J. P. Delgrande
A. C. Nayak
M. Pagnucco

Gricean Belief Change

Abstract. One of the standard principles of rationality guiding traditional accounts of belief change is the *principle of minimal change*: a reasoner's belief corpus should be modified in a minimal fashion when assimilating new information. This rationality principle has stood belief change in good stead. However, it does not deal properly with all belief change scenarios. We introduce a novel account of belief change motivated by one of Grice's maxims of *conversational implicature*: the reasoner's belief corpus is modified in a minimal fashion to assimilate *exactly* the new information. In this form of belief change, when the reasoner revises by new information $p \vee q$ their belief corpus is modified so that $p \vee q$ is believed but stronger propositions like $p \wedge q$ are not, no matter what beliefs are in the reasoner's initial corpus. We term this *conservative belief change* since the revised belief corpus is a conservative extension of the original belief corpus given the new information.

Keywords: Belief change, belief revision and update, Gricean conversational implicature.

1. Introduction and overview

A reasoning entity has need to maintain a stock of beliefs—its belief corpus—and modify this corpus to assimilate new information as it is acquired. Methods for modifying a belief corpus are subject to criteria for *rationality*. In traditional accounts of belief change [1, 3] the most common guiding criterion is the *principle of minimal change*. This principle states that a belief corpus should be modified in a minimal fashion when assimilating new information. The forms in which this principle have been employed are various [12, 14]. In accounts of belief change like that proposed by Alchourrón, Gärdenfors and Makinson [1, 3] the principle of minimal change is clearly evident in the *epistemic entrenchment* construction of Gärdenfors and Makinson [4] and the *system of spheres* construction of Grove [6].

We can also distinguish different fundamental forms of belief change. In the literature at least two are ubiquitous: *revision* and *update*. Intuitively, revision deals with the change that occurs when the reasoner's beliefs about its environment are incomplete and possibly incorrect. Any new information fills in these gaps or rectifies mistaken beliefs. However, the environment is assumed static and does not change. Update on the other hand deals with

Presented by **P. Peppas and G. Brewka**; *Received* August 1, 2004

Studia Logica **68**: 1–18, 2004.

dynamic environments in which the new information reflects changes brought about by actions that have occurred.

Here, we investigate a novel form of belief change that is orthogonal to the notions of revision and update where the notion of "minimal change" is interpreted in terms of the new information. As the reasoner assimilates new information into its belief corpus, its aim is to incorporate *all and only* the new information that it has received. This principle is modelled after one of Grice's [5] maxims of conversational implicature in which, when interpreting a speaker, we assume that the meaning of their utterance is *exactly* what they say and no more or less than that. Analogously, in the form of belief change we introduce, the reasoner's aim is to assimilate *exactly* the new information. As such, if the reasoner's belief corpus includes $p \wedge q$ and they were then told $p \vee q$, their belief corpus would be modified to include $p \vee q$ but not $p \wedge q$. This represents a weakening of the reasoner's belief corpus. In traditional accounts of belief change, like the AGM, no modification to the reasoner's belief corpus would be necessary in this case. A consequence of our approach is that the modified belief corpus is a conservative extension of the new information and so we term our proposal *conservative belief change*. We provide both rationality postulates and a construction in terms of orderings over logical interpretations together with a representation result linking these characterisations. We also discuss a context-dependent version of our approach.

In the next section we motivate our proposal through a series of examples. Section 3 provides the necessary background material. In Section 4 we outline our proposed method of belief change. Section 5 discusses the significance of these results and Section 6 presents our conclusions.

2. Motivation

We begin with an example of how a traditional account of belief change based on the principle of minimal change would modify a reasoner's belief corpus when faced with disjunctive information. Here we can take $K * \phi$ to be an AGM revision operation specifying how the reasoner's belief corpus K is to be revised so as to assimilate new information ϕ.

EXAMPLE 2.1 (Exclusive disjunctive revision). Leslie and Robin are two students who share a flat above your's. They are independent and have their own circles of friends. One evening, you believe that both are out of town, $K \equiv \neg l \wedge \neg r$. However, you subsequently hear unmistakable sounds of domestic activity. You modify your beliefs minimally to account for this new information, and so you conclude just that one of them has not gone out, i.e. $K * (l \vee r) \equiv (l \leftrightarrow \neg r)$.[1]

Note that the outcome above is only one of the possible outcomes that is consistent with the AGM approach and others are possible. It does however comply with distance based approaches like those of Dalal [2] and Winslett [15] and arguably captures the spirit of minimal change.

However, this outcome is not always the intended one. The following example illustrates another common possibility.

EXAMPLE 2.2 (Inclusive disjunctive revision). [2] There are two rooms in a warehouse, on the left and on the right. Let l and r denote the fact that the respective rooms are not empty. You believe that there are a number of boxes outside the warehouse and the rooms are empty, and so $K \equiv \neg l \wedge \neg r$. You are later informed that it had been raining, and the boxes had been moved inside. You conclude just that the rooms are not empty, i.e. $K*(l\vee r) \equiv (l\vee r)$.

This example conflicts with the previous one which appears to adhere to the principle of minimal change. Applying a distance based approach to this example would dictate that all the boxes be in exactly one of the rooms.

EXAMPLE 2.3 (Generalised inclusive disjunctive revision). A robbery has taken place; with no other information, we have $K \equiv \exists x R(x)$, that someone is a robber. We then learn that there were exactly three people A, B, and C present at the time of the robbery, i.e. $\phi = (R(A) \vee R(B) \vee R(C))$. We conclude that $K*\phi \equiv (R(A)\vee R(B)\vee R(C))$ – i.e. the robber(s) constitutes a (nonempty) subset of $\{A, B, C\}$. However standard accounts of minimization (e.g., [2]) stipulate that $K*\phi$ entail that R is true of exactly one of $\{A, B, C\}$.

Another well-known example of this phenomenon is due to Reiter:

EXAMPLE 2.4 (Inclusive disjunctive update). A coin is thrown onto a chessboard. If l and r denote the fact that the coin touches a black or a red square, then initially $K \equiv \neg l \wedge \neg r$. One concludes that $K * (l \vee r) \equiv (l \vee r)$, again counter to the principle of minimal change.

[1] We use $\leftrightarrow$ for *material biconditional* and $\equiv$ for *logical equivalence*.

[2] It is contentious whether this example illustrates update or revision. We take it to be revision since information about the current state of the world is learned.

3. Background

We shall consider a framework based on a finitary propositional language $\mathcal{L}$, over a set of atoms, or propositional letters, $\mathbf{P} = \{a, b, c, \dots\}$, and truth-functional connectives $\neg$, $\wedge$, $\vee$, $\implies$, and $\leftrightarrow$. $\mathcal{L}$ also includes the truth-functional constants $\top$ and $\bot$. To clarify the presentation we shall use the following notational conventions. Upper-case Roman characters (A, $B, \dots$) denote *consistent conjunctions of literals* from $\mathcal{L}$. Lower-case Greek characters (ϕ, ψ, ξ, $\dots$) denote arbitrary sentences of $\mathcal{L}$.

An *interpretation* of $\mathcal{L}$ is a function from $\mathbf{P}$ to $\{T, F\}$; M is the set of interpretations of $\mathcal{L}$. A *model* of a sentence ϕ is an interpretation that makes ϕ true, according to the usual definition of truth. A model can be equated with its defining set of literals. $|\phi|_{\mathcal{L}}$ denotes the set of models of sentence ϕ over language $\mathcal{L}$. For interpretation ω we write $\omega \models \phi$ for ϕ is true in ω. For $\phi \in \mathcal{L}$, we will define $\mathcal{L}(\phi)$, the language in which ϕ is expressed, as comprising the minimum set of atoms required to express ϕ, as follows, where ϕ^p_q is the result of substituting atom q everywhere for p in ϕ:

$$\mathcal{L}(\phi) = \{p \in \mathbf{P} \mid \phi^p_\top \not\equiv \phi^p_\bot\} \cup \{\top, \bot\}$$

This set of atoms is unique. Thus $\mathcal{L}(p \wedge (q \vee \neg q)) = \mathcal{L}(p) = \{p, \top, \bot\}$. This can be extended to sets of sentences in the obvious way. It follows trivially that if $\models \phi \leftrightarrow \psi$ then $\mathcal{L}(\phi) = \mathcal{L}(\psi)$. Also note that if $\models \phi$ then $\mathcal{L}(\phi) = \{\top, \bot\}$.

We will also make use of the notion of a *conservative extension* of one set of sentences by another.

DEFINITION 3.1. For sets of sentences $\Gamma_1 \subseteq \Gamma_2 \subseteq \mathcal{L}$ we have that Γ_2 is a *conservative extension* of Γ_1 iff for every $\phi \in \mathcal{L}(\Gamma_1)$, if $\Gamma_2 \models \phi$ then $\Gamma_1 \models \phi$.

Intuitively Γ_2 is a conservative extension of Γ_1 iff Γ_2 extends Γ_1 but tells us nothing more about sentences that are in the language of Γ_1. Γ_2 may entail sentences in its extended language of course but as far as the language which it shares with Γ_1 is concerned, it says no more than Γ_1.

Belief Revision

Belief revision aims to model the rational ways of modifying a reasoner's belief corpus when the reasoner attempts to assimilate newly acquired information. In this work we shall follow the AGM paradigm of belief change [1, 3]. The reasoner's belief corpus is represented by a *belief set*; a set of sentences closed under logical consequence Cn. Belief sets then have the following property: $\phi \in K$ if and only if K logically entails ϕ.

In the AGM paradigm, a revision function $*$ takes the reasoner's current belief set K and assimilates a new sentence ϕ, representing acquired information, and returns the revised belief set $K * \phi$. Of all the possible functions that can do this, we are only interested in those that follow some pre-imposed rationality criteria, of which the principle of minimal change is one of the most widely applied. The class of rational belief change functions can be characterised by a set of *rationality postulates*. We reproduce the AGM rationality postulates for belief revision for the purposes of contrast.

$(K*1)$ $K * \phi$ is a belief set.

$(K*2)$ $\phi \in K * \phi$.

$(K*3)$ $K * \phi \subseteq K + \phi$.

$(K*4)$ If $\neg\phi \notin K$, then $K + \phi \subseteq K * \phi$.

$(K*5)$ $K * \phi = K_{\perp}$ iff $\models \neg\phi$.

$(K*6)$ If $\models \phi \leftrightarrow \psi$, then $K * \phi = K * \psi$.

$(K*7)$ $K * (\phi \wedge \psi) \subseteq (K * \phi) + \psi$.

$(K*8)$ If $\neg\psi \notin K * \phi$, then $(K * \phi) + \psi \subseteq K * (\phi \wedge \psi)$.

The first postulate says that if the reasoner's initial belief corpus is a belief set (i.e., deductively closed) then it's revised corpus should also be a belief set. $(K*2)$ dictates that the new information be assimilated into the revised belief corpus. Postulate $(K*3)$ says that a revised corpus should be no larger than purely expanding the corpus by the new information while $(K*4)$ says that when the new information is consistent with the new information, the revision and expansion of the corpus amount to the same thing (when $(K*3)$ and $(K*4)$ are considered together). $(K*5)$ says that the revised corpus is inconsistent exactly when the new information itself is inconsistent. $(K*6)$ specifies that the syntactic form of the new information is irrelevant to the revised belief corpus. Postulates $(K*7)$ and $(K*8)$ are generalisations of $(K*3)$ and $(K*4)$ for conjunctive information.

One way of constructing a function satisfying these postulates is to place an ordering over logical interpretations [6].

$$|K * \phi|_{\mathcal{L}} = \min_{\leq_{SOS}} \{\omega \in M \mid \omega \models \phi\} \tag{i}$$

4. Conservative Belief Change

While the AGM approach is quite flexible, its rationality postulates do not allow for all interesting types of belief change. In this section we introduce the notion of conservative belief change in which the principle of minimal change can be viewed as applying to the newly acquired information. Once the reasoner acquires new information and assimilates it into its belief corpus, all that the reasoner believes about the new information in its revised corpus is the new information itself. As we have seen above, if the reasoner were to undergo the revision $K * (p \vee q)$ then all it would come to believe in its revised corpus about p and q is $p \vee q$. Of course, what it believes regarding other propositions in its language may be affected in other ways.

We shall now see how this can be achieved in an AGM-like formal setting.

Conservative Belief Revision

The type of belief change we are interested in is motivated by one of Grice's maxims of conversational implicature: when interpreting a speaker, assume that they have conveyed all the information they know regarding the subject matter. In our approach this means that when the reasoner receives new information ϕ, they revise their belief corpus so as to believe ϕ but not to believe any sentence, in the language of ϕ, that is implied by ϕ. Therefore, by way of illustration, if the reasoner revises their belief corpus to assimilate $(p \vee q) \wedge r$, their revised corpus includes precisely $(p \vee q) \wedge r$. This can be viewed as a constraint on the truth values of the atoms p, q and r. In particular, stronger versions of the sentence $(p \vee q) \wedge r$ such as p or $p \wedge r$ will not be included in the reasoner's revised corpus. This holds true even when the reasoner's initial belief corpus includes p or $p \wedge r$, in which case the revised corpus represents a weakening of the initial corpus. We refer to this operation as conservative belief revision or C-revision for short and denote it $\hat{*}$.

The semantics for C-revision is illustrated in Figure 1. Here we consider a language generated by the propositional atoms x, y and z. Initially, the reasoner's belief corpus contains $x \wedge \neg y \wedge z$. They acquire new information $\neg x \vee \neg y$. The set of interpretations is partitioned into four cells; each one corresponding to an interpretation over the language $\mathcal{L}(\neg x \vee \neg y)$. From each cell, the 'best' interpretation is chosen as part of the revised belief corpus. Relative to the language $\mathcal{L}(\neg x \vee \neg y)$, the revised corpus will be precisely $\neg x \vee \neg y$. Beliefs regarding z will depend on extralogical factors, determined by the plausibility of different worlds.

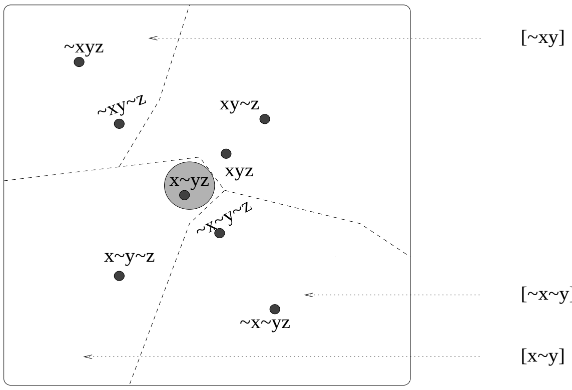

Figure 1. Partitioning the worlds according to evidential language $\mathcal{L}(x \vee y)$

To fully determine the semantics for C-revision we now impose an ordering over all interpretations of $\mathcal{L}$ signifying their level of 'plausibility'. This ordering has exactly the properties of Grove's systems of spheres construction for AGM revision [6]. That is, it is a total pre-order over all interpretations of the language in which the models of the belief set K are minimal under this ordering.[3] Figure 2 illustrates the system of spheres model applied to C-revision. The concentric 'rings' represent the total preorder over interpretations with the inner rings (or spheres) considered more plausible than the outer ones. This figure also shows the partitions from the previous example; the most plausible interpretations being chosen from each partition. Therefore, from the $[x\neg y]$ partition, interpretation $x\neg yz$ is selected while interpretation $\neg xy\neg z$ is chosen from the $[\neg xy]$ partitions and $\neg x\neg y\neg z$ from the $[\neg x\neg y]$ partitions.[4] It is evident from the selected interpretations $\{x\neg yz,\ \neg xy\neg z,\ \neg x\neg y\neg z\}$ that, as far as the atomic propositions x and y are concerned, $\neg x \vee \neg y$ is believed in the revised corpus. However, the reasoner is now indifferent to z since there are selected interpretations in which z is true and others in which $\neg z$ is true. The reasoner's attitude towards z in its revised belief corpus can be gleaned from beliefs $x \leftrightarrow z$ and $y \vee z$.

[3]The *Limit Assumption* applied by Grove is not necessary here as we deal with a finitary language.

[4]Of course, it is possible that more than one interpretation is selected from a given partition.

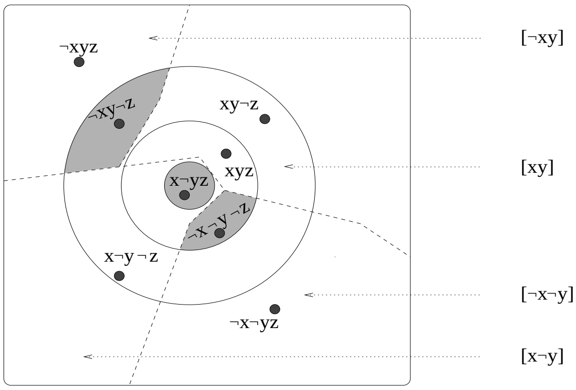

Figure 2. Conservative Revision – Semantics

The semantic condition for determining C-revision, as described above, can be formalised in a manner similar to Grove's condition for AGM revision.

$$|K \hat{*} \phi|_{\mathcal{L}} = \bigcup_{\sigma \in |\phi|_{\mathcal{L}(\phi)}} \min_{\leq_{SOS}} \{\omega \in M \mid \omega \models \sigma\}. \tag{ii}$$

The following significant result now follows.

THEOREM 4.1. *For any belief set K and input sentence ϕ, $K \hat{*} \phi$ is a* conservative extension *of ϕ, i.e., for $\psi \in \mathcal{L}(\phi)$, if $K \hat{*} \phi \models \psi$ then $\phi \models \psi$.*

In other words, $K \hat{*} \phi$ tells us no more about sentences in the language of ϕ ($\mathcal{L}(\phi)$) than ϕ does.

We also have the following results that relate C-revision to AGM revision.

THEOREM 4.2. *Let $\hat{*}$ be obtained from a systems of spheres $\leq_{SOS}$ and let $*$ be the AGM revision obtained from $\leq_{SOS}$.*

1. $K \hat{*} \phi \subseteq K * \phi$.
2. $K \hat{*} A = K * A$.[5]

These results lead to the question of whether it is possible to determine a specific C-revision operation via Grove's standard definition of revision (i) using some suitably constructed system of spheres. In other words, can C-revision be reduced to AGM revision? In general the answer turns out to be 'no'. A counterexample is easily obtained. Consider $\mathcal{L} = \{p, q\}$ and

[5] Recall that formulas A, B, ..., are conjunctions of literals by convention.

a C-revision function such that, given $K \equiv \neg p \wedge \neg q$, $K \hat{*} p = K \hat{*} (p \wedge q)$. This entails the following constraints on the ordering: $\{\neg p, \neg q\} < \{p, q\}$, $<$ $\{p, \neg q\}$. However, it is straightforward to verify that $K \hat{*} (p \vee q) \equiv p \vee q$. This cannot be obtained by standard revision which requires $\{p, q\}$, $\{\neg p, q\}$ and $\{p, \neg q\}$ to be at the same level of the system of spheres ordering, which is in conflict with the ordering above.

A system of spheres, then, leads to a unique C-revision via (ii). The following example shows that the converse situation does not hold in general.

EXAMPLE 4.3. Consider two SOS's: SOS_1: $\ldots < xyz < x\neg y\neg z$ and SOS_2: $\ldots < x\neg y\neg z < xyz$, where the $\ldots$ in the orderings represent an identical subsequence. The C-revision based on these SOS's (using (ii)) exhibit identical behaviour since no cell of any partition based on a sub-language of $\{x, y, z\}$ will pick up exactly the set $\{xyz, x\neg y\neg z\}$.

For a given AGM revision function $*$, if we have a fixed belief set K, we have a unique system of spheres. This is not the case with a given C-revision function $\hat{*}$ as it corresponds to a class of systems of spheres for a belief set K. We can however characterise the class of system of spheres that a particular C-revision function $\hat{*}$ determines.

DEFINITION 4.4. Two systems of spheres, $\leq_1$ and $\leq_2$ are $\hat{*}$-equivalent iff for every sentence $\phi \in \mathcal{L}$, $K \hat{*}_{\leq_1} \phi = K \hat{*}_{\leq_2} \phi$, where $|K|$ is the set of $\leq_{\{1,2\}}$-minimal[6] worlds and $\hat{*}_{\leq_1}$ and $\hat{*}_{\leq_2}$ are defined from $\leq_1$ and $\leq_2$ using (ii).

We want to be able to say when two systems of spheres are $\hat{*}$-equivalent.

DEFINITION 4.5. Let $\leq$ be a given SOS. We say an SOS $\leq'$ is a C-transform (conservative transform) of $\leq$ iff the former can be constructed from the latter in the following manner: (1) Consider any two worlds ω and ω'. If there is a consistent set S of literals over $\mathcal{L}$ such that both $\omega \models \bigwedge(S)$ and $\omega' \models \bigwedge(S)$, and ω is $\leq$-minimal among all worlds satisfying $\bigwedge(S)$, then $\omega \leq \omega'$ iff $\omega \leq' \omega'$ (note that since $\bigwedge(\omega) \equiv \omega$ we obtain a reflexive $\leq'$); and (2) After obtaining all those constraints on $\leq'$, we complete it as we wish to get a total preorder $\leq'$.

It is straightforward to show that C-transformation is symmetric. An example illustrates the idea.

[6] The $\leq_1$-minimal worlds and the $\leq_2$-minimal worlds are identical and equal to $|K|$ otherwise they are not appropriate for revising K.

EXAMPLE 4.6. Assume a language based on atoms $\{p, q, r\}$. Let $\leq$ be: $\{\neg p\neg q\neg r, \neg p\neg qr\} < \{\neg pq\neg r, \neg pqr\} < \{p\neg q\neg r\} < \{p\neg qr\} < \{pq\neg r, pqr\}$. If we compare worlds $pq\neg r$ and $p\neg qr$, the only relevant conjuncts are p and $\top$. Since neither of these worlds are $\leq$-minimal either in $|\top|$ (all worlds) or $|p|$ (worlds satisfying p), no particular constraint on $\leq'$ is generated by this comparison. On the other hand, if we compare $pq\neg r$ and $\neg pqr$, we notice that the relevant conjuncts are q and $\top$. Since, among worlds satisfying q, we have $\neg pqr$ as one of the $\leq$-minimal elements, and also $\neg pqr < pq\neg r$ it follows that $\neg pqr <' pq\neg r$

The following result shows the relationship between C-transformation and $\hat{*}$-equivalence.

THEOREM 4.7. *Two preorders $\leq$ and $\leq'$ are C-transforms of each other iff they are $\hat{*}$-equivalent.*

Postulates[7]

We now supply postulates for C-revision functions $\hat{*} : 2^{\mathcal{L}} \times \mathcal{L} \to 2^{\mathcal{L}}$. It will be convenient to remind the reader at this point of our linguistic convention that upper-case Roman characters A, B, ... denote *consistent conjunctions of literals* while lower-case Greek letters ϕ, ψ, χ, ... denote arbitrary sentences in $\mathcal{L}$.

($K\hat{*}1$) $K\hat{*}\phi$ is a belief set

($K\hat{*}2$) $\phi \in K\hat{*}\phi$

($K\hat{*}3$) $K\hat{*}A \subseteq K + A$

($K\hat{*}4$) If $\neg A \notin K$, then $K + A \subseteq K\hat{*}A$

($K\hat{*}5$) $K\hat{*}\phi = K_\perp$ iff $\models \neg\phi$.

($K\hat{*}6$) If $\models \phi \leftrightarrow \psi$, then $K\hat{*}\phi = K\hat{*}\psi$

($K\hat{*}7$) $K\hat{*}(A \wedge B) \subseteq (K\hat{*}A) + B$

($K\hat{*}8$) If $\neg B \notin K\hat{*}A$, then $(K\hat{*}A) + B \subseteq K\hat{*}(A \wedge B)$.

($K\hat{*}9$) If $\phi \not\models \perp$, then there is an $A \not\models \perp$ such that, $A \models \phi$ and for all B, $\neg\phi \notin K\hat{*}B$ implies $A \wedge B \models K\hat{*}B$

[7] We would like to acknowledge the assistance of Pavlos Peppas in formulating these postulates and the accompanying representaion theorem.

$(K \hat{*} 10)$ If $A \models \phi$ and $\mathcal{L}(A) \subseteq \mathcal{L}(\phi)$ then $K \hat{*} \phi \subseteq K \hat{*} A$.

$(K \hat{*} 11)$ If $\neg A \notin K \hat{*} \phi$, then there is a C such that $C \models \phi$ and $\neg A \notin K \hat{*} C$, and $\mathcal{L}(C) \subseteq \mathcal{L}(\phi)$.

Notice that we don't have the AGM version of $(K*4)$. Our previous example illustrates this nicely where $K = x \wedge \neg y \wedge z$. $K + (\neg x \vee \neg y) = K = x \wedge \neg y \wedge z$. However $K \hat{*} (\neg x \vee \neg y) = (x \leftrightarrow z) \wedge (y \vee z)$ Postulates $(K \hat{*} 3)$ – $(K \hat{*} 8)$ are simply the standard AGM postulates for belief revision with the restriction that $(K \hat{*} 3)$, $(K \hat{*} 4)$, $(K \hat{*} 7)$ and $(K \hat{*} 8)$ are restricted to consistent conjunctions of literals. However, it is also possible to show the following.

LEMMA 4.8. The AGM postulates $(K * 3)$ and $(K * 7)$ follow from the conservative revision postulates $(K \hat{*} 1)$ – $(K \hat{*} 11)$.

$(K \hat{*} 9)$ makes up for the weakening of the AGM postulates; for every consistent sentence ϕ, a stronger consistent conjunction of literals A exists that is capable of capturing any other C-revision by a conjunction of literals. Postulate $(K \hat{*} 10)$ stipulates that if $A \models \phi$, the only way for $K \hat{*} \phi$ not to be included in $K \hat{*} A$ is when the minimum language of A is not in the minimal language of ϕ. $(K \hat{*} 11)$ can be viewed as the converse of $(K \hat{*} 10)$. It says that when a consistent conjunction of literals A is in $K \hat{*} C$ where C is a consistent conjunction of literals (a prime implicant of ϕ, in fact) then A is in $K \hat{*} \phi$. Postulates $(K \hat{*} 10)$ and $(K \hat{*} 11)$ can be rephrased in terms of prime implicants as follows.

$(K \hat{*} 10')$ If $A \models \phi$ and $K \hat{*} \phi \not\subseteq K \hat{*} A$, then there is a literal L such that $A \models L$, and L is neither entailed nor contradicted by any prime implicant of ϕ.

$(K \hat{*} 11')$ If $\neg A \notin K \hat{*} \phi$, then $\exists C$ such that $C \models \phi$ and $\neg A \notin K \hat{*} C$, and for all literals L, if $C \models L$, then L is either entailed or contradicted by a prime implicant of ϕ.

The following postulate is also equivalent to $(K \hat{*} 11)$ in the presence of $(K \hat{*} 1)$ – $(K \hat{*} 10)$.

$(K \hat{*} 12)$ $K \hat{*} A$ is the largest theory satisfying postulates $(K \hat{*} 1)$ – $(K \hat{*} 10)$.

The soundness of the C-revision postulates can be shown quite straightforwardly.

THEOREM 4.9. *Let K be a theory and $\leq_{SOS}$ a system of spheres centered at $|K|_{\mathcal{L}}$. The function $\hat{*}$ induced from $\leq_{SOS}$ via (ii) satisfies* $(K\hat{*}1)$ – $(K\hat{*}11)$.

Completeness of the C-revision postulates is achieve in two stages: (i) the special case of consistent conjunctions of literals, where C-revision reduces to classical AGM revision, and we prove the completeness of $(K\hat{*}1)$ – $(K\hat{*}9)$, and (ii) extending the results to arbitrary sentences and include $(K\hat{*}10)$ and $(K\hat{*}11)$. More precisely, we have:

THEOREM 4.10.

1. *Let K be a theory and $\hat{*}$ a revision function satisfying* $(K\hat{*}1)$ – $(K\hat{*}9)$. *There exists a system of spheres $\leq_{SOS}$ centered on $|K|_{\mathcal{L}}$, such that for any consistent conjunction of literals A, $|K\hat{*}A|_{\mathcal{L}} = \min_{\leq_{SOS}}\{\omega \in M \mid \omega \models A\}$*
2. *Let K be a theory and $\hat{*}$ a revision function satisfying* $(K\hat{*}1)$ – $(K\hat{*}11)$. *There exists a system of spheres $\leq_{SOS}$ centered on $|K|_{\mathcal{L}}$, such that $\hat{*}$ is identical to the C-revision function induced from $\leq_{SOS}$.*

Conservative Belief Update

We can repeat the preceding development for belief update, using the formulation of [9]. For our purposes, these authors associate a well-founded partial preorder with each interpretation ω, such that ω is the unique minimum in the preorder.[8] The update of K by ϕ, denoted $K \diamond \phi$, is defined by:

$$Mod_{\mathcal{L}}(K \diamond \phi) = \bigcup_{\sigma \in Mod(K)} \min_{\leq_\sigma}\{\omega \in M \mid \omega \models \phi\}. \tag{iii}$$

Our definition for C-update is as follows:

$$Mod_{\mathcal{L}}(K \hat{\diamond} \phi) = \bigcup_{\omega_1 \in Mod(K)} \bigcup_{\sigma \in Mod_{\mathcal{L}(\phi)}\phi} \min_{\leq_{\omega_1}}\{\omega \in M \mid \omega \models \sigma\}. \tag{iv}$$

We obtain the following postulates.

THEOREM 4.11. *Let K be a belief set, ϕ, $\psi \in \mathcal{L}$ and let $\hat{\diamond}$ be defined via (iv), then $\hat{*}$ satisfies:*

[8] There are two differences with the Grove system of spheres: each interpretation has its own preorder, and the ordering is partial rather than total.

($K \hat{\diamond} 0$) *$K \hat{\diamond} \phi$ is a belief set.*

($K \hat{\diamond} 1$) *$\phi \in K \hat{\diamond} \phi$.*

($K \hat{\diamond} 2$) *If $\phi \in K$ then $K \hat{\diamond} \phi = K$.*

($K \hat{\diamond} 3$) *If $K \not\models \bot$ and $\phi \not\models \bot$ then $K \hat{\diamond} \phi \neq K_{\bot}$.*

($K \hat{\diamond} 4$) *If $\models \phi \leftrightarrow \psi$, then $K \hat{\diamond} \phi = K \hat{\diamond} \psi$.*

($K \hat{\diamond} 5$) *If $\psi \in K \hat{\diamond} \phi$ then $K \hat{\diamond} \phi = K \hat{\diamond} (\phi \wedge \psi)$.*

($K \hat{\diamond} 7$) *If K is complete then $K \hat{\diamond} (\phi \vee \psi) \subseteq K \hat{\diamond} \phi \cup K \hat{\diamond} \psi$.*

($K \hat{\diamond} 8$) *$(K_1 \cap K_2) \hat{\diamond} \phi = (K_1 \hat{\diamond} \phi) \cap (K_2 \hat{\diamond} \phi)$.*

($K \hat{\diamond} 9$) *If $\psi \in \mathcal{L}(\phi)$ then $K \hat{\diamond} \phi \models \psi$ implies $\phi \models \psi$.*

Again ($K \hat{\diamond} 9$) states $K \hat{\diamond} \phi$ is a conservative extension of ϕ.

Note that update postulate ($K \diamond 6$), is a straightforward consequence of ($K \hat{\diamond} 5$). As well, ($K \diamond 5$) does not hold for C-update. A counterexample is given by that for ($K * 7$).

Context-Dependent Revision

The nature of C-revision dictates that all conjunctions of literals corresponding to models of ϕ over the language $\mathcal{L}(\phi)$ are satisfiable given the C-revised belief corpus $K \hat{*} \phi$. The language $\mathcal{L}(\phi)$ can be viewed as a context for ϕ. This notion can be extended to a context C where $\mathcal{L}(\phi) \subseteq C \subseteq \mathbf{P}$. The resulting revision operation $K \hat{*} (C, \phi)$ is a revision where ϕ is all that is believed about the context C after revision.

To motivate the notion of context sensitive revision, consider the following example.

Example 4.12. Assume a belief set represented by $\{a \wedge b \wedge c \wedge d\}$, where a, b and c stand respectively for Albert, Becky and Charles being involved in a bank robbery, and d stands for Doug being a geologist. The input $a \vee b$ suggests that the context is the relevant bank robbery, represented by $\{a, b, c\}$. We would expect that after the revision, we would no longer suspect Charles of robbery; and whether or not we would still believe that Doug is a geologist would depend on extraneous factors. On the other hand, if we let c stand for Charles being a nice dad (and a, b and d as before) we would expect that the context is simply $\{a, b\}$, and as a result of the revision, whether or not c would be maintained will depend on extraneous factors.

A semantic construction like that for C-revision in (ii) can be provided for context-dependent revision. For $K \subseteq \mathcal{L}$, $\phi \in \mathcal{L}$, and $\mathcal{L}(\phi) \subseteq C \subseteq \mathbf{P}$ define:

$$|K \mathbin{\hat{*}} (C, \phi)|_{\mathcal{L}} = \bigcup_{\sigma \in |\phi|_C} \min_{\leq_{SOS}} \{\omega \in M \mid \omega \models \sigma\}. \tag{v}$$

Again, we can prove that $K \mathbin{\hat{*}} (C, \phi)$ is a conservative extension of C where ϕ holds.

THEOREM 4.13. *For any belief set K, context C, and input sentence ϕ, $K \mathbin{\hat{*}} (C, \phi)$ is a* conservative extension *of C in which ϕ is true.*

Furthermore, we have the following results about context-dependent revision.

THEOREM 4.14. *Let $\hat{*}$ and $*$ (representing AGM revision) be obtained from a system of spheres $\leq_{SOS}$.*

1. $K \mathbin{\hat{*}} (C, \phi) \subseteq K * \phi$.
2. *If* $\mathcal{L}(A) = C$ *then* $K \mathbin{\hat{*}} (C, A) = K * A$.
3. $K \mathbin{\hat{*}} (\mathbf{P}, \phi) \equiv Cn(\phi)$.

The first result states that whatever is believed in context-dependent revision will be believed in AGM revision. The second shows that if the context corresponds to a conjunction of literals, C-revision coincides with AGM revision. The last result states that when the context is the entire set of atoms, context-dependent revision coincides with AGM full-meet revision.

By making the context the entire language of the input sentence, we can define C-revision in terms of context-dependent revision, that is, $K \mathbin{\hat{*}} \phi \equiv K \mathbin{\hat{*}} (\mathcal{L}(\phi), \phi)$. Context-dependent revision, on the other hand, can be defined in terms of C-revision as follows. $|K \mathbin{\hat{*}} (C, \phi)|_{\mathcal{L}} = \bigcup_{A \in |\phi|_C} |K \mathbin{\hat{*}} A|_{\mathcal{L}}$. A context-dependent version of belief update can be constructed in a similar manner.

5. Discussion

When viewed through their semantic constructions, the difference between AGM revision and C-revision can be seen in similar light to the distinction between belief revision and belief update. In AGM revision, the models of K are considered as a whole; revision is determined by taking the 'closest' (i.e., minimal) models of ϕ to the models of K in a system of spheres. In belief update, each model of K is considered separately; for each we take the closest ϕ models and collect them together to characterise the updated belief corpus. In C-revision, we consider each model of ϕ with respect to the language of ϕ (i.e., $\mathcal{L}(\phi)$) separately and use it to revise K as usual, collecting the resulting models to characterise the C-revised belief corpus. This gives a more complete classification of belief change operations in terms of whether the models of the belief corpus are considered as a whole (revision) or individually (update) or whether the models of the new information are considered as a whole (revision and update) or separately (C-revision and C-update). C-revision also corresponds to a Gricean view of how the new information should be treated during the belief change process.

C-revision also provides a way of dealing with a problem with the contentious *recovery postulate* in AGM belief contraction

$$K \subseteq (K \dot{-} \phi) + \phi$$

This postulate is manifestation of the principle of minimal change and says that when contracting and then expanding by the same sentence, no beliefs should be lost. Hansson [7] gives the following counterintuitive example.

Example 5.1 (Paraphrased from Hansson [7]). Let K entail that "Cleopatra had a son and a daughter" (say, $s \wedge d$). New information is received that Cleopatra didn't have a child, expressed by $K \dot{-} (s \vee d)$. Then one learns that she had a child, expressed by $(K \dot{-} (s \vee d)) + (s \vee d)$. Recovery says that $s \wedge d$ is believed – that Cleopatra had a son and daughter. Intuitively it seems that just $s \vee d$ should be believed, that all that is known after these changes is that she had a child.

Firstly, if $K \models \phi$ then recovery can be written in terms of revision rather than expansion. $(K \dot{-} \phi) + \phi \ = \ (K \dot{-} \phi) * \phi$. The example above can be better captured using conservative revision; concerning $\{s, d\}$, we are informed *at most* that Cleopatra has a child $s \vee d$. Using C-revision we have possibly $K \not\subseteq (K \dot{-} \phi) \hat{*} \phi$. In the example, if $K \equiv (s \wedge d)$, then we would expect that $K \dot{-} (s \vee d) \equiv \neg s \wedge \neg d$ and so $(K \dot{-} (s \vee d)) \hat{*} (s \vee d) \equiv (s \vee d)$, in accord with intuitions.

AGM belief contraction has also attracted criticism for removing too little information. Makinson [11] shows that when we consider the class of functions that do not necessarily satisfy the recovery postulate (known as withdrawal functions) and partition them into equivalence classes of functions that produce the same revision behaviour under the Levi Identity: $K * \phi = (K \dot{-} \neg\phi) + \phi)$, an AGM revision function is the maximal element of each equivalence class in the sense that it removes the fewest beliefs with respect to set inclusion. A proposal by Rott and Pagnucco for a more severe form of withdrawal [13] has been similarly criticised for removing too much information. Hansson [7] suggests that belief contraction should lie between these two extremes. If we were to use a C-revision operator $\hat{*}$ to define a *C-contraction* operation via the Harper Identity: $K \hat{-} \phi = K \cap (K \hat{*} \neg\phi)$ we would obtain such a form of contraction.

6. Conclusion

We have presented a novel form of belief change that focuses on the content of the new information acquired by the reasoner and to be used during the belief change process. It is motivated by one of Grice's maxims of conversational implicature. Since it leads to a conservative extension of the newly acquired information, we term this conservative belief change. One of the more interesting features of our proposal is that it gives an added dimension to the classification of belief change operations in a way that is orthogonal to the usual distinction between revision and update. Conservative revision gives a way of dealing with a well known problem with the recovery postulate. C-revision also resolves a problem pointed out by Herzig and Rifi [8] due to disjunction and does so in a syntax independent manner.

We have given a semantic construction and postulates for C-revision, and have proven completeness for these postulates with respect to this semantics. However we have not provided a complete set of postulates for C-update, nor for the context-dependent versions of C-revision and for C-update. Furthermore, as noted above, C-contraction (and its context-dependent versions) is a good candidate for a viable belief contraction operation. A proper account of such operations is a good opportunity for future work.

Acknowledgements

The authors would especially like to thank Pavlos Peppas for his help with some of the details in this paper. They would also like to thank the anonymous referees and participants of the Fifth International Workshop on Nonmonotonic Reasoning and Change, 2003 for feedback on an earlier version of this work.

References

[1] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50:510–530, 1985.

[2] Mukesh Dalal. Investigations into a theory of knowledge base revision: Preliminary report. In *Proceedings of the Seventh National Conference on Artificial Intelligence*, pages 449–479, St. Paul, MN, 1988.

[3] Peter Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. Bradford Books, MIT Press, Cambridge Massachusetts, 1988.

[4] Peter Gärdenfors and David Makinson. Revisions of knowledge systems using epistemic entrenchment. In *Proceedings of the Second Conference on Theoretical Aspect of Reasoning About Knowledge*, pages 83–96, 1988.

[5] H. P. Grice. *Studies in the Way of Words*. Harvard University Press, Cambridge Massachusetts, 1989.

[6] Adam Grove. Two modellings for theory change. *Journal of Philosophical Logic*, 17:157–170, 1988.

[7] Sven Ove Hansson. A textbook of belief dynamics: Theory change and database updating. Kluwer Academic Publishers, 1999.

[8] Andreas Herzig and Omar Rifi. Propositional belief base update and minimal change. *Artificial Intelligence*, 115:107–138, 1999.

[9] Hirofumi Katsuno and Alberto O. Mendelzon. Propositional knowledge base revision and minimal change. *Artificial Intelligence*, 52:263–294, 1991.

[10] David Lewis. *Counterfactuals*. Harvard University Press, Cambridge, Massachusetts, 1973.

[11] David Makinson. On the status of the postulate of recovery in the logic of theory change. *Journal of Philosophical Logic*, 16:383–394, 1987.

[12] David Makinson. Five faces of minimality. *Studia Logica*, 52:339–379, 1993.

[13] H. Rott and M. Pagnucco. Severe withdrawal (and recovery). *Journal of Philosophical Logic*, 28(5), 1999.

[14] Hans Rott. Two dogmas of belief revision. *Journal of Philosophy*, 97:503–522, 2000.

[15] Marianne Winslett. *Updating Logical Databases*. Cambridge University Press, 1990.

James P. Delgrande
Department of Computing Science
Simon Fraser University
Burnaby AC, Canada V5A 1S6
`jim@cs.sfu.ca`

Abhaya C. Nayak
Department of Computing
Division of Information and Communication Sciences
Macquarie University
North Ryde, NSW 2109, Austrlia
`abhaya@ics.mq.edu.au`

Maurice Pagnucco
National ICT Australia
The University of New South Wales
Sydney, NSW 2052, Australia
Maurice.Pagnucco@nicta.com.au