the subset of the belief base which is relevant for a certain belief change operation. Using this structure, we have defined local inference operators that can be implemented efficiently. The key idea of the method described is to use a relation of relatedness between formulas of the belief base.

Given a relatedness relation $\mathcal{R}$, we can represent a belief base as a (possibly disconnected) graph where each node is a formula and there is an edge between $\varphi$ and $\psi$ if and only if $\mathcal{R}(\varphi, \psi)$. This graph representation gives us immediately a notion of degree of relatedness: the shorter the path between two formulas of the base, the more related they are. Another notion made clear is that of connectedness: the connected components partition the graph into unrelated "topics" or "subjects". Sentences in the same connected component are somehow related, even if far apart. Formally:

DEFINITION 3.4. *(*Wassermann, 2001*) Let $B$ be a belief base and $\mathcal{R}$ be a relation between formulas. An* **$\mathcal{R}$-path** *between two formulas $\varphi$ and $\psi$ in a belief base $B$ is a sequence $P = (\varphi_0, \varphi_1, ..., \varphi_n)$ of formulas such that:*

1. *$\varphi_0 = \varphi$ and $\varphi_n = \psi$*
2. *$\{\varphi_1, ..., \varphi_{n-1}\} \subseteq B$*
3. *$\mathcal{R}(\varphi_i, \varphi_{i+1})$, $0 \leq i < n$.*

*If it is clear from the context to which relation we refer, we will talk simply about a path in $B$.*

*We represent the fact that $P$ is a path between $\varphi$ and $\psi$ by $\varphi \overset{P}{\leadsto} \psi$.*

*The* **length** *of a path $P = (\varphi_0, \varphi_1, ..., \varphi_n)$ is $l(P) = n$*

Note that the extremities of a path in $B$ are not necessarily elements of $B$. This is due to the fact that we may want to check whether some sentence outside the belief base is related to elements of the belief base.

DEFINITION 3.5. (Wassermann, 2001) *Let $B$ be a belief base and $\mathcal{R}$ a relation between formulas of the language. We say that two formulas $\varphi$ and $\psi$ are* **related in** *$B$ by $\mathcal{R}$ if and only if there is an $\mathcal{R}$-path $P$ in $B$ such that $\varphi \overset{P}{\rightsquigarrow} \psi$.*

Given two formulas $\varphi$ and $\psi$ and a belief base $B$, we can use the length of the shortest path between them in $B$ as the degree of unrelatedness of the formulas. If the formulas are not related in $B$, the degree of unrelatedness is set to infinity. Formulas with a shorter path between them in $B$ are more closely related in $B$.

DEFINITION 3.6. (Wassermann, 2001) *Let $B$ be a belief base, $\mathcal{R}$ a relation between formulas of the language and $\varphi$ and $\psi$ formulas. The* **degree of unrelatedness** *of $\varphi$ and $\psi$ in $B$ is given by:*

$$u(\varphi, \psi) = \begin{cases} 0 & \text{if } \varphi = \psi \text{ and } \varphi \in B \\ min\{l(P) | \varphi \overset{P}{\rightsquigarrow} \psi, P \text{ in } B\} & \text{if } \mathcal{R}(\varphi,\psi) \text{ in } B,\ \varphi \neq \psi \\ \infty & \text{otherwise} \end{cases}$$

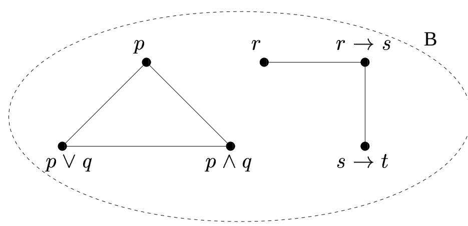


*Figure 1.* Structured belief base

In Figure 1 we see an example of a belief base structured by a relation $\mathcal{R}$ defined by:

$\mathcal{R}(\varphi, \psi)$ if and only if $\varphi$ and $\psi$ share (at least) one atom.

We now show, given the structure of a belief base, how to retrieve the set of formulas relevant for a given formula $\alpha$:

DEFINITION 3.7. (Wassermann, 2001) *The set of formulas of $B$ which are* **relevant for $\alpha$ with degree** *$i$ is given by:*

$\Delta^i(\alpha, B) = \{\varphi \in B | u(\alpha, \varphi) = i\}$ *for* $i \geq 0$

DEFINITION 3.8. *(*Wassermann, 2001*)*

*The set of formulas of $B$ which are* **relevant for $\alpha$ up to degree** *$n$ is given by:*

$\Delta^{\leq n}(\alpha, B) = \bigcup_{0 \leq i \leq n} \Delta^i(\alpha, B)$ *for* $n \geq 0$

*We say that* $\Delta^{<\omega}(\alpha, B) = \bigcup_{i \geq 0} \Delta^i(\alpha, B)$ *is the set of* **relevant formulas for $\alpha$**.

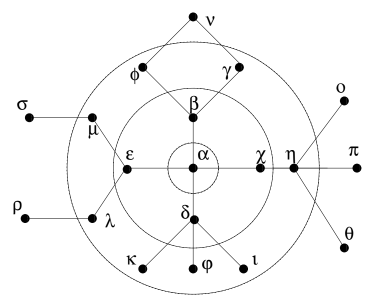


*Figure 2.* Degrees of relevance

In Figure 2, we see an example of a structured belief base $B = \{\alpha, \beta, \gamma, \delta, \varepsilon, \eta, \theta, \iota, \kappa, \lambda, \mu, \nu, o, \pi, \rho, \sigma, \phi, \varphi, \chi\}$. The circles represent different levels of relevance for $\alpha$. We have:

$\Delta^0(\alpha, B) = \{\alpha\}$
$\Delta^1(\alpha, B) = \{\beta, \chi, \delta, \varepsilon\}$
$\Delta^2(\alpha, B) = \{\gamma, \eta, \iota, \varphi, \kappa, \lambda, \mu, \phi\}$
$\Delta^3(\alpha, B) = \{\nu, o, \pi, \theta, \rho, \sigma\}$
$\Delta^{\leq\omega}(\alpha, B) = \Delta^0(\alpha, B) \cup \Delta^1(\alpha, B) \cup \Delta^2(\alpha, B) \cup \Delta^3(\alpha, B) = B$

PROPOSITION 3.9. *(*Wassermann, 2001*) Let $B$ be a finite belief base and $\alpha$ a formula. For every natural number $n$ and any inference operator $C$, if $C$ is monotonic and compact, then the local inference operations defined as $C^n_\alpha(B) = C(\Delta^{\leq n}(\alpha, B))$ are monotonic and compact. Moreover, if $C$ satisfies non-contravention, then $C^n_\alpha$ satisfies $\alpha$-local non-contravention (if $\neg\alpha \in C^n_\alpha(B \cup \{\alpha\})$, then $\neg\alpha \in C^n_\alpha(B)$).*

The proposition above implies that for any $n$, $\Delta^{\leq n}$ can be used to define local operations characterized by the results presented in (Hansson and Wassermann, 1999).

It is not difficult to see that we can depart from any relation $\mathcal{R}$ to obtain a local inference operator satisfying the relevant properties. For the proof of Proposition 3.9, no properties of $\mathcal{R}$ were used, we only needed conditions on the initial inference operator $C$. This gives us a very general framework. Depending on the application, one can use the most appropriate notion of relatedness and still obtain a local inference operator satisfying the properties needed for the representation results given in (Hansson and Wassermann, 1999).

Up to now, we have not said where the relatedness relation comes from. When no extra information about the domain is given, we can take the relatedness relation to be a syntactical relation between formulas, for example, we can say that two formulas are related if they share a propositional variable. In some applications, such a relation is given with the problem.

Both sorts of local inference, the one defined in (Hansson and Wassermann, 1999) ($C_A(B) = C(c(A, B))$) and the one defined in (Wassermann, 2001) ($C^n_\alpha(B) = Cn(\Delta^{\leq n}(\alpha, B))$) can be used with the generalized belief change framework giving rise to belief change operators that take only (the most) relevant beliefs into account. In this way, we can model the reasoning of an agent with limited memory recall, i.e., one that cannot think of all his beliefs at once.

What do the rationality postulates tell us about such an operation?

Of the four postulates for contraction, inclusion ($B \dot{-} \alpha \subseteq B$) is the only one which does not depend on the logic being used.

Consider the success postulate for contraction where the underlying logic is that of local inferences as defined in (Hansson and Wassermann, 1999):

$$\text{If } \alpha \notin C_A(\emptyset), \text{ then } \alpha \notin C_A(B \dot{-} \alpha)$$

In this case, $A$ may be any set of formulas. The set $c(A, B)$ represents the part of the belief base which the agent takes into account when reasoning about the formulas in $A$. And since the agent cannot consider all his beliefs at the same time, it may be the case that $\alpha$ is still a classical consequence of the contracted belief base. But the agent cannot see it, so he believes that the contraction was successful. Note that $\alpha$ and $A$ may be completely unrelated. If $\alpha$ is an element of the set $A$, then it is not difficult to see that local success ($\alpha \notin C_A(B \dot{-} \alpha)$) implies classical success ($\alpha \notin Cn(B \dot{-} \alpha)$) if $B \dot{-} \alpha$ is classically consistent.

Using the other notion of local inference, $C^n_\alpha(B) = C(\Delta^{\leq n}(\alpha, B))$, the agent first collects whatever he can see as being relevant for $\alpha$ and then performs the operation. Success in this case, means that if $\alpha$ is not a tautology, then it does not follow from the most relevant beliefs of the contracted base. Since the degree of relevance where the collection of the relevant beliefs stops is determined by the available resources, local success means in this case that the agent cannot find a proof for

$\alpha$ in the contracted belief base, he does not have enough memory for that.

### 3.2. Inconsistency Tolerance

One of the problems with the use of classical logic to model the behavior of a rational agent is that it trivializes in the presence of inconsistencies, i.e., whenever an agent happens to have inconsistent beliefs, his belief set contains every formula of the language. This is a highly undesirable property, since any real agent (natural or artificial) with a reasonable amount of beliefs may happen to accept inconsistent beliefs without even noticing it. Or even if he notices, he may not be able to solve the inconsistency.

We would like to model agents that tolerate some inconsistency, i.e., agents that can isolate the inconsistencies instead of letting them "contaminate" the whole belief base. A logic is said to be *paraconsistent* if for some pair of formulas $\alpha$ and $\beta$, we have $\alpha, \neg\alpha \not\vdash \beta$. Several logics with this characteristic have been proposed in the literature. We will concentrate here on the logic $C_1$ (da Costa, 1963).

We will first briefly introduce the calculus $C_1$, and then see how it fits the generalized belief change framework.

DEFINITION 3.10. A formula $\alpha$ is said to be *well behaved* if the principle of non-contradiction holds for $\alpha$, i.e., if $\neg(\alpha \wedge \neg\alpha)$ holds. We use $\alpha^o$ to denote $\neg(\alpha \wedge \neg\alpha)$.

Da Costa's calculus $C_1$ (da Costa, 1963) was introduced in order to deal with possible inconsistencies that should not damage reasoning by trivializing it. The idea is to block derivations from formulas which are not well behaved, isolating inconsistencies.

The following is an axiomatization of $C_1$:

($\rightarrow_1$) $\alpha \rightarrow (\beta \rightarrow \alpha)$

($\rightarrow_2$) $(\alpha \rightarrow \beta) \rightarrow (\alpha \rightarrow (\beta \rightarrow \gamma)) \rightarrow (\alpha \rightarrow \gamma)$

($\rightarrow_3$) $\alpha, \alpha \rightarrow \beta / \beta$

($\wedge_1$) $\alpha \wedge \beta \rightarrow \alpha$

($\wedge_2$) $\alpha \wedge \beta \rightarrow \beta$

($\wedge_3$) $\alpha \rightarrow (\beta \rightarrow \alpha \wedge \beta)$

($\vee_1$) $\alpha \rightarrow \alpha \vee \beta$

($\vee_2$) $\beta \rightarrow \alpha \vee \beta$

($\vee_3$) $(\alpha \rightarrow \gamma) \rightarrow ((\beta \rightarrow \gamma) \rightarrow (\alpha \vee \beta \rightarrow \gamma))$

$(\neg_1)$ $\beta^o \to (\alpha \to \beta) \to ((\alpha \to \neg\beta) \to \neg\alpha)$
$(\neg_2)$ $\alpha^o \wedge \beta^o \to ((\alpha \to \beta)^o \wedge (\alpha \wedge \beta)^o \wedge (\alpha \vee \beta)^o)$
$(\neg_3)$ $\alpha \vee \neg\alpha$
$(\neg_4)$ $\neg\neg\alpha \to \alpha$

This calculus is interesting since, except for negation, all other connectives behave classically. A semantic for this system was given in (da Costa and Alves, 1977):

DEFINITION 3.11. An *$\mathcal{N}$-valuation* $v_{\mathcal{N}}$ is a function, $v_{\mathcal{N}} : \mathcal{L}_C \to \{0, 1\}$, that extends a propositional valuation $v_p$ (i.e., $v_{\mathcal{N}}(p) = v_p(p)$), satisfying the following restrictions:

| | | |
|---|---|---|
| $(i)$ | $v_{\mathcal{N}}(\alpha \wedge \beta) = 1$ | $\Leftrightarrow v_{\mathcal{N}}(\alpha) = v_{\mathcal{N}}(\beta) = 1$ |
| $(ii)$ | $v_{\mathcal{N}}(\alpha \vee \beta) = 0$ | $\Leftrightarrow v_{\mathcal{N}}(\alpha) = v_{\mathcal{N}}(\beta) = 0$ |
| $(iii)$ | $v_{\mathcal{N}}(\alpha \to \beta) = 0$ | $\Leftrightarrow v_{\mathcal{N}}(\alpha) = 1$ and $v_{\mathcal{N}}(\beta) = 0$ |
| $(iv)$ | $v_{\mathcal{N}}(\alpha) = 0$ | $\Rightarrow v_{\mathcal{N}}(\neg\alpha) = 1$ |
| $(v)$ | $v_{\mathcal{N}}(\neg\neg\alpha) = 1$ | $\Rightarrow v_{\mathcal{N}}(\alpha) = 1$ |
| $(vi)$ | $v_{\mathcal{N}}(\beta^o) = v_{\mathcal{N}}(\alpha \to \beta) = v_{\mathcal{N}}(\alpha \to \neg\beta) = 1$ | $\Rightarrow v_{\mathcal{N}}(\alpha) = 0$ |
| $(vii)$ | $v_{\mathcal{N}}(\alpha^o) = v_{\mathcal{N}}(\beta^o) = 1$ | $\Rightarrow v_{\mathcal{N}}((\alpha * \beta)^o) = 1, * \in \{\to, \wedge, \vee\}$ |

An $\mathcal{N}$-valuation has the following properties:

LEMMA 3.12. *(1)* $v_{\mathcal{N}}(\alpha) = 1 \Leftrightarrow v_{\mathcal{N}}(\neg\alpha \wedge \alpha^o) = 0$
*(2)* $v_{\mathcal{N}}(\alpha^o) = 0 \Leftrightarrow v_{\mathcal{N}}(\alpha) = v_{\mathcal{N}}(\neg\alpha)$

PROPOSITION 3.13. *Let* $C_1(B) = \{\alpha | B \vdash_{C_1} \alpha\}$. *Then* $C_1$ *satisfies monotony, compactness, and* $\alpha$*-local non-contravention.*

**Proof:** Monotony and compactness are well known properties of $C_1$ (the notion of deduction is the usual). For $\alpha$-local non-contravention, suppose that $\neg\alpha \notin C_1(B)$. Then there is an $\mathcal{N}$-valuation $v$ such that for any formula $\beta \in B$, $v(\beta) = 1$ and $v(\neg\alpha) = 0$. By restriction $(iv)$ of the definition of an $\mathcal{N}$-valuation, if $v(\neg\alpha) = 0$, then $v(\alpha) = 1$. Hence, $v$ is an $\mathcal{N}$-valuation such that for any formula $\beta \in B \cup \{\alpha\}$, $v(\beta) = 1$, and $v(\neg\alpha) = 0$. Thus, $\neg\alpha \notin C_1(B \cup \{\alpha\})$. □

This means that $C_1$ can be used with the generalized belief change framework in order to model an agent that tolerates some inconsistency.

The agent may have some inconsistent beliefs which he does not care about. For those formulas about which the agent does care about consistency, he believes $\alpha^o$. The success postulate in this case indicates that $B \dot{-} \alpha$ may be classically inconsistent, but it is consistent with respect to $\alpha$ and $\alpha$ cannot be inferred from it.

### 3.3. Limited Logical Ability

Another problem of the use of classical logic is that it is usually computationally intractable. If we want to model realistic agents, we must take into account limits in computational power and logical ability. We will present in this section the idea of approximate entailment as a response to these limitations.

In (Schaerf and Cadoli, 1995), Schaerf and Cadoli define two approximations of classical entailment: $\models^1_S$ which is complete but not sound, and $\models^3_S$ which is sound and incomplete. These approximations are carried out over a set of atoms $S \subseteq L$ which determines their closeness to classical entailment. In the trivial extreme of approximate entailment, i.e., when $S = L$, classical entailment is obtained. At the other extreme, when $S = \emptyset$, $\models^1_S$ holds for any two formulas (i.e., for all $\alpha$,$\beta$, we have $\alpha \models^1_S \beta$) and $\models^3_S$ corresponds to Levesque's logic for explicit beliefs (Levesque, 1984), which bears a connection to relevance logics such as those of Anderson and Belnap (Anderson and Belnap, 1975).

In an $S_1$ assignment, if $x \in S$, then $x, \neg x$ are given opposite truth values; if $x \notin S$, then $x, \neg x$ both get the value 0. In an $S_3$ assignment, if $x \in S$, then $x, \neg x$ get opposite truth values, while if $x \notin S$, $x, \neg x$ do not both get 0, but may both get 1. The set of formulas for which we are testing entailments is assumed to be in clausal form.[4]

Since $\models^3_S$ is sound but incomplete, it can be used to approximate $\models$, i.e., if for some $S$ we have that $B \models^3_S \alpha$, then $B \models \alpha$. On the other hand, $\models^1_S$ is unsound but complete, and can be used for approximating $\not\models$, i.e., if for some $S$ we have that $B \not\models^1_S \alpha$, then $B \not\models \alpha$.

LEMMA 3.14. *(*Schaerf and Cadoli, 1995*) Let simplify-1$(B, S)$ be the result of deleting all literals of $B$ which mention atoms outside $S$. $B$ is $S_1$-satisfiable if and only if simplify-1$(B, S)$ is classically satisfiable.*

[4] All results can be extended for negation normal form cf. (Cadoli and Schaerf, 1995), but further generalization implies missing the good complexity upper bounds, as shown in (Cadoli and Schaerf, 1996). Besides the increase in complexity, the standard translation of formulas into NNF makes use of De Morgan's law and does not preserve truth values under Schaerf and Cadoli's non-standard semantics. This has been solved in (Finger and Wassermann, 2001a; Finger and Wassermann, 2001b), where extensions of $S_3$ to full propositional logic were presented.

THEOREM 3.15. (Schaerf and Cadoli, 1995) *Let* $\alpha$ *be* $\alpha_S \vee \alpha_{\overline{S}}$, *where* $Var(\alpha_S) \subseteq S$ *and* $Var(\alpha_{\overline{S}}) \cap S = \emptyset$. *Then* $B \models_S^1 \alpha$ *iff* $B \cup \{\neg\alpha_S\}$ *is not* $S_1$ *satisfiable.*[5]

This means that, in order to test whether $B \not\models_S^1 \alpha$, for every literal of $\alpha$ of the form $p$, where $p \in S$, we add the clause $\neg p$ to $B$ and for every literal of $\alpha$ of the form $\neg p$, where $p \in S$, we add the clause $p$ to $B$. Let $B'$ be this expanded set of clauses. We now must check whether $B'$ is $S_1$ satisfiable. Using Lemma 3.14, we can reduce this problem to testing the classical satisfiability of a simplified set of clauses.

LEMMA 3.16. (Schaerf and Cadoli, 1995) *Let simplify-3*$(B, S)$ *be the result of deleting all clauses of* $B$ *which contain an atom outside* $S$. *Then* $B$ *is* $S_3$*-satisfiable if and only if simplify-3*$(B, S)$ *is classically satisfiable.*

THEOREM 3.17. (Schaerf and Cadoli, 1995) *Let* $L(\alpha) \subseteq S$. *Then* $B \models_S^3 \alpha$ *iff* $B \cup \{\neg\alpha\}$ *is not* $S_3$ *satisfiable.*

As in the case of $S_1$, Lemma 3.16 and Theorem 3.17 together provide a constructive method for testing $S_3$ entailment.

In (Chopra et al., 2000), it was shown that the notion of approximate entailment described above can be used for implementing generalized change efficiently. As we have noted above in Section 2, if an inference operation satisfies monotony, compactness, and $\alpha$-local non-contravention, it can be used together with the axiomatizations given in (Hansson and Wassermann, 1999). Approximate entailment satisfies these requirements:

DEFINITION 3.18. (Chopra et al., 2000)$C_S^i(B) = \{\alpha | B \models_S^i \alpha\}$, *where* $S$ *is a set of propositional variables and* $i \in \{1, 3\}$.

PROPOSITION 3.19. (Chopra et al., 2000) *Let* $S$ *be a fixed set of propositional variables. Then* $C_S^3$ *satisfies monotony, compactness, consistency, and* $\alpha$*-local non-contravention for every formula* $\alpha$.

PROPOSITION 3.20. (Chopra et al., 2000) *Let* $S$ *be a fixed set of propositional variables. Then* $C_S^1$ *satisfies monotony, compactness, consistency, and* $\alpha$*-local non-contravention if* $Var(\alpha) \subseteq S$.

As an example of how $C_S^1$ can fail to satisfy the conditions above when $Var(\alpha) \not\subseteq S$, take $\alpha = p$, $B = \emptyset$, and $p \notin S$. Then any $S_1$

[5] This can only be done because $\alpha_S$ behaves classically and we can compute its negation in clausal form (as a set of clauses).

interpretation assigns both $\alpha$ and $\neg\alpha$ the value 0. We have that $\neg\alpha \in C_S^1(B \cup \{\alpha\}) = Cn(\bot)$, but $\neg\alpha \notin C_S^1(B) = Cn(\emptyset)$.

The propositions above imply that both inference operations can be used to define operations of belief change. This allows us to combine the computational efficiency of Schaerf and Cadoli's method for approximate entailment with the logical characterization of belief change operations given in (Hansson and Wassermann, 1999). The idea is that these approximate inferences give us an approximation of the revised belief base and the larger the set $S$ grows, the closer we get to the classical definition. At any point in the revision process, we can stop to check our results and depending upon resource availability, we can choose to carry on the revision process further or stop (Chopra et al., 2000).

Note that in keeping with the approximated inference operation used above, we will obtain two kinds of contraction operations. Operations based on $S_1$ entailment will model more radical contraction operations (leading to a greater loss of beliefs) and those based on $S_3$ will model cautious operations (fewer beliefs than warranted are dropped). The choice of which contraction operation to use will be a context-sensitive one. At any stage of the approximated revision process, we know that contraction based on classical consequence lies between contraction based on approximated consequence relations, and that adding more propositional variables to the context $S$ provides a better approximation. If $S = L$, then $B \dot{-}_{C_S^i} \alpha = B \dot{-}_{Cn} \alpha$. Usually, the approximations will converge for a proper subset $S$ of $L$.

We now look once more at the success postulate. In the case of $S_1$ entailment, we have that $\alpha \notin C_S^1(B \dot{-} \alpha)$ implies that $\alpha \notin Cn(B \dot{-} \alpha)$, i.e., since the operation may remove more than what is needed but never less, classical success obtains. This is not the case with $S_3$ entailment. The fact that $\alpha \notin C_S^3(B \dot{-} \alpha)$ only means that as far as the agent can see, $\alpha$ does not follow from the contracted base. It may follow classically, but the agent, with his present resource limitations, will not derive it.

### 3.4. Combining them all

We have started by saying that if we want to model realistic agents, we have to deal with:

1. Limited memory.

2. Limited logical ability.

3. Inconsistency tolerance.

We have seen three different sorts of logics which address each of the three desiderata and that fitted well in the generalized change paradigm. But what we are actually looking for is a single formalism which addresses the three points simultaneously. Can we combine the three kinds of logics?

Let us first examine some extra properties of the logics proposed. In (Finger and Wassermann, 2001a), the relation between $S_3$ entailment and the calculus $C_1$ was studied. The extension of $S_3$ to full propositional logic given in (Finger and Wassermann, 2001b) was shown to be very close to $C_1$. And with respect to inconsistency tolerance, both present the same behavior. For some formulas (those not in $S$ for $S_3$, those not well-behaved in $C_1$), the agent may believe the formula and its negation without believing every formula of the language. This means that $S_3$ entailment alone satisfies two of our three desiderata.

If we look at the notions of local inferences, we see that they also tolerate some inconsistencies. Consider the following example, taken from (Hansson and Wassermann, 1999):

**Example:**

When at home I hear on the radio that my friend Carol has been murdered yesterday night and that there were no traces of doors or windows having been forced. I talked to her yesterday on the phone and she was at home with her flat-mates Ann and Bill. I know that no one else, except for Ann, Bill and Carol had the keys to their apartment. I conclude that Ann or Bill must have done it. But I have known Ann for quite some time and cannot believe that she would be able to murder anyone. I believe that she did not do it. For similar reasons, I believe that Bill did not do it. This is clearly inconsistent with my belief that one of them did it, but I am sure of one thing: I do not believe I am asleep!

Let $a$ stand for the proposition "Ann is the murderer", $b$ for "Bill is the murderer", and $s$ for "I am asleep". My belief base $B$ contains: $\{a \vee b, \neg a, \neg b, \neg s\}$. The belief base is classically inconsistent, but if one is only interested in reasoning about being asleep or not, both methods for local inference will retrieve only the subset $\{\neg s\}$, which is consistent. This means that if $\alpha = s$, neither $C_\alpha(B)$ nor $C_\alpha^n(B)$ will contain the whole language.

From this, we see that we only need to combine local inference and approximate entailment in order to satisfy the three desiderata. In (Chopra et al., 2000), some steps in this direction were taken. The idea there was to use approximate entailment together with some notion of relevance. Relevance was used in order to guide the approximations, i.e., in order to find the context set $S$.

We want now to mix the notions of local and approximate entailment even further. The idea is to first find the relevant compartment and then use approximate entailment on the relevant part. For finding the relevant compartment we use the local inference defined in (Wassermann, 2001), which is computationally more efficient.

DEFINITION 3.21. *Let* $C_\alpha^{S,n}(B) = C_S^3(\Delta^{\leq n}(\alpha, B))$.

Note that this notion of inference is parametrized by $S$ an $n$. We can play with the parameters according to the available resources.

PROPOSITION 3.22. *For* $S$, $n$, *and* $\alpha$ *fixed,* $C_\alpha^{S,n}$ *satisfies monotony, compactness and* $\alpha$*-local non-contravention.*

**Proof:** Monotony follows directly from the monotony of $S_3$ and of local inference: if $B \subseteq B'$, then $\Delta^{\leq n}(\alpha, B) \subseteq \Delta^{\leq n}(\alpha, B')$ and from the monotony of $S_3$, $C_S^3(\Delta^{\leq n}(\alpha, B)) \subseteq C_S^3(\Delta^{\leq n}(\alpha, B'))$. Compactness also follows from the compactness of the local inference, since $\Delta^{\leq n}$ always retrieves a finite set, all proofs are finite. For $\alpha$-local non-contravention, we only have to see that $\Delta^{\leq n}(\alpha, B \cup \{\alpha\}) = \Delta^{\leq n}(\alpha, B) \cup \{\alpha\}$. If $\neg\alpha \notin C_\alpha^{S,n}(B)$, then there is an $S-3$ valuation $v$ such that $v$ takes all formulas in $\Delta^{\leq n}(\alpha, B)$ into 1 and $v(\neg\alpha) = 0$. By the definition of $S-3$ valuations, we know that $v(\alpha) = 1$, i.e., $v$ takes all formulas of $\Delta^{\leq n}(\alpha, B) \cup \{\alpha\}$ into 1 and $v(\neg\alpha) = 0$. Hence, $\neg\alpha \notin C_\alpha^{S,n}(B \cup \{\alpha\})$. □

This proposition shows that we can indeed combine local inference and approximate entailment and obtain a family of logics parametrized by the resource limitations of the agent and still keep the logical results of the generalized change paradigm.

## 4. Conclusions and Related Work

In this paper, we have shown how different logics can be used together with the generalized belief change proposed in (Hansson and Wassermann, 1999). We have seen what each one brings in terms of modeling the behavior of more realistic agents.

We have also proposed a new combination of the existent logics which satisfies all of our three desiderata.

Restall and Slaney (Restall and Slaney, 1995) have done something similar for the revision of belief sets. They showed that the notion of consequence used can be weakened to that of first-degree entailment keeping the representation results. First-degree entailment is is known to satisfy properties of paraconsistent and relevant logics. For this, it can be seen as a formalism addressing two of our three desiderata. However, it presents no gains in terms of computational complexity. [6]

Marquis and Porquet (Marquis and Porquet, 2001) combine approximate reasoning with some form of inconsistency tolerance. They designed a family of inference operators that are based on $S_3$ entailment, but that instead of considering the whole belief base, takes into account only a maximal consistent subset of it. For selecting this maximal subset, it depends on a given ranking of the formulas of the base, as described by (Benferhat et al., 1995). When an inconsistency is found, the formula with the lowest rank is left out. This gives rise to a non-monotonic inference operator, that cannot be used together with the generalized change paradigm. But the whole point of the approach in (Benferhat et al., 1995) is to keep everything in the belief base and only infer from the consistent parts and not to perform revision, as is our goal. Marquis and Porquet claim that there was no treatment in the literature aiming at both inconsistency tolerance and computational efficiency. But (Chopra et al., 2000) and (Wassermann, 2001) address these two problems, and also the need of some notion of relevance for dealing with memory limitations.

Our message here is that in order to model realistic agents, we need logics which are more sophisticated than classical logic in the sense that they model more features of the agents' reasoning. But we also need logics which give us partial answers when the available resources are not sufficient for finding the classical answers. The solution proposed in this paper is a logic that "does as good as it can", the more resources we have available, the more we can increase the parameters $n$ and $S$, and the closer we get to classical logic.

## References

Alchourrón, C., P. Gärdenfors, and D. Makinson: 1985, 'On The Logic of Theory Change'. *Journal of Symbolic Logic* **50**(2), 510–530.

Alchourrón, C. and D. Makinson: 1982, 'On The Logic of Theory Change: contraction functions and their associated revision functions'. *Theoria* **48**, 14–37.

Anderson, A. and N. Belnap: 1975, *Entailment: The Logic of Relevance and Necessity, Vol. 1.* Princeton University Press.

Benferhat, S., D. Dubois, and H. Prade: 1995, 'How to infer from inconsistent beliefs without revising?'. In: *Proceedings of the International Joint Conference on Artificial Intelligence.*

Cadoli, M. and M. Schaerf: 1995, 'Approximate inference in default logic and circumscription'. *Fundamenta Informaticae* **23**, 123–143.

[6] The same is true of Da Costa's $C_1$.

Cadoli, M. and M. Schaerf: 1996, 'The Complexity of Entailment in Propositional Multivalued Logics'. *Annals of Mathematics and Artificial Intelligence* **18**(1), 29–50.

Chopra, S., R. Parikh, and R. Wassermann: 2000, 'Approximate Belief Revision - Preliminary Report'. In: *Proceedings of the Seventh Workshop on Logic, Language, Information and Comunication (WoLLIC), Natal, Brazil.* To appear in the Journal of the IGPL.

Coniglio, M. E. and W. A. Carnielli: 2001, 'Transfers between Logics and their Applications'. Technical Report Vol. 1(4), CLE e-Prints- UNICAMP. available at http://www.cle.unicamp.br/e-prints/.

da Costa, N. C.: 1963, 'Calculs Propositionnels pour les Systémes Formels Inconsistants'. *Comptes Rendus d'Academie des Sciences de Paris* **257**.

da Costa, N. C. A. and E. H. Alves: 1977, 'A semantical analysis of the calculi $C_n$'. *Notre Dame Journal of Formal Logic* **16**.

Finger, M. and R. Wassermann: 2001a, 'Approximate Reasoning and Paraconsistency - Preliminary Report'. In: *Proceedings of the Eigth Workshop on Logic, Language, Information and Comunication (WoLLIC), Brasília, Brazil.*

Finger, M. and R. Wassermann: 2001b, 'Tableaux for Approximate Reasoning'. In: *Proceedings of the IJCAI-01 Workshop on Inconsistency in Data and Knowledge.*

Fuhrmann, A.: 1991, 'Theory Contraction Through Base Contraction'. *Journal of Philosophical Logic* **20**, 175–203.

Hansson, S. O.: 1989, 'New Operators for Theory Change'. *Theoria* **55**, 114–132.

Hansson, S. O.: 1991, 'Belief Base Dynamics'. Ph.D. thesis, Uppsala University.

Hansson, S. O.: 1992, 'A dyadic representation of belief'. In: P. Gärdenfors (ed.): *Belief Revision*, Vol. 29 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, pp. 89–121.

Hansson, S. O.: 1994, 'Kernel Contraction'. *Journal of Symbolic Logic* **59**, 845–859.

Hansson, S. O.: 1999, *A Textbook of Belief Dynamics.* Kluwer Academic Press.

Hansson, S. O. and R. Wassermann: 1999, 'Local Change'. Technical Report PP-1999-17, ILLC, University of Amsterdam. To appear in Studia Logica.

Herzig, A.: 1997, 'How to change factual beliefs using laws and independence information'. In: D. M. Gabbay, R. Kruse, A. Nonnengart, and H. J. Ohlbach (eds.): *Qualitative and Quantitative Practical Reasoning - First Int. Joint Conf. on Qual. and Quant. Practical Reasoning; ECSQARU-FAPR'97.* pp. 311–321, Springer.

Levesque, H.: 1984, 'A logic of implicit and explicit belief'. In: *Proceedings of AAAI-84.*

Marquis, P. and N. Porquet: 2001, 'Resource-bounded inference from inconsistent belief bases'. In: *Proceedings of the Seventeenth International Joint Conference on Artificial Intelligence (IJCAI).* pp. 103–108, Morgan Kaufmann.

Nebel, B.: 1992, 'Syntax Based Approaches to Belief Revision'. In: P. Gärdenfors (ed.): *Belief Revision.* Cambridge University Press.

Restall, G. and J. Slaney: 1995, 'Realistic Belief Revision'. In: *Proceedings of the Second World Conference on Foundations of Artificial Intelligence (WOCFAI).* pp. 367–378.

Schaerf, M. and M. Cadoli: 1995, 'Tractable Reasoning via Approximation'. *Artificial Intelligence* **74**(2), 249–310.

Wassermann, R.: 2001, 'On Structured Belief Bases'. In: H. Rott and M.-A. Williams (eds.): *Frontiers in Belief Revision.* Kluwer.