

# Construction of Revisions by Explanations

Marcelo A. Falappa

Guillermo R. Simari

Department of Computer Science - Universidad Nacional del Sur  
Bahía Blanca - ARGENTINA  
[ccfalapp,grs]@criba.edu.ar

**Key Words:** Belief Revision, Knowledge Representation, Explanations, Belief Dynamics.

## Abstract

Belief Revision systems are logical frameworks to modeling the dynamics of knowledge. That is, how to modify our beliefs when we receive new information. The main problem arises when that information is inconsistent with the beliefs that represent our epistemic state. For instance, suppose we believe that a Ferrari coupe is the fastest car and then, we found that some Porsche cars are faster than Ferrari cars. Surely, we need to revise our beliefs in order to accept the new information preserving as much old information as possible.

There are many different models for belief revision but AGM is the most popular one. Almost any others are based on the foundations of AGM. They present an epistemic model (the formalism in which the beliefs will be represented) and then define different kinds of operators. The basic representation of epistemic states is through belief sets (set of sentences closed under logical consequence) or belief bases (set of sentences not necessarily closed). Each operator may be presented in two ways: by giving an explicit construction (algorithm) for the operator, or by giving a set of rationality postulates to be satisfied. Rationality postulates determine constraints that the respective operators should satisfy. They treat the operators as black boxes; after receiving certain inputs (of new information) we know what the response will be but not the internal mechanisms used.

The operators for change use selection functions to determine which beliefs will be erased from the epistemic state. Partial meet contractions (AGM model) are based on a selection among subsets of the original set that do not imply the information to be retracted. The kernel contraction approach is based on a selection among the sentences that imply the information to be retracted. Revision operators can be defined through Levi identity; in order to revise an epistemic state  $K$  with respect to a sentence  $\alpha$ , we contract with respect to  $\neg\alpha$  and then expand the new set with respect to  $\alpha$ . On the other hand, consolidations are operators that make set of sentences (non closed under logical consequence) consistent.

One of the most discussed properties of the revision operators is success. Success specifies that the new information has primacy over the beliefs of an agent. We propose a kind of non prioritized revision operator in which the new information is supported by an explanation. Each explanation is a set of sentences with some restrictions. The operator we propose is built in terms of kernel contractions and consolidations. This presentation contains several examples that justify the intuitions behind our model.

# Construction of Revisions by Explanations

**Key Words:** Belief Revision, Knowledge Representation, Explanations, Belief Dynamics.

## 1 Use of Explanations

We can motivate the use of explanations with an example. Suppose that Michael believes that  $(\alpha)$  *all birds fly* and that  $(\beta)$  *Tweety is a bird*. Thus, he will believe that  $(\delta)$  *Tweety flies*. Then, Johana tells him that *Tweety does not fly*. As a consequence, Michael will have to drop the belief in  $\alpha$  or the belief in  $\beta$  forced by having to drop  $\delta$ . However, it does not seem a rational attitude to incorporate any external belief without pondering it. Usually, an *intelligent* agent demands an explanation supporting the provided information. Even more if that information contradicts its own set of beliefs. Being rational, Michael will demand an explanation for the belief  $\neg\delta$ . For instance, Johana accompanies her contention of *Tweety does not fly* with the sentences *Tweety does not fly because it is a penguin* and *penguins are birds but they do not fly*. Perhaps convinced, Michael would have to check his beliefs in order to determine whether he believes *Tweety flies*.

The main role of an explanation is to rationalize facts. At the base of each explanation rests a why-question [Res70]: “Why Tweety does not fly?”, “Why he said what he did?”, “Why is it raining?”. We think that a rational agent, before incorporating a new belief that contradicts its knowledge, demands an explanation for the provided information by means of a why-question. Then, if the explanation resists the discussion, the new belief, together with its explanation, is incorporated to the knowledge.

## 2 Preliminaries

We will adopt a propositional language  $\mathcal{L}$  with a complete set of boolean connectives:  $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$ . Formulae in  $\mathcal{L}$  will be denoted by lowercase Greek characters:  $\alpha, \beta, \delta, \dots$ . Sets of sentences in  $\mathcal{L}$  will be denoted by uppercase Latin characters:  $A, B, C, \dots$ . The symbol  $\top$  represents a tautology or *truth*. The symbol  $\perp$  represents a contradiction or *falsum*. The  $\gamma$  character will be reserved to represent selection functions for change operators. We also use a consequence operator  $Cn$ .  $Cn$  takes sets of sentences in  $\mathcal{L}$  and produces new sets of sentences. The operator  $Cn$  satisfies the following three conditions:

- *Inclusion*:  $A \subseteq Cn(A)$  for all  $A \subseteq \mathcal{L}$ .
- *Iteration*:  $Cn(A) = Cn(Cn(A))$  for all  $A \subseteq \mathcal{L}$ .
- *Monotony*: If  $A \subseteq B$  then  $Cn(A) \subseteq Cn(B)$  for all  $A, B \subseteq \mathcal{L}$ .

We will assume that the consequence operator includes classical consequences and verifies the standard properties of deduction and compactness:

- *Supraclassicality*: If  $\alpha$  can be derived from  $A$  by deduction in classical logic, then  $\alpha \in Cn(A)$ .
- *Deduction*:  $\beta \in Cn(A \cup \{\alpha\})$  if and only if  $(\alpha \rightarrow \beta) \in Cn(A)$ .
- *Compactness*: If  $\alpha \in Cn(A)$  then  $\alpha \in Cn(A')$  for some finite subset  $A'$  of  $A$ .

To simplify notation, we write  $Cn(\alpha)$  for  $Cn(\{\alpha\})$  where  $\alpha$  is any individual sentence in  $\mathcal{L}$ . We also write  $\alpha \in Cn(A)$  as  $A \vdash \alpha$ .

Let  $K$  be a set of sentences. As in the AGM model, we will assume three different epistemic attitudes:

- **Accepted:**  $\alpha$  is accepted in  $K$  if  $\alpha \in Cn(K)$ .
- **Rejected:**  $\alpha$  is rejected in  $K$  if  $\neg\alpha \in Cn(K)$ .
- **Undetermined:**  $\alpha$  is undetermined in  $K$  if  $\alpha \notin Cn(K)$  and  $\neg\alpha \notin Cn(K)$ .

### 3 Formal Definition of Explanation

In order to present a revision operator based on explanations, we will define first different kinds of explanation. We will use a set of sentences as the *explanans* (beliefs that support the new belief) and a single sentence as the *explanandum* (the conclusion of the explanans).

#### 3.1 Previous works about explanations

In [Gär88] Gärdenfors presents some references to works about explanations. He present an account of explanations where the connections between the explanans sentences (in our framework, a set of sentences  $A$ ) and the explanandum ( $\alpha$ ) is always evaluated relative to an epistemic state.

In [WPFS95] the authors used a notion of explanation using transmutations of information systems. They used the Spohn's notion of *reason for* within the general setting of transmutations and extended this to characterize most plausible explanations. Each sentence has assigned a informational value by an ordinal that represents its *firminess* in the epistemic state.

Gärdenfors [Gär88] define a kind of explanation measured by a surprise degree. According to Gärdenfors, when a person (or agent) wants an explanation of a fact  $\alpha$ , he normally already knows that  $\alpha$  is true. So the information conveyed by the explanans does not change the belief in  $\alpha$ . Since the fact  $\alpha$  may be more or less *surprising* or *unexpected*, the principal effect of a successful explanation is that the surprise in  $\alpha$  is decreased. This model is different of our model because request an explanation for a belief in the epistemic state and it use a numerical measure for the surprise degree.

#### 3.2 Evolution of the concept of explanation

Now, we will present different definitions of explanation. We begin with the most simple form of explanation.

**Definition 3.1:** *First Approach.* The set  $A$  is an *explanation* for  $\alpha$  if and only if  $A \vdash \alpha$ .

■

However, note that in this definition, any inconsistent set is an explanation for any belief.

**Definition 3.2:** *Second Approach.* The set  $A$  is an *explanation* for  $\alpha$  if and only if:

1.  $A \vdash \alpha$ .
2. *Consistency:*  $A \not\vdash \perp$ .

■

Omer [Tuo72] has defined a model of explanation using deductive logic. He gives the following criterion for explanation:

In the explanans no sentence which is less informative in the topic should be given when it is possible to give a more informative one.

Next, Omer states as “an important subcase” of this requirement [Tuo72]:

No sentence in the explanans should be of less informative content than the explanandum.

One important question left open by the above requirement is how to measure the information content of a sentence. It is possible to use some kind metrical measure as probability. However, we do not use this kind of measures; we will use a deductive notion of informational content.

**Definition 3.3:** Let  $\alpha$  and  $\beta$  be sentences in  $\mathcal{L}$ . We define the following relationships about informational content:

- If  $\alpha \not\equiv \perp$  then:
  - “ $\alpha$  has more informational content than  $\beta$ ”, noted by  $\alpha >_{\text{ic}} \beta$ , whenever  $\alpha \vdash \beta$  and  $\beta \not\vdash \alpha$ .
  - “ $\alpha$  has the same informational content than  $\beta$ ”, noted by  $\alpha =_{\text{ic}} \beta$ , whenever  $\alpha \vdash \beta$  and  $\beta \vdash \alpha$ .
  - “ $\alpha$  has at least the same informational content as  $\beta$ ”, noted by  $\alpha \geq_{\text{ic}} \beta$ , whenever  $\alpha >_{\text{ic}} \beta$  or  $\alpha =_{\text{ic}} \beta$  (that is,  $\alpha \vdash \beta$ ).
  - “ $\alpha$  is incomparable with  $\beta$ ”, noted by  $\alpha \times_{\text{ic}} \beta$ , whenever  $\alpha \not\geq_{\text{ic}} \beta$  and  $\beta \not\geq_{\text{ic}} \alpha$  (that is,  $\alpha \not\vdash \beta$  and  $\beta \not\vdash \alpha$ ).
- If  $\alpha \equiv \perp$  then  $\beta \geq_{\text{ic}} \alpha$  for all  $\beta \in \mathcal{L}$ . ■

For instance, we have that  $\alpha \geq_{\text{ic}} \alpha \vee \beta$  and  $\alpha \wedge \beta \geq_{\text{ic}} \alpha$  (except in the limit case in which  $\beta = \neg\alpha$ ).

**Observation 3.1:** Relations between  $\geq_{\text{ic}}$ ,  $>_{\text{ic}}$ ,  $\times_{\text{ic}}$  and  $\vdash$ :

1. If  $\alpha \not\equiv \perp$  then  $\alpha \geq_{\text{ic}} \beta$  if and only if  $\alpha \vdash \beta$ .
2. If  $\alpha \not\vdash \beta$  then  $\beta >_{\text{ic}} \alpha$  or  $\alpha \times_{\text{ic}} \beta$ . ■

**Observation 3.2:** The deduction relation  $\vdash$  induces a partial order  $\geq_{\text{ic}}$  among the sentences of the language.

PROOF: We need to show that the relation  $\geq_{\text{ic}}$  satisfies *reflexivity*, *transitivity* and *antisymmetry*.

1. **Reflexivity:**  $\alpha \geq_{\text{ic}} \alpha$  for all  $\alpha \in \mathcal{L}$ .
  - (a)  $\alpha \equiv \perp$ . Then trivially  $\alpha \geq_{\text{ic}} \alpha$ .
  - (b)  $\alpha \not\equiv \perp$ . Since  $\alpha \vdash \alpha$  by Observation 3.1 we have that  $\alpha \geq_{\text{ic}} \alpha$ .
2. **Transitivity:** If  $\alpha \geq_{\text{ic}} \beta$  and  $\beta \geq_{\text{ic}} \delta$  then  $\alpha \geq_{\text{ic}} \delta$  for all  $\alpha, \beta, \delta \in \mathcal{L}$ .
  - (a)  $\delta \equiv \perp$ . Then trivially  $\alpha \geq_{\text{ic}} \delta$ .
  - (b)  $\delta \not\equiv \perp$ . Since  $\beta \geq_{\text{ic}} \delta$  then  $\beta \vdash \delta$  and  $\beta \not\equiv \perp$ . Since  $\alpha \geq_{\text{ic}} \beta$  then  $\alpha \vdash \beta$  and  $\alpha \not\equiv \perp$ . From  $\alpha \vdash \beta$  and  $\beta \vdash \delta$  then  $\alpha \vdash \delta$ . Therefore,  $\alpha \geq_{\text{ic}} \delta$ .
3. **Antisymmetry:** If  $\alpha \geq_{\text{ic}} \beta$  and  $\beta \geq_{\text{ic}} \alpha$  then  $\alpha =_{\text{ic}} \beta$  for all  $\alpha, \beta \in \mathcal{L}$ .
  - (a)  $\alpha \equiv \perp$ . If  $\alpha \geq_{\text{ic}} \beta$  then  $\beta \equiv \perp$ . Therefore  $\alpha =_{\text{ic}} \beta$ .

(b)  $\alpha \neq \perp$ . If  $\alpha \geq_{\text{ic}} \beta$  then  $\alpha \vdash \beta$ . Since  $\beta \geq_{\text{ic}} \alpha$  then  $\beta \vdash \alpha$ . Therefore,  $\alpha =_{\text{ic}} \beta$ . ■

In the last definition of explanation, we can have the following explanation for the sentence  $\alpha$ :  $\{\alpha \vee \beta, \alpha \vee \neg\beta\}$ . In this case, the explanandum could be the explanans of each sentence of the original explanans and the informational content of the explanandum is greater than the informational content of each sentence in the explanans. Therefore, we need to add another constraint to the definition of explanation. We do not want the case in which for all  $\beta \in A$ ,  $\alpha \vdash \beta$  (i.e.,  $\alpha \geq_{\text{ic}} \beta$ ). We can request the following property:

*Strong Informational Content:* For all  $\beta \in A$ ,  $\beta \geq_{\text{ic}} \alpha$  or  $\beta \times_{\text{ic}} \alpha$ .

That is, all the sentences in the explanans have at least the same informational content as the explanandum or they are incomparable to it. However, we can not have the following explanation for  $\alpha$ :  $\{\delta, \delta \rightarrow \alpha\}$  because  $\alpha >_{\text{ic}} (\delta \rightarrow \alpha)$ . Therefore, we need to weaken the above requirement.

**Definition 3.4:** *Third Approach.* The set  $A$  is an *explanation* for  $\alpha$  if and only if:

1.  $A \vdash \alpha$ .
2. *Consistency:*  $A \not\vdash \perp$ .
3. *Weak Informational Content:* If there is some  $\beta \in A$  such that  $\alpha >_{\text{ic}} \beta$  then there is some  $\delta \in (A \setminus \{\beta\})$  such that  $\delta \geq_{\text{ic}} \alpha$  or  $\delta \times_{\text{ic}} \alpha$ . ■

This proposal is quite different from Olmer's approach [Tuo72]. According to Olmer, the explanandum should be incomparable with any of the sentences of the explanans.

In spite of the fact that a sentence in the explanans can get less informational content than the explanandum, the conjunction of the explanation has at least as informational content as the explanandum.

**Observation 3.3:** If  $A$  is an explanation for  $\alpha$  then the information content of  $\wedge(A)$  has at least the same informational content as  $\alpha$ .

PROOF: Since  $A$  is an explanation for  $\alpha$  then  $A \vdash \alpha$  and  $\alpha \neq \perp$ . Then  $\wedge(A) \vdash \alpha$ . Suppose that  $\wedge(A) \not\geq_{\text{ic}} \alpha$ . Due to Observation 3.1, if  $\wedge(A) \not\geq_{\text{ic}} \alpha$  then  $\wedge(A) \not\vdash \alpha$  (absurd). Therefore  $\wedge(A) \geq_{\text{ic}} \alpha$ . ■

However, we can have the following explanation for  $\alpha$ :  $\{\alpha \vee \beta, \alpha \vee \neg\beta, \delta\}$ . Therefore, we can request another restriction for the definition of explanation, following the principle formulated by Olmer [Tuo72]:

The explanans should contain redundant information only if it is relevant.

We can approximate this criterion demanding minimality in the explanation.

**Definition 3.5:** *Fourth Approach.* The set  $A$  is an *explanation* for  $\alpha$  if and only if:

1.  $A \vdash \alpha$ .
2. *Consistency:*  $A \not\vdash \perp$ .
3. *Weak Informational Content:* If there is some  $\beta \in A$  such that  $\alpha >_{\text{ic}} \beta$  then there is some  $\delta \in (A \setminus \{\beta\})$  such that  $\delta \geq_{\text{ic}} \alpha$  or  $\delta \times_{\text{ic}} \alpha$ .
4. *Minimality:* If  $B \subset A$  then  $B \not\vdash \alpha$ .

■

Due to the compactness of  $Cn$  and by minimality, all explanations are finite. This definition does not correspond exactly with the constraint suggested by Omer but we will adopt it. Again, we have a problem in some cases:  $\{\alpha \vee \beta, (\alpha \vee \neg\beta) \wedge \delta\}$  could be an explanation for  $\alpha$ . This explanation is logically equivalent to  $\{\alpha \vee \beta, \alpha \vee \neg\beta, \delta\}$  but syntactically different from it. Therefore, we need to add syntactic restrictions to the sentences in the explanans.

**Definition 3.6:** Let  $A$  be a set of sentences. The set  $A$  is in *conjunctive normal form* [Dav89] if and only if each sentence in  $A$  is:

- i) An atomic sentence (positive literal).
- ii) The negation of an atomic sentence (negative literal).
- iii) A disjunction of (positive or negative) literals.

■

Each formulae in  $\mathcal{L}$  can be represented as a conjunction of disjunctions of literals. Therefore, for all  $\alpha \in \mathcal{L}$  there is a set  $A$  such that  $A \vdash \alpha$  and  $A$  is in conjunctive normal form. We propose a more flexible way to represent sentences in conjunctive normal form.

**Definition 3.7:** Let  $A$  be a set of sentences. The set  $A$  is in *weak conjunctive normal form* if and only if each sentence in  $A$  is:

- i) An atomic sentence (positive literal).
- ii) The negation of an atomic sentence (negative literal).
- iii) A disjunction of (positive or negative) literals or a logical implication of the form  $(\alpha_1 \wedge \dots \wedge \alpha_n) \rightarrow \beta$  where  $\alpha_1, \dots, \alpha_n, \beta$  are positive or negative literals.

■

Finally, we will give the last definition of explanation.

**Definition 3.8:** *Fifth Approach.* The set  $A$  is an *explanation* for  $\alpha$  if and only if:

1.  $A$  is in *weak conjunctive normal form*.
2.  $A \vdash \alpha$ .
3. *Consistency:*  $A \not\vdash \perp$ .
4. *Weak Informational Content:* If there is some  $\beta \in A$  such that  $\alpha >_{\text{ic}} \beta$  then there is some  $\delta \in (A \setminus \{\beta\})$  such that  $\delta \geq_{\text{ic}} \alpha$  or  $\delta \times_{\text{ic}} \alpha$ .
5. *Minimality:* If  $B \subset A$  then  $B \not\vdash \alpha$ .

We will note the relationship “ $A$  is an explanation for  $\alpha$ ” as  $A \twoheadrightarrow \alpha$ .

■

## 4 Consolidation Operator

The consolidation operator provides a mechanism to treat local inconsistencies in belief bases. This operator is defined as a contraction with respect to  $\perp$ . The idea is to give up all possible contradictions in a belief base. The most common approach to contraction is partial meet contraction, proposed by AGM [AGM85] and defined as follow:

**Definition 4.1** (Alchourrón & Makinson [AM81]): Let  $K$  be a set of sentences and  $\alpha$  a sentence. The set  $K \perp \alpha$  is the set of sets such  $K' \in K \perp \alpha$  if and only if:

1.  $K' \subseteq K$ .
2.  $K' \not\models \alpha$ .
3. There is no set  $K''$  such that  $K' \subset K'' \subseteq K$  and  $K'' \not\models \alpha$ .

The set  $K \perp \alpha$  is called the *remainder set* of  $K$  with respect to  $\alpha$ . ■

**Definition 4.2** (Alchourrón, Gärdenfors & Makinson [AGM85]): A *selection function* for a set  $K$  of sentences is a function  $\gamma$  ( $\gamma : 2^K \Rightarrow 2^K$ ) such that, for all sentence  $\alpha$ :

1. If  $K \perp \alpha$  is non-empty, then  $\gamma(K \perp \alpha)$  is a non-empty set of  $K \perp \alpha$ .
  2. If  $K \perp \alpha$  is empty then  $\gamma(K \perp \alpha) = \{K\}$ .
- 

**Definition 4.3** (Alchourrón, Gärdenfors & Makinson [AGM85]): Let  $K$  be a set of sentences and  $\gamma$  a selection function for  $K$ . The *partial meet contraction* on  $K$  generated by  $\gamma$  is the operation  $\div \gamma$  such that, for all sentences  $\alpha$ :  $K \div \gamma \alpha = \cap \gamma(K \perp \alpha)$ . An operation  $-$  is a partial meet contraction if and only if there is a selection function  $\gamma$  such that for all sentences  $\alpha$  holds that  $K - \alpha = K \div \gamma \alpha$ . ■

The partial meet contraction of  $K$  with respect to  $\alpha$  is equal to the intersection of the set of *selected* maximal subsets of  $K$  that do not imply  $\alpha$ . However, we do not talk about restrictions on selection functions. The following definition refers to a special case when the selection function is based on a “preference” relation.

**Definition 4.4** (Alchourrón, Gärdenfors & Makinson [AGM85]): Let  $K$  be a set of sentences. A selection function  $\gamma$  for  $K$  and the contraction operator based on it are:

1. *relational* if and only if there is a binary relation  $\sqsubseteq$  such that for all sentences  $\alpha$ , if  $K \perp \alpha \neq \emptyset$  then:

$$\gamma(K \perp \alpha) = \{K' : K' \in K \perp \alpha \text{ and } K'' \sqsubseteq K' \text{ for all } K'' \in K \perp \alpha\}$$

2. *transitively relational* if and only if there is a relational binary relation that is transitive.
- 

Based on partial meet contraction, we can define partial meet consolidation as a partial meet contraction by falsum ( $\perp$ ) [Han97b]. The partial meet consolidation of  $K$  with respect to  $\perp$  is the intersection of the “most preferred” maximal consistent subsets of  $K$ .

The following representation theorem has been obtained by Hansson for partial meet consolidation.

**Theorem 4.1** (Hansson [Han97b]): An operation  $!$  is an operation of *partial meet consolidation* if and only if for all sets  $K$  of sentences:

*Strong Consistency:*  $K!$  is consistent.

*Inclusion:*  $(K!) \subseteq K$ .

*Relevance:* If  $\alpha \in K \setminus (K!)$  then there is some  $K'$  such that  $K! \subseteq K' \subseteq K$ ,  $K'$  is consistent and  $(K' \cup \{\alpha\})$  is inconsistent. ■

## 5 Kernel Consolidation

An alternative to partial meet consolidation is kernel contraction. The kernel contraction is based on the observation that a set  $K$  implies a sentence  $\alpha$  if and only if contains some minimal subset that implies  $\alpha$ . Therefore, in order to remove  $\alpha$  from  $K$ , it is necessary to give up at least one element of each minimal  $\alpha$ -implying subset of  $K$  [Han97b].

**Definition 5.1** (Hansson [Han93a]): Let  $K$  be a set of sentences and  $\alpha$  a sentence. Then  $K \perp\!\!\!\perp \alpha$  is the set such that  $K' \in K \perp\!\!\!\perp \alpha$  if and only if:

1.  $K' \subseteq K$ .
2.  $K' \vdash \alpha$ .
3. If  $K'' \subset K'$  then  $K'' \not\vdash \alpha$ .

The set  $K \perp\!\!\!\perp \alpha$  is called the *kernel set*, and its elements are called the  $\alpha$ -*kernels* of  $K$ . ■

**Definition 5.2** (Hansson [Han93a]): An *incision function*  $\sigma$  for  $K$  is a function such that for all  $\alpha$ :

1.  $\sigma(K \perp\!\!\!\perp \alpha) \subseteq \cup(K \perp\!\!\!\perp \alpha)$ .
2. If  $\emptyset \subset K' \in K \perp\!\!\!\perp \alpha$  then  $(K' \cap \sigma(K \perp\!\!\!\perp \alpha)) \neq \emptyset$ .

Furthermore, an incision function  $\sigma$  for  $K$  is *smooth* if and only if it holds for all subsets  $K'$  of  $K$  that if  $K' \vdash \beta$  and  $\beta \in \sigma(K \perp\!\!\!\perp \alpha)$  then  $(K' \cap \sigma(K \perp\!\!\!\perp \alpha)) \neq \emptyset$ . ■

**Definition 5.3** (Hansson [Han93a]): Let  $\sigma$  be an incision function for  $K$ . The *kernel contraction* on  $K$  generated by  $\sigma$  is the operation  $\div \sigma$  such that, for all sentences  $\alpha$ :

$$K \div \sigma \alpha = K \setminus (\sigma(K \perp\!\!\!\perp \alpha))$$

An operator  $-$  for  $K$  is a kernel contraction if and only if there is some incision function  $\sigma$  for  $K$  such that  $K - \alpha = K \div \sigma \alpha$  for all sentences  $\alpha$ . Furthermore, if  $\sigma$  is a *smooth* incision function then  $-$  is a *smooth kernel contraction*. ■

The consolidation is defined as a contraction with respect to falsum. Then, a (smooth) kernel consolidation is defined as a (smooth) kernel contraction with respect to falsum. The following representation theorem has been obtained by Hansson for kernel consolidation and smooth kernel consolidation.

**Theorem 5.1** (Hansson [Han97b]): An operation  $!$  is an operation of *kernel consolidation* if and only if for all sets  $K$  of sentences:

*Strong Consistency:*  $K!$  is consistent.

*Inclusion:*  $(K!) \subseteq K$ .



*Core-retainment:* If  $\alpha \in K \setminus (K!)$  then there is some  $K'$  such that  $K' \subseteq K$ ,  $K'$  is consistent and  $(K' \cup \{\alpha\})$  is inconsistent.

Furthermore,  $!$  is an operation of *smooth kernel consolidation* if and only if it satisfies *consistency, inclusion, core-retainment* and:

*Relative Closure:*  $K \cap Cn(K!) \subseteq K!$ . ■

The main difference between partial meet contraction (consolidation) and kernel contraction (consolidation) is that the first one uses an order among the subsets of  $K$  while the second one (by means of an incision function) uses an order among the sentences of  $K$ . Furthermore, since relevance implies core-retainment we can see that any partial meet consolidation is a kernel consolidation. We may see kernel consolidation (contraction) as a generalization of partial meet consolidation (contraction).

## 6 AGM Revision

The AGM model is the most popular model of belief revision. This model is defined upon belief sets (set of sentences closed under logical consequence). The revision operator is characterized by *rationately postulates* [AGM85, Gär88]. These postulates should be satisfied by any revision operator applied upon belief sets. Let  $\mathbf{K}$  be a belief set (*i.e.*,  $\mathbf{K} = Cn(\mathbf{K})$ ),  $*$  a revision operator and  $\alpha$  a sentence. Then, a AGM revision operator satisfies the following postulates:

- i) *Closure:*  $\mathbf{K} * \alpha = Cn(\mathbf{K} * \alpha)$ .
- ii) *Success:*  $\mathbf{K} * \alpha \vdash \alpha$ .
- iii) *Inclusion:*  $\mathbf{K} * \alpha \subseteq Cn(\mathbf{K} \cup \{\alpha\})$ .
- iv) *Vacuity:* If  $\mathbf{K} \not\vdash \neg\alpha$  then  $\mathbf{K} * \alpha = Cn(\mathbf{K} \cup \{\alpha\})$ .
- v) *Consistency:* If  $\not\vdash \neg\alpha$  then  $\mathbf{K} * \alpha \not\vdash \perp$ .
- vi) *Extensionality:* If  $\vdash \alpha \leftrightarrow \beta$  then  $\mathbf{K} * \alpha = \mathbf{K} * \beta$ .
- vii) *Superexpansion:*  $\mathbf{K} * (\alpha \wedge \beta) \subseteq Cn((\mathbf{K} * \alpha) \cup \{\beta\})$ .
- viii) *Subexpansion:* If  $\mathbf{K} * \alpha \not\vdash \neg\beta$  then  $Cn((\mathbf{K} * \alpha) \cup \{\beta\}) \subseteq \mathbf{K} * (\alpha \wedge \beta)$ .

These postulates allow to characterize AGM revision upon belief sets and it is called *partial meet revision*. A characterization of partial meet revision on belief bases was presented by Hansson [Han93b].

**Theorem 6.1** (Hansson [Han93b]): The operator  $*$  is an operator of partial meet revision for a belief base  $K$  if and only if it satisfies:

- i) *Consistency:* If  $\not\vdash \neg\alpha$  then  $K * \alpha \not\vdash \perp$ .
- ii) *Inclusion:*  $K * \alpha \subseteq K \cup \{\alpha\}$ .
- iii) *Relevance:* If  $\beta \in K$  and  $\beta \notin K * \alpha$  then there is some  $K'$  such that  $K * \alpha \subseteq K' \subseteq K \cup \{\alpha\}$ ,  $K'$  is consistent but  $K' \cup \{\alpha\}$  is inconsistent.
- iv) *Success:*  $K * \alpha \vdash \alpha$ .

- v) *Uniformity*: If for all  $K' \subseteq K$ ,  $K' \cup \{\alpha\}$  is inconsistent if and only if  $K' \cup \{\beta\}$  is inconsistent, then  $K \cap (K * \alpha) = K \cap (K * \beta)$ . ■

The AGM revision operator has some discussed properties such as success and consistency. We (and many other belief revision researchers) think that these postulates are the “weak side” of the model. Success gives more priority to new information. In the real world, we may not want to give top priority to new information. Moreover, if it conflict with the old information, we may wish to weight it against the old pieces of information, and if this is really “better” or “more plausible” than the old information, we may wish to accept it; elsewhere, we could wish to preserve all old information.

## 7 Revision using explanations in terms of AGM-revision and consolidation

Now, we will present a new definition of revision using explanations. The operator of revision by explanations is a non-prioritized operator [Han97a, Mak97] since it does not give more priority to the new information. In order to define this operator, we need to define the set of counter-explanations. The operator was previously defined in [FaSi95, FaSi96] but this presentation is given in terms of others change operators.

**Definition 7.1** (Hansson [Han96]): The function  $\mathbf{n}$  is a mapping from set of sentences to sentences ( $\mathbf{n} : 2^{\mathcal{L}} \Rightarrow \mathcal{L}$ ). Let  $A$  be a finite set of sentences. The *sentential negation* of  $A$ , noted by  $\mathbf{n}(A)$ , is defined as follows:

1.  $\mathbf{n}(\emptyset) = \perp$ .
2. If  $A$  is a singleton,  $A = \{\alpha\}$ , then  $\mathbf{n}(A) = \neg\alpha$ .
3. If  $A = \{\alpha_1, \dots, \alpha_m\}$  for some  $m > 1$ , then  $\mathbf{n}(A) = \neg\alpha_1 \vee \dots \vee \neg\alpha_m$ . ■

An revision operator can be defined by means of two suboperations: contraction and expansion. For instance, if we have to revise a set  $K$  with respect to a sentence  $\alpha$ , we first contract  $K$  by  $\neg\alpha$  and then expand by  $\alpha$ . This process is knowed as Levi identity. We may define a multiple partial meet revision operator using the Levi identity.

**Definition 7.2** (Hansson [Han93b, Han96]): Let  $K$  be a set of sentences,  $A$  a finite set of sentences and  $-$  a partial meet contraction for  $K$ . The *multiple partial meet revision* of  $K$  with respect to  $A$  is defined as  $K * A = (K - \mathbf{n}(A)) \cup A$ . ■

Since in our model the explanations are finite, we will use this revision operator. Hansson [Han92, Han93b] has showed that the revision of belief bases can be generalized for revision by finite sets.

**Definition 7.3**: Let  $K$  be a set of sentences and  $A$  an explanation for  $\alpha$ . The set of *counter-explanations* of  $A$  with respect to  $K$  is the set  $K \bowtie A$  such that  $K \bowtie A = \cup(K \perp \mathbf{n}(A))$ . ■

**Definition 7.4**: Let  $K$  be a set of sentences,  $*$  a partial meet revision operator,  $A$  an explanation for  $\alpha$  and  $!$  a (partial meet) consolidation operator [Han97b]. The *revision using explanations* of  $K$  with respect to  $A$  is the following:

$$K \circ A = \begin{cases} K * A & \text{if } A \subseteq ((K \bowtie A) \cup A) \\ K & \text{otherwise} \end{cases}$$

■

The idea is to attack the external explanation with our explanations against it.  $A \subseteq (K \bowtie A \cup A)!$  means that the external explanation has “survived” in the consolidation process. Then, we accept the external explanation and we make an multiple partial meet revision with respect to the explanation. If  $A$  does not survive, then the epistemic state remains without changes.

**Example 7.1:** Suppose we have an epistemic state represented by the set  $K$  such that:

$$K = \{\alpha, \alpha \rightarrow \beta, \beta \rightarrow \delta, \phi\}$$

Then, we receive the following external explanation for  $\neg\beta$ :

$$A = \{\alpha, \varphi, \alpha \wedge \varphi \rightarrow \neg\beta\}$$

It is easy to check that  $A$  is a correct explanation for  $\neg\beta$  (Definition 3.8). Now, we need to get the set of counter-explanations of  $\mathbf{n}(A)$ .

$$\mathbf{n}(A) \equiv \neg\alpha \vee \neg\varphi \vee \neg(\alpha \wedge \varphi \rightarrow \neg\beta) \equiv \neg\alpha \vee \neg\varphi \vee \neg(\alpha \wedge \varphi \wedge \beta)$$

$$\mathbf{n}(A) \equiv ((\neg\alpha \vee \neg\varphi \vee \alpha) \wedge (\neg\alpha \vee \neg\varphi \vee \varphi) \wedge (\neg\alpha \vee \neg\varphi \vee \beta)) \equiv \top \wedge \top \wedge (\neg\alpha \vee \beta \vee \neg\varphi) \equiv (\neg\alpha \vee \beta \vee \neg\varphi)$$

The set of counter-explanations of  $A$  with respect to  $K$  is:

$$K \bowtie A = \cup(K \perp \mathbf{n}(A)) = \{\alpha \rightarrow \beta\}$$

Now, we have to make a consolidation of the set:

$$H = \{\alpha \rightarrow \beta, \alpha, \varphi, \alpha \wedge \varphi \rightarrow \neg\beta\}$$

We need an order among the sentences of  $H$ . We can use the specificity relation [Poo85, MTDR92] to decide which sentence(s) will be erased. The set  $\{\alpha \wedge \varphi \rightarrow \neg\beta\}$  of rules seems more specific than the set  $\{\alpha \rightarrow \beta\}$ . Then, we could retract  $\alpha \rightarrow \beta$  and the set  $H$  consolidated could be:

$$H! = \{\alpha, \varphi, \alpha \wedge \varphi \rightarrow \neg\beta\}$$

Since  $A \subseteq H!$  then  $A$  has survived. Then, the revision by explanation of  $K$  with respect to  $A$  will be  $K \circ A = K * A$ . In order to make possible the AGM revision of  $K$  with respect to  $A$  we need an order (possibly different from the above one) among the sentences (or subsets) of  $K$ . A possible outcome of this change is:

$$K \circ A = \{\alpha, \varphi, \alpha \wedge \varphi \rightarrow \neg\beta, \beta \rightarrow \delta, \phi\}$$

■

**Example 7.2:** Suppose we have an epistemic state represented by the set  $K$  such that:

$$K = \{\alpha, \beta, \alpha \rightarrow \delta, \delta \rightarrow \phi\}$$

Then, we receive the following external explanation for  $\neg\delta$ :

$$A = \{\alpha, \alpha \rightarrow \varphi, \varphi \rightarrow \neg\delta\}$$

It is easy to check that  $A$  is a correct explanation for  $\neg\delta$ . Now, we need to get the set of counter-explanations of  $\mathbf{n}(A)$ .

$$\mathbf{n}(A) \equiv \neg\alpha \vee \neg(\alpha \rightarrow \varphi) \vee \neg(\varphi \rightarrow \neg\delta)$$

$$\mathbf{n}(A) \equiv \neg\alpha \vee \neg(\neg\alpha \vee \varphi) \vee \neg(\neg\varphi \vee \neg\delta) \equiv \neg\alpha \vee (\alpha \wedge \neg\varphi) \vee (\varphi \wedge \delta)$$

$$\begin{aligned}\mathbf{n}(A) &\equiv ((\neg\alpha \vee \alpha) \wedge (\neg\alpha \vee \neg\varphi)) \vee (\varphi \wedge \delta) \equiv (\top \wedge (\neg\alpha \vee \neg\varphi)) \vee (\varphi \wedge \delta) \\ \mathbf{n}(A) &\equiv (\neg\alpha \vee \neg\varphi \vee \varphi) \wedge (\neg\alpha \vee \neg\varphi \vee \delta) \equiv (\neg\alpha \vee \neg\varphi \vee \delta) \equiv (\neg\alpha \vee \delta \vee \neg\varphi)\end{aligned}$$

The set of counter-explanations of  $A$  with respect to  $K$  is:

$$K \bowtie A = \cup(K \perp \mathbf{n}(A)) = \{\alpha \rightarrow \delta\}$$

Now, we have to make a consolidation of the set:

$$H = \{\alpha \rightarrow \delta, \alpha, \alpha \rightarrow \varphi, \varphi \rightarrow \neg\delta\}$$

Again, if we use specificity [Poo85, MTDR92] we have that the set  $\{\alpha \rightarrow \delta\}$  of rules is better than the set  $\{\alpha \rightarrow \varphi, \varphi \rightarrow \neg\delta\}$ . So we need to retract  $\alpha \rightarrow \varphi$  or  $\varphi \rightarrow \neg\delta$  or both. Therefore,  $A$  does not survive in the consolidation process. Then, the outcome of the revision will be:

$$K \circ A = K = \{\alpha, \beta, \alpha \rightarrow \delta, \delta \rightarrow \phi\}$$

■

In this construction there is not a strong relation between the concept of explanation and the revision operator. The key is that an explanation is not defined with respect to a change operator. We are just formulating a multiple revision operator with respect to a set with some restrictions.

## 8 Conclusions

We have proposed a kind of non prioritized revision operator in which the new information is supported by an explanation. Each explanation is a set of sentences with some restrictions. The operator we have proposed is built in terms of kernel contractions and consolidations. We defined this operator because we think that the classical revision operator (as AGM) does not correspond with the real world. For instance, we do not agree with some postulates such as success and consistency. We think that the revisions using explanations are typical in several contexts.

But this model has a disadvantage; it needs two orders: one for sentences or subsets of the language, another one for the sentences or subsets of  $K$ . However, the two kinds of relation appear to have different sources. When we perform an ordinary revision (or contraction), the relation that is used concerns only the sentences in  $K$ . When we perform a revision by explanations, we are comparing new information (more exactly new explanations) typically from outside  $K$ , so we need a superior order among sentences or subsets not necessarily included in  $K$ . This “external ordering” (with respect to  $K$ ) should be more fundamental than the ordering in  $K$  (first, sentences or subsets of  $\mathcal{L}$ ; second, sentences or subsets of  $K$ ). A similar problem is present in the second definition of screened revision [Mak97], where the order among sentences of the language is more fundamental than the order among the subsets of  $K$  (first, sentences in  $\mathcal{L}$ ; second, subsets of  $K$ ).

## Acknowledgements

We would like to thank Sven Ove Hansson and Eduardo Fermé for discussion on some concepts included in this paper, and an anonymous referee who provided useful comments.

## References

- [AGM85] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. *On the Logic of Theory Change: Partial Meet Contraction and Revision Functions*. **The Journal of Symbolic Logic**, 50: 510–530, 1985.
- [AM82] Carlos Alchourrón and David Makinson. *On the Logic of Theory Change: Contraction Functions and their Associated Revision Functions*. **Theoria**, 48: 14–37, 1982.
- [AM81] Carlos Alchourrón and David Makinson. *Hierarchies of Regulation and Their Logic*. **New Studies in Deontic Logic**, D. Reider Publishing Company: 125–148, 1981.
- [Dav89] Davis, Ruth: *Truth, Deduction and Computation: Logic and Semantics for Computer Science*. **Computer Science Press**, New York, United States of America, 1989.
- [FaSi95] Marcelo Falappa and Guillermo Falappa. *Propiedades del Operador de Revisión mediante Argumentos*. **Congreso Argentino de Ciencias de la Computación, CACIC’95**, Universidad Nacional del Sur, 1995.
- [FaSi96] Marcelo Falappa and Guillermo Falappa. *Actualización de Bases de Conocimiento utilizando Explicaciones Externas*. **Encuentro de Trabajo en Temas de Lógica No Estándar, Vaquerías’96**, Universidad Nacional de Córdoba, 1996.
- [Gär88] Peter Gärdenfors. *Knowledge in Flux: Modelling the Dynamics of Epistemic States*. The MIT Press, Bradford Books, Cambridge, Massachusetts, 1988.
- [Han92] Sven Ove Hansson. *A Dyadic Representation of Belief*. In **Belief Revision**, pages 89–121, 1992.
- [Han93a] Sven Ove Hansson. *Kernel Contraction*. **The Journal of Symbolic Logic**, 1993.
- [Han93b] Sven Ove Hansson. *Reversing the Levi Identity*. **The Journal of Philosophical Logic**, 22: 637–669, 1993.
- [Han96] Sven Ove Hansson. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Uppsala University, Department of Philosophy, Uppsala, Sweden, 1996.
- [Han97a] Sven Ove Hansson. *Theoria: Special Issue on Non-Prioritized Belief Revision*. Department of Philosophy, Uppsala University, 1997.
- [Han97b] Sven Ove Hansson. *Semi-Revision*. **Journal of Applied Non-Classical Logic**, in press, 1997.
- [Mak97] David Makinson. *Screened Revision*. In **Theoria** [Han97a], 1997.
- [PNF94] Maurice Pagnucco, Abhaya Nayak, and Norman Foo. *Abduction Expansion: Abductive Inference and the Process of Belief Change*. **Proceedings of the Seventh Australian Conference on Artificial Intelligence**, 1994.
- [Pag96] Maurice Pagnucco. *The Role of Abductive Reasoning within the Process of Belief Revision*. PhD. Thesis. Basser Department of Computer Science. University of Sydney, Australia. 1996.

- [Poo85] David Poole. *On the Comparison of Theories: Preferring the Most Specific Explanation*. **Proceedings of IJCAI-85, Los Angeles (CA)**, pages 144–147, 1985.
- [Res70] Nicolas Rescher. *Scientific Explanation*. **The Free Press**, New York, 1970.
- [Tuo72] Raimo Tuomela. *Deductive Explanation of Scientific Laws*. **Journal of Philosophical Logic**, 1, pages 369–392, 1972
- [MTDR92] Guillermo R. Simari and Ron P. Loui. *A Mathematical Treatment of Defeasible Reasoning and its Implementation*. **Artificial Intelligence**, 53: 125–127, 1992.
- [WPFS95] Mary-Anne Williams, Maurice Pagnucco, Norman Foo and Brailey Sims. *Determining Explanations using Transmutations*. In proceedings of **IJCAI 1995**: 822–830, 1995